



МОНИТОРИНГ ОБЩЕСТВЕННОГО МНЕНИЯ

ЭКОНОМИЧЕСКИЕ И СОЦИАЛЬНЫЕ ПЕРЕМЕНЫ

№ 1 (161)

январь — февраль 2021

ЦИФРОВИЗАЦИЯ

МЕТОДЫ
И МЕТОДОЛОГИЯ



СОЦИОЛОГИЯ
ТРУДА



ПРАКТИКИ
ГОВОРЯТ

18+

ISSN 2219-5467



9 772219 546006 >

Главный редактор журнала:
Федоров Валерий Валерьевич —
кандидат политических наук, генеральный директор ВЦИОМ,
профессор НИУ ВШЭ

Заместители главного редактора:
Седова Наталья Николаевна —
помощник гендиректора по науке ВЦИОМ
Подвойский Денис Глебович —
кандидат философских наук, ведущий научный сотрудник
Института социологии ФНИСЦ РАН, доцент РУДН

Ответственный редактор:
Кулешова Анна Викторовна —
кандидат социологических наук, член российской
Ассоциации научных редакторов и издателей (АНРИ) (Россия)

M77 Мониторинг общественного мнения: экономические и социальные перемены. — М.: АО «ВЦИОМ», 2021. — № 1 (161). — 409 с.

ISSN 2219-5467

Объективная, точная, регулярная и свежая информация «Мониторинга» полезна всем, кто принимает управленческие решения, занимается прогнозированием и анализом развития общества. Наш журнал пригодится сотрудникам научных и аналитических центров, работникам органов управления, ученым, преподавателям, молодым исследователям, студентам и аспирантам, журналистам.

Тематика материалов охватывает широкий круг социальных, экономических, политических вопросов, основные рубрики посвящены теории, методам и методологии социологических исследований, вопросам взаимодействия государства и общества, социальной диагностике. Каждый номер журнала содержит двухмесячный дайджест основных результатов еженедельных общероссийских опросов ВЦИОМ.

Мы публикуем статьи специалистов, представляющих ведущие научные социологические центры, институты, организации, а также ВУзы России и зарубежных стран. Широкая тематика журнала представляет возможность выступить на его страницах представителям смежных специальностей (политологам, историкам, экономистам и т.д.), опирающимся в своих исследованиях на эмпирические социологические данные.

Журнал издается с 1992 г.

СОДЕРЖАНИЕ

ВСТУПИТЕЛЬНОЕ СЛОВО ПРИГЛАШЕННОГО РЕДАКТОРА

А. В. Резаев, Н. Д. Трегубова

- Искусственный интеллект и искусственная социальность:
новые явления, проблемы и задачи для социальных наук 4

ТЕОРИЯ И ИСТОРИЯ

A. V. Rezaev

- Twelve Theses on Artificial Intelligence and Artificial Sociality 20

A. Karatzogianni

- Research Design for an Integrated Artificial Intelligence Ethical Framework 31

N. Klowait, M. A. Erofeeva

- The Rise of Interactional Multimodality in Human-Computer Interaction 46

СОЦИОЛОГИЯ ТРУДА, ОРГАНИЗАЦИЙ И ПРОФЕССИИ

N. Hosogaya

- Development of Telework and its Effects on Japan's Human Resource Management:
Artificial Sociality as a Focus 71

A. Tadesse, W. R. Allen, C. Mitchell-Kernan

- Integrating Educational Technology in East Africa: One Size Does Not Fit All 91

МОНИТОРИНГ МНЕНИЙ

- Мониторинг мнений: январь — февраль 2021 109

ЦИФРОВИЗАЦИЯ

А. В. Смирнов

- Цифровое общество: теоретическая модель
и российская действительность 129

А. П. Климович

- Риски концентрации власти в эпоху искусственного интеллекта:
между следящим капитализмом и цифровым Левиафаном 154

Т. С. Мартыненко, Д. Е. Добринская

- Социальное неравенство в эпоху искусственного интеллекта:
от цифрового к алгоритмическому разрыву 171

<i>А. М. Корбут</i>	
Одомашнивание искусственного интеллекта: умные колонки и трансформация повседневной жизни.....	193
<i>А. М. Давыдова, М. А. Солянова, К. Соренсен</i>	
Дисциплинарные практики цифрового селф-трекинга: между эмансиацией и контролем	217
<i>А. И. Егорова, Н. Кловайт</i>	
Как попрощаться с роботом? Исследование достижения разговорного закрытия	241

МЕЖДИСЦИПЛИНАРНЫЕ ИССЛЕДОВАНИЯ

<i>И. А. Блохин, С. П. Морозов, В. Ю. Чернина, А. Е. Андрейченко, И. В. Шахабов, А. В. Смышляев, В. А. Гомболовский</i>	
Использование искусственного интеллекта в здравоохранении: опыт валидации алгоритма искусственного интеллекта в медицинских организациях в условиях пандемии COVID-19	271

МЕТОДЫ И МЕТОДОЛОГИЯ

<i>Н. Д. Трегубова, М. Л. Ни, А. А. Китаева</i>	
Транснационализм в режиме онлайн: миграционные исследования в условиях искусственной социальности.....	283
<i>М. Б. Богданов, И. Б. Смирнов</i>	
Возможности и ограничения цифровых следов и методов машинного обучения в социологии.....	304
<i>М. Ю. Александрова</i>	
Методы машинного обучения в социологическом исследовании: предсказание частичного неответа с использованием наивного байесовского классификатора	329
<i>Н. В. Ярцева</i>	
Анализ больших объемов данных: возможности Gdelt Project при использовании языка программирования Python. Опыт гуманитария, решившего постичь Big Data.....	351

ПРАКТИКИ ГОВОРЯТ

<i>Т. А. Нестик, Н. Н. Седова, Е. Г. Климанова</i>	
Будущее исследовательской индустрии: от конкуренции за бюджеты к поиску партнеров.....	368

РЕЦЕНЗИЯ НА КНИГУ

А. А. Иванова

Взаимоотношения «человек — алгоритм — человек»
как социальная проблема и проблема для социологов.

Рец. на кн.: Cathy O'Neil. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. New York: Crown Publishers, 2016 387

D. M. Zhikharevich

How Algorithms Became Capitalist.

Book Review of: Zuboff Sh. *Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. New York: PublicAffairs, 2019..... 398

ВСТУПИТЕЛЬНОЕ СЛОВО ПРИГЛАШЕННОГО РЕДАКТОРА

DOI: [10.14515/monitoring.2021.1.1905](https://doi.org/10.14515/monitoring.2021.1.1905)



А. В. Резаев, Н. Д. Трегубова

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ И ИСКУССТВЕННАЯ СОЦИАЛЬНОСТЬ: НОВЫЕ ЯВЛЕНИЯ, ПРОБЛЕМЫ И ЗАДАЧИ ДЛЯ СОЦИАЛЬНЫХ НАУК

Правильная ссылка на статью:

Резаев А. В., Трегубова Н. Д. Искусственный интеллект и искусственная социальность: новые явления, проблемы и задачи для социальных наук // Мониторинг общественного мнения: экономические и социальные перемены. 2021. № 1. С. 4—19. <https://doi.org/10.14515/monitoring.2021.1.1905>.

For citation:

Rezaev A. V., Tregubova N. D. (2021) Artificial Intelligence and Artificial Sociality: New Phenomena and Challenges for the Social Sciences. *Monitoring of Public Opinion: Economic and Social Changes*. No. 1. P. 4–19. <https://doi.org/10.14515/monitoring.2021.1.1905>. (In Russ.)

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ И ИСКУССТВЕННАЯ СОЦИАЛЬНОСТЬ: НОВЫЕ ЯВЛЕНИЯ, ПРОБЛЕМЫ И ЗАДАЧИ ДЛЯ СОЦИАЛЬНЫХ НАУК

РЕЗАЕВ Андрей Владимирович — доктор философских наук, профессор, руководитель Международной исследовательской лаборатории **ТАНДЕМ**, Санкт-Петербургский государственный университет, Санкт-Петербург, Россия
E-MAIL: rezaev@hotmail.com
<https://orcid.org/0000-0002-4245-0769>

ТРЕГУБОВА Наталья Дамировна — кандидат социологических наук, ассистент кафедры сравнительной социологии, Санкт-Петербургский государственный университет, Санкт-Петербург, Россия
E-MAIL: n.tregubova@spbu.ru
<https://orcid.org/0000-0003-3259-5566>

Благодарность. Работа выполнена при поддержке РФФИ в рамках научного проекта № 20-04-60033.

ARTIFICIAL INTELLIGENCE AND ARTIFICIAL SOCIALITY: NEW PHENOMENA AND CHALLENGES FOR THE SOCIAL SCIENCES

Andrey V. REZAEV¹ — Prof. Dr. habil., Director of International Research Laboratory **TANDEM**
E-MAIL: rezaev@hotmail.com
<https://orcid.org/0000-0002-4245-0769>

Natalia D. TREGUBOVA¹ — Cand. Sci. (Soc.), Assistant Professor, Chair of Comparative Sociology
E-MAIL: n.tregubova@spbu.ru
<https://orcid.org/0000-0003-3259-5566>

¹ Saint Petersburg State University, Saint Petersburg, Russia

Acknowledgments. The study was conducted in the framework of the research project supported by RFBR grant No. 20-04-60033.

В последние годы мы наблюдаем всплеск дискуссий и публикаций о технологиях искусственного интеллекта (ИИ), которые становятся неотъемлемой частью повседневной жизни общества. Эти дискуссии ставят ряд вопросов, которые оказываются ключевыми для современной социальной науки. На каких основаниях социальные ученые могут/должны принимать участие в экспертной оценке и в общественном обсуждении того, как обеспечить прогресс в области ИИ? Должны ли социологи опираться на абстрактные философские теории или их дело — проводить конкретную эмпирическую работу, фиксирующую текущее положение дел? Следует ли социологам присоединяться к междисциплинарным коллективам? На каком уровне целесообразно проводить критический анализ технологий ИИ? Наконец, каким образом социальные ученые реально могут включиться в дискуссии и исследовательскую работу в области ИИ, которой традиционно занимаются представители других дисциплин?

Данные вопросы становятся еще более актуальными в период пандемии COVID-19. С начала пандемии три характеристики стали решающими для социально-экономического развития современных обществ: а) возвращение опре-

деляющей роли национальному государству в большинстве его функций; б) опора на авторитет науки и на научное понимание того, как справиться с вирусом; в) необходимость включения в повседневную жизнь людей технологий ИИ и онлайн-культуры. Наиболее важными агентами ИИ для осмыслиения опыта пандемии COVID-19 являются алгоритмы, анализирующие данные о пользовательской активности в социальных медиа, реконструирующие предпочтения пользователей на основании этих данных и определяющие приоритет показа тех или иных новостей на основании этих предпочтений. Также оказываются важны агенты ИИ, которые действуют в сфере медицины в качестве диагностов и посредников между пациентами и медицинским персоналом. Наконец, появляются алгоритмы, отслеживающие контакты индивида и предупреждающие его/её о возможном заражении. Таким образом, ситуация с пандемией выявляет разнообразие агентов ИИ, включенных в повседневную жизнь общества, открывая тем самым новое поле исследований для социальных ученых.

Специальный выпуск журнала, который мы представляем читателю, посвящен обсуждению того, как социальные ученые понимают *pro et contra* в развитии искусственного интеллекта (ИИ) и искусственной социальности (ИС) в современных обществах.

Следует сделать три замечания о базовых допущениях и общем характере дискуссий, которые читатель найдет в этом номере. Во-первых, мы утверждаем, что в рамках современной социальной аналитики искусственный интеллект, как бы ученые ни трактовали данный термин, следует рассматривать в связке с искусственной социальностью — с реальностью новых форм взаимодействия между людьми и машинами/алгоритмами [Резаев, Трегубова, 2019]. Дистанция между исследованиями в области ИИ и ИС далее будет только сокращаться. Во-вторых, развитие новых технологий и онлайн-культуры сопровождается появлением взаимных «обязательств» между ИИ и ИС, а также «обязательств» ИИ перед обществом. В-третьих, мы полагаем, что сегодня человечество имеет все возможности для того, чтобы ИИ и ИС служили на благо обществу и человеку. Вопрос в том, как сделать так, чтобы не утратить эти возможности.

Статьи в данном выпуске ориентированы на концептуальный и эмпирический анализ технологий ИИ. Специальный выпуск не просто осмысляет практики, связанные с технологиями ИИ, которые явно или скрыто представлены в повседневной жизни людей, — он призывает к дальнейшим исследованиям, где большее внимание будет уделено социальным, этическим, культурным аспектам создания, распространения и использования новых технологий. Чтобы возникли новые формы и методы анализа и оценки развития ИИ, необходимо проводить новые исследования, пересекающие границы дисциплин, исследования, ориентированные на недисциплинарные подходы.

Главная цель специального выпуска проста: мы стремимся понять, что же представляет собой искусственный интеллект, продвинуться в поиске теоретических оснований для социальных ученых, интересующихся проблемами ИИ и ИС, рассмотреть характеристики, от которых другие ученые могли бы отталкиваться, развивая их или обосновывая более аргументированные положения. Иными словами, мы полагаем, что сегодня наиболее продуктивно работать с проблемами ИИ и ИС в формате дискуссии.

Из этого следует два обстоятельства, на которые нам бы хотелось обратить внимание читателя заранее.

Во-первых, многие темы неизбежно остаются за рамками данного выпуска. Поэтому мы приносим извинения читателям, ожидающим найти здесь детальный анализ того, как новые технологии влияют на экономику или как можно работать с социальными и культурными стереотипами агентов ИИ: развернутое обсуждение данных проблем выходит за рамки данного номера. Во-вторых, мы стремились включить в специальный выпуск зарубежные исследования, но достигли в этом лишь ограниченного успеха: в номер входят три статьи исследователей из Великобритании, США, и Японии.

Структура дальнейшего изложения достаточно традиционна. Мы начнем с общего обзора специальной литературы, с тем чтобы определить систему координат или то, что в английском языке называется *background*. Перейдем от него к формулировке концептуальных допущений, которые, мы полагаем, могут стать путеводной нитью для тех, кто исследует феномены ИИ и ИС. Далее обратим внимание на наиболее важные проблемы, встающие перед социальной аналитикой ИИ. Указывая на аспекты ИИ, представляющие интерес для социального знания, экономики, социологии повседневности, этики, мы надеемся дополнить существующие исследования в компьютерных науках и других дисциплинах, а также проложить дорогу кросс-дисциплинарному взаимодействию. Далее представим краткое описание материалов, включенных в выпуск. В конце мы сделаем несколько заключительных замечаний о специальном выпуске.

Обзор литературы

О различных аспектах развития технологий ИИ и их включении в общество написано уже немало. В рамках вступительной статьи, наверное, нет необходимости представлять глубокий и детальный анализ публикаций по данной проблематике. Мы отсылаем читателя к аналитическим обзорам исследований и разработок ИИ: [Boden, 2016; McCorduck, 1979; Ford, 2018]. Тем, кто ориентирован на профессиональную литературу об ИИ в системе компьютерных наук, рекомендуем фундаментальную работу С. Рассела и П. Норвига [Russel, Norvig, 2016]. Также надеемся, что данный выпуск станет в какой-то мере помощником и проводником в область исследований ИИ тем читателям журнала, для которых данная проблематика не является приоритетной или их специализацией.

В массиве академических публикаций об ИИ сегодня доминирует «большая тройка» — компьютерные науки, когнитивные науки и философия (см. [Резаев, Трегубова, 2019]). В последнее время в западной литературе (особенно в США) к «большой тройке» присоединяются исследования, организованные в рамках гуманитарных дисциплин, в особенности критической теории, которые обсуждают вопросы социальной справедливости и индивидуальной приватности в мире машин и алгоритмов.

Вместе с тем активно заявляют о своем интересе к исследованию ИИ такие уважаемые и признанные области знания, как STS и теория коммуникации. В их рамках появляются перспективные теоретико-методологические подходы [Woolgar, 1985; Collins, 2018; Lee, Larsen, 2019; Ruppert, Law, Savage, 2013; Esposito, 2017b;

Floridi, 2015] и весьма интересные эмпирические исследования (см., например, [Ziewitz, 2016; Beer, 2017; Esposito, 2017a]). При этом нельзя не отметить: каждая из областей, явно или скрыто, претендует на монополию исследования ИИ в рамках социальной науки.

В этой картине место социологии оказывается везде и нигде: везде в той мере, в какой социологи считают STS, исследования коммуникаций и критическую теорию «своими»; нигде, поскольку каждая из этих областей с успехом отмежевывается от социологии и противопоставляет себя ей. Тем не менее начиная с 1990-х годов появляются работы, трактующие проблемы ИИ и ИС с опорой на концептуальные и методологические ресурсы социологической дисциплины [Wolfe, 1993; Joerges, 1990; Carley, 1996; Коллинз, 2004; Couldry, 2015; Etzioni, Etzioni, 2017].

Корпус публикаций об ИИ может быть далее структурирован на основании двух обобщений. С одной стороны, представители компьютерных наук настаивают на «порядке и прогрессе», тогда как исследователи из гуманитарных дисциплин обсуждают «свободу и равенство». С другой стороны, теоретики сосредотачиваются на своих излюбленных теориях, а эмпирики — на том, чтобы все «работало» и подтверждалось собранными данными.

Еще один важный вопрос — соотношение исследований в академии и в рамках бизнес-организаций. Прежде всего, следует понимать, что граница между ними сегодня размыта: корпорации создают мощные исследовательские отделы, производящие в том числе научные публикации; университеты организуют программы аспирантуры и магистратуры вместе с бизнес-организациями; внутри самих университетов происходит дифференциация между коллективами, ориентированными на академическую и на прикладную науку. Причем зачастую и те, и другие считают, что их наука лучше [Hoffman, 2016]. Однако представляется, что в результате данных процессов выигрывают корпорации: в них (вследствие накопления ресурсов разного рода) развиваются и наука, и бизнес. Университеты же — за немногими исключениями (такими как Stanford и MIT), — пытаясь угнаться за корпорациями, не получают ни того, ни другого в сопоставимом объеме и качестве.

Сегодня заметным трендом в развитии исследований ИИ, который в определенной степени объединяет усилия ученых из разных дисциплин, исследователей из академии и бизнес-организаций, становится кристаллизация человеко-ориентированного подхода (*human-centered AI*). Его основная идея заключается в том, чтобы в центр развития технологий ИИ поставить благо человека и человечества, чтобы подстраивать технологии под людей, а не людей под технологии. В последние годы крупные исследовательские центры человеко-ориентированного ИИ открываются в США¹ и Европе².

Наконец, следует сказать несколько слов об исследованиях ИИ, представленных в отечественной литературе. По сути, здесь мы видим все те же тенденции развития, с поправкой на некоторое запаздывание и на местную специфику.

¹ Наиболее крупные центры: Stanford Institute for Human-Centered AI (<https://hai.stanford.edu/>) и MIT Sloan (<https://mitsloan.mit.edu/experts/human-centered-ai-how-can-technology-industry-fight-bias-machines-and-people>).

² См., например: European Network of Human-Centred Artificial Intelligence. URL: <https://www.humane-ai.eu/> (дата обращения: 20.02.2021); Human-Centred AI Lab (Holzinger Group). URL: <https://human-centered.ai/> (дата обращения: 20.02.2021).

Во-первых, в России наблюдается то же соотношение академических и бизнес-исследований в сфере ИИ. Так, крупнейшая конференция по проблематике ИИ в 2020 г. организована корпорацией³; академические мероприятия на этом фоне выглядят скромно, хотя вполне достойно⁴.

Во-вторых, в русскоязычной науке доминирует та же «большая тройка» дисциплин, оформленная еще в рамках советской науки: компьютерные науки / информатика и математика, когнитивные исследования / нейронаука, и философия.

В-третьих, в социальных науках происходит выборочная рецепция идей, подходов и методов, развиваемых в Западной Европе и Северной Америке. При этом особую популярность получают отдельные направления STS и акторно-сетевой теории [Утехин, 2012; Земнухова, 2018; Сивков, 2018; Кузнецов, 2020], а также микросоциологические подходы [Хониева, 2017; Максимова, Глазков, 2018; Кловайт, 2018; Корбут, 2018]. Кроме того, появляются публикации, обсуждающие проблематику неравенства в связи с развитием технологий ИИ [Капелюшников, 2017; Ним, 2018], и мы полагаем, что их число будет расти, следуя общему международному тренду.

Инструментальность, экзистенциальность, социальность

Важность человеко-ориентированного подхода к ИИ, противостоящего чисто технологической модернизации, заключается в его способности исследовать, ставить под вопрос и предлагать альтернативы нормативным практикам повседневного использования технологий. Иными словами, смысл включения ИИ в общество не в том, что люди получают новые инструменты, а в том, что они получают их от кого-то, для реализации конкретных целей и в рамках некоторых отношений.

Онлайн-культура⁵ — это культура плотных информационно-инструментальных обменов и, одновременно, культура экзистенциальной атомизации и одиночества. Человеческое одиночество в онлайн-активностях только усиливается.

Сформулируем следующую гипотезу: онлайн-культура переводит социальность в инструментальность, а экзистенциальность — в социальность. Онлайн-культура инструментализирует социальные взаимодействия, трактуя их как информационные обмены и как потенциальный источник прибыли. Онлайн-культура сводит экзистенциальное измерение человеческого общения и одиночества к коллективным эмоциям и замыканию внутри «информационных пузырей». Как следствие, в условиях развития технологий ИИ человеческое общение — равно как и возможность побывать одному — все чаще воспринимается как роскошь.

Еще один аспект экзистенциального измерения развития технологий ИИ состоит в том, что искусственный интеллект — наше создание — вступает с нами

³ См.: AI Journey. URL: <https://ai-journey.ru/> (дата обращения: 20.02.2021).

⁴ См., например: От искусственного интеллекта к искусственной социальности: новые проблемы и задачи социологии // ВЦИОМ. 12.11.2020. URL: <https://profi.wciom.ru/index.php?id=2287&uid=20> (дата обращения: 20.02.2021); Возможен ли Хайдеггерянский искусственный интеллект? // ЕУ СПб. 04.12.2020. URL: <https://eusp.org/events/vozmozen-li-khaydeggerianskiy-iskusstvennyy-intellekt> (дата обращения: 20.02.2021).

⁵ Мы рассматриваем развитие онлайн-культуры как логически независимое, но эмпирически взаимосвязанное по отношению к развитию технологий ИИ. Под «онлайн-культурой» мы понимаем «ансамбль (гармоническую совокупность) сетей коммуникаций, технических устройств, алгоритмов, формальных и неформальных правил взаимодействия, паттернов поведения, культурных символов, которые делают возможной и структурируют активность людей в сети интернет и аналогичных сетях, обеспечивающих удаленный доступ к созданию, обмену и получению информации» [Резаев, Трегубова, 2019: 42].

во взаимодействие. Что при этом происходит с человеком? Что каждый из нас узнает о самом себе, взаимодействуя с машинами? Как ИИ отражает в измененном виде то, что получает от человека?⁶ По сути, это старая история Пигмалиона и Галатеи, учителя и ученика. Однако старая история развертывается в новых условиях. Мы учим ИИ взаимодействовать с людьми на данных, взятых от людей, накопленных в ходе взаимодействия людей друг с другом. Дальше ИИ учится сам, получает данные от своего взаимодействия с людьми, при этом накапливаются отличия в способах взаимодействия и в том, как они отображаются в данных. Одновременно новые поколения, выросшие в условиях искусственной социальности, взаимодействуют между собой уже по-другому, так как взаимодействуют преимущественно через машины. К чему приводит этот сдвиг? К каким новым типам взаимодействия и социальных отношений? Проблематика искусственной социальности — об этом, а не только о работающих машинах.

Технологии ИИ никогда еще не играли столь разнообразные и столь неоднозначные роли в жизни общества. Современное общество сталкивается с проблемами социального и антисоциального поведения, связанного с использованием технологий ИИ, препятствующего тому, чтобы люди создавали и поддерживали общественные отношения. Дискуссии о том, что искусственный интеллект «должен» и «может» делать, увы, часто концентрируются на структурном и институциональном неравенстве, которое пронизывает отношения между создателями и пользователями технологий.

Одна из задач, стоящих перед учеными, состоит в том, чтобы, организуя исследования, не упускать из виду, как их собственная позиция соотносится с тем сложным и иногда морально неудобным содержанием, которое ИИ привносит в жизнь людей. Человеческий опыт во взаимодействии с агентами ИИ весьма разнообразен, и очевидно, что многим людям такие взаимодействия приносят пользу. Однако это не отменяет необходимость критически исследовать «наплыв» и нерефлексивное использование таких технологий, которые приносят не пользу, а вред.

Развитие технологий ИИ идет рука об руку с подъемом онлайн-культуры, и вместе они делают неизбежным будущее под знаком искусственной социальности. Это ставит социальных аналитиков перед необходимостью задавать все новые вопросы о том, для кого и для какой цели предназначены технологии ИИ.

Социальная аналитика под знаком искусственной социальности

В 2018 г. во время конгресса Международной социологической ассоциации (MCA) в Торонто мы задались вопросом: готовы ли социологи к анализу искусственной социальности? [Rezaev, Tregubova, 2018]. Тогда мы начали исследования, которые думали и думаем продолжить в рамках форумов MCA. Пандемия нарушила планы, и в данном номере мы не можем представить результаты нового исследования. Тем не менее вопрос не только не утратил своей актуальности, но стал

⁶ Наиболее интересный пример — Replica, приложение для романтического или дружеского общения, подстраивающееся под взаимодействие с конкретным пользователем и обучающееся во взаимодействии с этим пользователем (<https://replika.ai/>). Replica создает цифрового двойника, который «отзеркаливает» то, что получает от человека — партнера по взаимодействию.

еще более значимым. Что изменилось за два года? Каковы тенденции развития интересов социальных ученых в отношении проблематики ИИ?

В 2018 г. ответ на вопрос о готовности социологов был отрицательным: интерес социальных ученых к ИИ и ИС был обусловлен скорее особенностями их биографических траекторий, нежели институциональными факторами. Если бы в 2020 г. мы имели возможность провести исследование в рамках форума МСА, мы бы сформулировали следующую гипотезу: интерес социологов к ИИ возрастает, однако это происходит под давлением извне: со стороны прикладных исследований рынков и организаций и со стороны критических исследований угнетения и неравенства. Иными словами, вспоминая «штампы» советского времени, социология — или служанка капитализма, или совсем наоборот.

Пандемия COVID-19 сделала такое положение дел очевидным. К социальным ученым обращаются за рекомендациями по вопросам оптимизации, увеличения эффективности онлайн-взаимодействий, за прогнозами в отношении общественного мнения по поводу использования новых технологий и новых способов регуляции повседневной жизни. В то же время от них ждут критики ошибок, искажений, пристрастных решений алгоритмов, которые, предположительно, приводят к усилению классового, расового, гендерного и иных видов неравенства в условиях современного капитализма.

Если наша гипотеза верна, то текущая ситуация может трактоваться как оптимистично — социологи нужны, так и пессимистично — от социологов хотят то, чего они не хотят или не могут делать.

Здесь мы придерживаемся осторожного оптимизма, с оговоркой: необходимо понимание того, что социальные ученые могут и чего не могут в принципе.

Социальная аналитика сегодня призвана не предсказывать результаты, а формулировать проблемы. В индустриальную эпоху, до развития искусственной социальности социологи были в состоянии, как минимум, предлагать варианты решения проблем. Сейчас это не так, потому что нарушается принцип, происходит сбой: мы не знаем, что вносит машина во взаимодействие с человеком, и не можем предсказать ее действия. Один из способов увидеть это — сказать, что ИИ не понимает экзистенциальные измерения человеческого существования: боль, одиночество, смерть, рождение, радость общения.

Отсюда ключевой вопрос для социальных ученых: как измерить искусственную социальность? Мы должны научиться измерять, чтобы сравнивать и оценивать эффекты воздействия ИС на человеческое общество и общение. Задачи социологов сегодня — в анализе взаимодействий людей с ИИ, а не в попытках старыми способами, опираясь на старый терминологический аппарат исследовать новые феномены ИИ и искусственной социальности. И здесь становится ясной необходимость кросс-дисциплинарного и, в пределе, антидисциплинарного исследования ИИ и ИС.

Как и на каких основаниях это может быть осуществлено? Очевидно, что сейчас мы находимся в самом начале пути. В рамках данного выпуска мы предлагаем 12 тезисов об ИИ и ИС, которые могут стать основанием для дальнейших исследований. Мы также адресуем читателя к обсуждению вопросов об антидисциплинарности и о новой социальной аналитике ИИ в нашей предыдущей работе (см. [От искусственного интеллекта..., 2020]).

ИИ в повседневных практиках: побочные эффекты капитализма

Одна из наиболее популярных тем в обсуждении проблем искусственной социальности — поиск все новых ошибок и искажений в работе агентов ИИ (см. рецензию в этом номере на работу [O’Neil, 2016]). Однако при этом не так много говорится об ошибках и искажениях, связанных с развитием капитализма и рыночной экономики.

Сегодня отношения между ИИ и повседневной жизнью общества определяются формой экономического рационализма, ориентированной на различные варианты потребительской активности. Происходит сдвиг от технологического и культурного к экономическому восприятию ИИ. Товарные ценности сводят ИИ к машине, жестко ориентированной на прибыль. При этом прогресс в технологиях ИИ идет рука об руку с развитием искусственной социальности, которая, в свою очередь, оказывает влияние на общественные отношения и общение.

Идеология конкуренции и потребительство разделяют людей, вместо того чтобы дать им возможность объединиться вокруг общих интересов. Искусственный интеллект, разработанный на основании таких установок, вероятно, принесет больше вреда, чем пользы. Практики, выставляющие эгоизм добродетелью, лишенные всякого понимания общественного блага и общительности, порождающие стыд и отчуждение у обездоленных, не должны стать основанием для разработок в сфере ИИ.

Еще одна опасность состоит в том, что правила, на которых основано взаимодействие ИИ и человека, меняются «по ходу игры». Рынок реагирует на общественный запрос: корпорации вводят квазиэтические ограничения, которые приводят к новым ошибкам и искажениям. Затем цикл повторяется, а прибыль получает тот, кто первым извлек из данных тренды, фиксирующие общественные настроения. При этом в условиях онлайн-культуры важным оказывается не решить проблему, а убедить общественность в том, что она решена⁷.

Когда Адам Смит писал о чудесах рынка, он имел в виду рынок, помещенный в рамках морального порядка людей, ориентированных на общество других и чувствительных к нему. Смит полагал, что моральный порядок будет поддерживать благосостояние бедных в рабочей среде. Не понимая, что рыночная экономика может подорвать моральный порядок, Смит предсказывал общество всеобщего материального достатка для всех. Но что происходит с таким видением, когда технологии ИИ разделяют людей по все новым линиям?

Ирония рыночного общества состоит в том, что оно увеличивает разрыв между богатыми и бедными, богатыми и всеми остальными. Более того, в современной рыночной экономике технологии ИИ создают избыточные потребности, одновременно препятствуя удовлетворению потребностей, которые есть у всякого человека: потребностей в регулярном питании, полноценном сне, общении, возможности побывать одному.

⁷ Здесь можно привести пример с популярной сетью кофеен, описанный в [O’Neil, 2016]. В 2014 г. в New York Times вышла статья о том, что алгоритмы ИИ составляют для сотрудников компаний крайне неудобное и непредсказуемое расписание, которое пагубно влияет на их сон, режим питания, а также на возможность совмещать работу с учебой. После скандала, вызванного статьей, представители компаний публично пообещали поменять практики составления расписания, так чтобы скорректировать решения алгоритмов. Однако на следующий год оказалось, что практики так и не поменялись, в основном из-за сопротивления младшего управляющего персонала, который должен достигать собственных показателей эффективности.

Содержание номера

Статьи, представленные в специальном выпуске, могут быть сгруппированы по трем основаниям: 1) теоретические/концептуальные и эмпирические; 2) отвечающие на вопросы что и как исследовать, 3) ориентированные на анализ собственно ИИ или онлайн-культуры. Данная классификация, однако, не отражает тематического и методологического разнообразия материалов номера: скорее, она служит для удобства читателя.

В первом разделе представлены статьи на английском языке. Раздел начинается с «Двенадцати тезисов об ИИ и ИС» («Twelve Theses on Artificial Intelligence and Artificial Sociality») А. В. Резаева, которые предлагают определение ИИ в системе координат научного знания, формулируют базовые характеристики и принципы исследования ИИ, призваны обратить внимание читателя на возможность и необходимость осмыслиения ИИ и ИС на основе человеко-ориентированного подхода, дать общее понимание того, как и почему ИИ становится предметом интереса для социальных ученых. Тезисы определяют основные направления анализа проблем, связанных с включением ИИ в повседневную жизнь общества, и обращают внимание на необходимость соотнесения проблем ИИ с развитием капитализма.

Афина Караджоянни рассматривает проблему ИИ в соотношении с этическими характеристиками и представляет план исследования, где основное внимание уделяется этическим суждениям и практическим рекомендациям, которые используются для осуществления политики в области ИИ в технологических корпорациях, государственных органах и в организациях, представляющих гражданское общество.

Нильс Кловайт и Мария Ерофеева в своей работе обращаются к вопросу о том, как этнотехнологические подходы расширяют наши возможности в аналитике ИИ. Авторы соотносят их с исследованиями в области Human-Computer Interaction, чтобы показать «конкурентные преимущества» мультимодального анализа.

Следующие две статьи рассматривают процессы ИС, которые имеют место в рамках организаций и затрагивают интересы и динамику повседневной жизни различных акторов. Нобуко Хосогая анализирует распространение онлайн-коммуникаций в сфере управления персоналом в Японии и последствия этого для данной области как пример развития ИС. Уолтер Аллен с соавторами сосредотачиваются на другом аспекте развития онлайн-технологий — на том, как принимаются решения, — на примере внедрения онлайн-образования в Восточной Африке.

Русскоязычный блок открывает статья А. В. Смирнова, в которой автор стремится ответить на вопрос: в какой мере и в каком смысле Россия является «цифровым обществом»? Ответ на поставленный вопросдается на основании анализа данных массовых опросов, а также онлайн-данных о населении России. Особое внимание автор уделяет проблематике «цифрового неравенства», его причин и возможных последствий.

Две статьи номера критически анализируют динамику онлайн-культуры и ее социальные и экономические последствия. А. П. Климович рассматривает вопрос о концентрации власти в эпоху цифрового капитализма, сравнивая случаи США и Китая как примеры взаимодействия демократического и авторитарного государства с IT-гигантами. Т. С. Мартыненко и Д. Е. Добринская ставят вопрос о том, как повсеместное распространение алгоритмов влияет на характер «цифрового

разрыва» и как их совместное воздействие меняет структуры социального неравенства в современных обществах.

Далее следуют микросоциологические решения проблемы того, как изучать ИС. А. М. Корбут, используя разные типы данных и разные методы их анализа, исследует, как появление в доме умной колонки меняет повседневность его обитателей. А. М. Давыдова с соавторами анализирует, как пользователи устройств селф-трекинга оценивают свое взаимодействие с ними, в том числе как реагируют на контроль со стороны умных устройств. А. И. Егорова и Н. Кловайт применяют конверсационный анализ для оценки умения современных технологий ИИ вести беседу на примере детального разбора диалога с участием Google Duplex.

Специальный выпуск был подготовлен в период пандемии COVID-19, которая поставила ученых, экспертов и управленцев перед необходимостью лавировать между сохранением жизни граждан своей страны, обеспечением их базовых прав и свобод и стимулированием экономического и социального развития. Статья И. А. Блохина и соавторов характеризует один из аспектов текущей пандемии — использование ИИ в сфере медицины. Авторы обсуждают сильные и слабые стороны ИИ в качестве диагноза в период пандемии и оценивают дальнейшие перспективы использования данной технологии в сходных ситуациях.

Следующие четыре статьи ориентированы на методологическую рефлексию о том, как проводить исследования в условиях развития онлайн-культуры и ИС. Н. Д. Трегубова с соавторами на примере исследования миграционных процессов на пространстве СНГ рассматривает, каким образом в онлайн-исследовании возникает «разделение труда» между людьми и алгоритмами. М. Б. Богданов и И. Б. Смирнов на основании анализа широкого круга специальной литературы выделяют преимущества и ограничения анализа «цифровых следов» в социальных науках. М. Ю. Александрова раскрывает возможности использования методов машинного обучения на примере предсказания частичного неответа респондентов с помощью наивного байесовского классификатора. Н. В. Ярцева представляет собственный опыт проведения исследования на платформе GDELT. Автор характеризует типы данных, возможности их анализа и визуализации, доступные для социальных ученых.

Т. А. Нестик, Н. Н. Седова и Е. Г. Климанова в своей статье характеризуют основные тенденции развития рынка прикладных социологических исследований в России на ближайшие годы. Авторы отмечают, что структура рынка исследований и специализация компаний будут меняться в связи с автоматизацией сбора и анализа данных в онлайн-среде, в том числе за счет развития интернета вещей и алгоритмов ИИ.

Специальный выпуск завершается рецензиями, знакомящими читателя с двумя монографиями, формулирующими ключевые тенденции становления ИС в условиях современного капитализма. А. А. Иванова обсуждает идею «оружия математического поражения» (*weapons of math destruction*), которую Кэти О'Нил развивает на материалах исследования использования алгоритмов в различных сферах общественной жизни. Д. М. Жихаревич представляет свои размышления о труде Шушаны Зубофф, посвященном концепции «надзорного капитализма» (*surveillance capitalism*).

Настоящий выпуск — одна из первых попыток обсуждения проблематики ИИ на страницах русскоязычных журналов по социальным наукам. Статьи, которые читатель найдет здесь, представляют модели (парадигмы) для дальнейших исследований ИС и указывают на «проблемные точки» во взаимодействии человека и машины/алгоритма. Выводы, к которым приходят авторы, предполагают разное понимание сущности ИИ и фиксируют разные аспекты вхождения технологий ИИ в повседневную жизнь общества. Мы рассматриваем разнообразие материала в специальном выпуске как преимущество: преждевременное закрытие проблемы не принесло бы пользу никому.

Заключительные замечания

Мы заканчиваем вводное слово к специальному выпуску серией вопросов, которые фиксируют парадоксальную ситуацию в современной социальной науке.

Что современный просвещенный читатель знает об искусственном интеллекте? Что он/она ожидает от специального выпуска журнала по социальным наукам, посвященного профессиональному интеллекту? Эти ожидания — ожидания социального ученого или обычного человека? Иными словами, каков источник знания и мнения об ИИ? Нужно ли здесь проводить границу между социологом и гражданином? А если нужно, то должен ли социолог знать больше (шире) — или меньше (уже, глубже)?

Мы уверены, что большинство читателей «Мониторинга общественного мнения» имеют некоторое представление о развитии технологий ИИ и об их влиянии на жизнь людей. В повседневной жизни все мы постоянно сталкиваемся как с самими технологиями ИИ, так и с их осмыслением в социальных медиа, СМИ, научно-популярной и художественной литературе, кино и т. п. Вместе с тем как социальные ученые мы оказываемся замкнуты в дисциплинарных рамках, которые, на первый взгляд, на удивление мало могут помочь в осмыслении проблем ИИ и ИС. То, что актуально для социологов, интересующихся ИИ, в конечном счете оказывается заимствовано из критической теории, STS (включая акторно-сетевую теорию), исследований коммуникаций или некоторого сочетания данных концептуальных построений. При этом важность социальных проблем, связанных с распространением ИИ, сегодня не вызывает сомнения даже у представителей технических наук.

Представляется, что социальный ученый в исследованиях ИИ сегодня знает, что он не знает, и в этом его сила и слабость. Наша задача — с одной стороны, понять и принять это, с другой — сделать шаг вперед к тому, чтобы знать, от очевидного — к неочевидному.

Несколько лет назад мы писали: что может быть социальное общение людей друг с другом? Сегодня пришло время задать вопрос: что может быть «социальное» искусственного интеллекта, понимаемого как источник новых социальных проблем — проблем для социальных наук и проблем для общества? Цель обоих вопросов — в том, чтобы побудить социологов обратиться к игнорируемой ими проблематике. Сегодня эти два вопроса соединяются под рубрикой исследования искусственной социальности.

Мы надеемся, что настоящий выпуск позволит, пусть в малой степени, разобраться в понимании технологий ИИ и последствий их использования и станет

проводником для социальных ученых в мир исследования искусственной социальности. Мы также надеемся, что тем, кто отстаивает гуманистические взгляды, материалы выпуска помогут в осмыслении и в поиске решений проблем, порождаемых нашим высокотехнологичным обществом.

В завершение мы хотели бы выразить признательность всем, кто принимал участие в создании специального выпуска и без чьего участия выпуск бы не состоялся.

Прежде всего, мы благодарим членов Международной исследовательской лаборатории СПбГУ ТАНДЕМ за интеллектуальную и эмоциональную поддержку, за точные и здравые суждения, которые очень помогли нам в подготовке специального выпуска.

Особую благодарность мы приносим коллективу журнала «Мониторинга общественного мнения» и лично Анне Кулешовой за профессиональную помощь в самых разных аспектах работы над выпуском. Для нас большая радость — работать с ними.

Наконец, мы благодарим всех авторов, подавших свои статьи в настоящий выпуск. Количество и качество более чем 60 материалов, поступивших на рассмотрение, сделали выбор приглашенного редактора очень непростым. Вместе с тем они отражают рост интереса научного сообщества к проблематике искусственной социальности и повышение общего уровня академической дискуссии.

Список литературы (References)

Земнухова Л. В. Социотехническое в цифровой социологии: методологические возможности и ограничения // Социология власти. 2018. Т. 30. № 3. С. 54—68. <https://doi.org/10.22394/2074-0492-2018-3-54-68>.

Zemnukhova L. V. (2018) The Sociotechnical in Digital Sociology: Methodological Possibilities and Limitations. *Sociology of Power*. Vol. 30. No. 3. P. 54—68. <https://doi.org/10.22394/2074-0492-2018-3-54-68>. (In Russ.)

Капелишников Р. И. Технологический прогресс — пожиратель рабочих мест? // Вопросы экономики. 2017. № 11. С. 111—140. <https://doi.org/10.32609/0042-8736-2017-11-111-140>.

Kapeliushnikov R. I. (2017) Is Technological Change a Devourer of Jobs? *Voprosy Ekonomiki*. No. 11. P. 111—140. <https://doi.org/10.32609/0042-8736-2017-11-111-140>. (In Russ.)

Кловайт Н. Рефлексивный антропоморфизм: неведение онтологии или невежественная онтология? // Социологический журнал. 2018. Т. 24. № 1. С. 8—33. <https://doi.org/10.19181/socjour.2018.24.1.5711>.

Klowait N. (2018) Reflexive Anthropomorphism: Ontological Ignorance, or Ignorant Ontology? *Sotsiologicheskiy Zhurnal*. Vol. 24. No. 1. P. 8—33. <https://doi.org/10.19181/socjour.2018.24.1.5711>. (In Russ.)

Коллинз Р. Может ли социология создать искусственный разум? // Личностно-ориентированная социология / П. Бергер, Б. Бергер, Р. Коллинз. М.: Академический проект, 2004. С. 566—598.

Collins R. (2004) Can Sociology Create an Artificial Intelligence? In: Berger B., Berger R., Collins R. *Personality-Oriented Sociology*. Moscow: Academic project. P. 566—598. (In Russ.)

Корбут А. «Простите, я никак не могу понять»: способы реагирования на непонимание во взаимодействии человека и робота // *Laboratorium: Журнал социальных исследований*. 2018. Т. 10. № 3. С. 57—78. <https://doi.org/10.25285/2078-1938-2018-10-3-57-78>.

Korbut A. (2018) ‘Sorry, I Can Not Understand’: Ways of Dealing With Non-Understanding in Human-Robot Interaction. *Laboratorium: Russian Review of Social Research*. Vol. 10. No. 3. P. 57—78. <https://doi.org/10.25285/2078-1938-2018-10-3-57-78>. (In Russ.)

Кузнецов А. Г. Туманности нейросетей: «черные ящики» технологий и наглядные уроки непрозрачности алгоритмов // *Социология власти*. 2020. Т. 32. № 2. С. 157—182. <https://doi.org/10.22394/2074-0492-2020-2-157-182>.

Kuznetsov A. G. (2020) Neural Network Nebulae: ‘Black Boxes’ of Technologies and Object-Lessons From Opacities of Algorithms. *Sociology of Power*. Vol. 32. No. 2. P. 157—182. <https://doi.org/10.22394/2074-0492-2020-2-157-182>. (In Russ.)

Максимова А. С., Глазков К. П. (2018) Существует ли цифровая микросоциология? // *Социология власти*. 2018. Т. 30. № 3. С. 14—37. <https://doi.org/10.22394/2074-0492-2018-3-14-37>.

Maximova A. S., Glazkov K. P. (2018) Is There a Digital Microsociology? *Sociology of Power*. Vol. 30. No. 3. P. 14—37. <https://doi.org/10.22394/2074-0492-2018-3-14-37>. (In Russ.)

Ним Е. Г. Селф-трекинг как практика квантификации телесности: концептуальные контуры // *Антропологический форум*. 2018. № 38. С. 172—192. <https://doi.org/10.31250/1815-8870-2018-14-38-172-192>.

Nim E. (2018) Self-Tracking as a Practice of Quantifying the Body: Conceptual Outlines. *Antropologicheskij forum*. No. 38. P. 172—192. <https://doi.org/10.31250/1815-8870-2018-14-38-172-192>. (In Russ.)

От искусственного интеллекта к искусственной социальности: новые исследовательские проблемы современной социальной аналитики / под ред. А. В. Резаева. М.: ВЦИОМ, 2020.

Rezaev A. V. (ed.) (2020) Artificial Intelligence on the Way to Artificial Sociality: New Research Agenda for Social Analytics. Moscow: VCIOM.

Резаев А. В., Трегубова Н. Д. «Искусственный интеллект», «онлайн-культура», «искусственная социальность»: определение понятий // Мониторинг общественного мнения: экономические и социальные перемены. 2019. № 6. С. 35—47. <https://doi.org/10.14515/monitoring.2019.6.03>.

Rezaev A. V., Tregubova N. D. (2019) Artificial Intelligence, On-Line Culture, Artificial Sociality: Definition of the Terms. *Monitoring of Public Opinion: Economic and Social Changes*. No. 6. P. 35—47. <https://doi.org/10.14515/monitoring.2019.6.03>. (In Russ.)

Сивков Д. Ю. Дигитальная социология: напряжения, неопределенности, нерешенные проблемы // Социология власти. 2018. Т. 30. № 3. С. 8—13. <https://doi.org/10.22394/2074-0492-2018-3-8-13>.

Sivkov D. Y. (2018) Digital Sociology: Tensions, Ambiguities, Unsolved Problems. *Sociology of Power*. Vol. 30. No. 3. P. 8—13. <https://doi.org/10.22394/2074-0492-2018-3-8-13>. (In Russ.)

Утехин И. Взаимодействие с «умными вещами»: введение в проблематику // Антропологический форум. 2012. № 17. С. 134—156.

Utekhin I. (2012) Interaction With ‘Smart Things’: Introduction. *Antropologicheskij forum*. No. 17. P. 134—156. (In Russ.)

Хониева Е. А. Гендер и дисплей: коммуникативные жанры и способы категоризации во взаимодействии с голосовыми ассистентами // Журнал социологии и социальной антропологии. 2017. № 5. С. 95—112. <https://doi.org/10.31119/jssa.2017.20.5.6>.

Khonieva E. A. (2017) Gender and the Display: Communicative Genres and Ways of Categorization in Interaction With Voice Assistants. *Journal of Sociology and Social Anthropology*. No. 5. P. 95—112. <https://doi.org/10.31119/jssa.2017.20.5.6>. (In Russ.)

Beer D. (2017) The Social Power of Algorithms. *Information, Communication & Society*. Vol. 20. No. 1. P. 1—13. <https://doi.org/10.1080/1369118X.2016.1216147>.

Boden M. (2016) AI: Its Nature and Future. Oxford: Oxford University Press.

Carley K. M. (1996) Artificial Intelligence within Sociology. *Sociological Methods & Research*. Vol. 25. No. 1. P. 3—30. <https://doi.org/10.1177/0049124196025001001>.

Collins H. (2018) Artifictional Intelligence: Against Humanity’s Surrender to Computers. Madford, MA: Polity Press.

Couldry N. (2015) Researching Social Analytics: Cultural Sociology in the Face of Algorithmic Power. In: Hanquinet L., Savage M. (eds.) *Routledge International Handbook of the Sociology of Art and Culture*. Routledge International Handbooks. London: Routledge. P. 383—395.

Esposito E. (2017a) Algorithmic Memory and the Right to be Forgotten on the Web. *Big Data & Society*. Vol. 4. No. 1. P. 1—11. <https://doi.org/10.1177/2053951717703996>.

Esposito E. (2017b) Artificial Communication? The Production of Contingency by Algorithms. *Zeitschrift für Soziologie*. Vol. 46. No. 4. P. 249—265. <https://doi.org/10.1515/zfsoz-2017-1014>.

Etzioni A., Etzioni O. (2017) Should Artificial Intelligence Be Regulated? *Issues in Science and Technology*. Vol. 33. No. 4. P. 32—36.

Floridi L. (ed.) (2015) The Onlife Manifesto. Being Human in a Hyperconnected Era. Cham: Springer.

- Ford M. (2018) Architects of Intelligence: The Truth About AI From People Building It. Birmingham: Packt Publishing.
- Hoffman S. G. (2016) Managing Ambiguities at the Edge of Knowledge: Research Strategy and Artificial Intelligence Labs in an Era of Academic Capitalism. *Science, Technology, & Human Values*. Vol. 42. No. 4. P. 703—740. <https://doi.org/10.1177/0162243916687038>.
- Joerges B. (1990) Images of Technology in Sociology: Computer as Butterfly and Bat. *Technology and Culture*. Vol. 31. No. 2. P. 203—227. <https://doi.org/10.2307/3105658>.
- Lee F., Larsen L. B. (2019) How Should We Theorize Algorithms? Five Ideal Types in Analyzing Algorithmic Normativities. *Big Data & Society*. Vol. 6. No. 2. P. 1—6. <https://doi.org/10.1177/2053951719867349>.
- McCorduck P. (1979) Machines Who Think. San Francisco: W. H. Freeman.
- O’Neil C. (2016) Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. New York, NY: Crown Publishers.
- Rezaev A. V., Tregubova N. D. (2018) Are Sociologists Ready for ‘Artificial Sociality’? Current Issues and Future Prospects for Studying Artificial Intelligence in the Social Sciences. *Monitoring of Public Opinion: Economic and Social Changes*. No. 5. P. 91—108. <https://doi.org/10.14515/monitoring.2018.5.10>.
- Ruppert E., Law J., Savage M. (2013) Reassembling Social Science Methods: The Challenge of Digital Devices. *Theory, Culture & Society*. Vol. 30. No. 4. P. 22—46. <https://doi.org/10.1177/0263276413484941>.
- Russell S., Norvig P. (2016) Artificial Intelligence: A Modern Approach. Harlow, UK: Pearson Education Limited.
- Wolfe A. (1993) The Human Difference: Animals, Computers, and the Necessity of Social Science. Berkley: University of California Press.
- Woolgar S. (1985) Why not a Sociology of Machines? The Case of Sociology and Artificial Intelligence. *Sociology*. No. 4. P. 557—572. <https://doi.org/10.1177/0038038585019004005>.
- Ziewitz M. (2016) Governing Algorithms: Myth, Mess, and Methods. *Science, Technology, & Human Values*. Vol. 41. No. 1. P. 3—16. <https://doi.org/10.1177/0162243915608948>.

DOI: [10.14515/monitoring.2021.1.1894](https://doi.org/10.14515/monitoring.2021.1.1894)



A. V. Rezaev

TWELVE THESES ON ARTIFICIAL INTELLIGENCE AND ARTIFICIAL SOCIALITY

For citation:

Rezaev A. V. (2021) Twelve Theses on Artificial Intelligence and Artificial Sociality. *Monitoring of Public Opinion: Economic and Social Changes*. No. 1. P. 20–30. <https://doi.org/10.14515/monitoring.2021.1.1894>.

Правильная ссылка на статью:

Резаев А. В. Двенадцать тезисов об искусственном интеллекте и искусственной социальности // Мониторинг общественного мнения: экономические и социальные перемены. 2021. № 1. С. 20—30. <https://doi.org/10.14515/monitoring.2021.1.1894>. (In Eng.)

TWELVE THESES ON ARTIFICIAL INTELLIGENCE AND ARTIFICIAL SOCIALITY

*Andrey V. REZAEV¹ — Prof. Dr. habil., Director of International Research Laboratory TANDEM
E-MAIL: rezaev@hotmail.com
<https://orcid.org/0000-0002-4245-0769>*

¹ Saint Petersburg State University, Saint Petersburg, Russia

Abstract. The Theses deal with the theoretical foundations and methodological implications for scholarly research that arise from the development and implementation of artificial intelligence (AI) technologies into society's daily life. The reader is introduced to age-old intellectual debates about AI and recent research concerning human-centered AI, artificial sociality (AS), and online culture. The paper presents the working definition of AI. It claims that AI has to be examined in relation to AS. The paper argues that the human-machine-interdependence is a new reality of artificial sociality. It envisages AI research as multidisciplinary and potentially a-disciplinary scientific activity. The questions the Theses raise: What should we be concerned about as artificial intelligence advances? Can AI technologies solve modern society's problems and bring human beings to a new level of community and well-being? Are there 'no-AI areas' in society? Do human biases and prejudices influence AI technologies? The paper's essential assertion is that the challenges posed by AI technologies and AS should be addressed apropos three P's of the capitalist society: private property, profit, price.

ДВЕНАДЦАТЬ ТЕЗИСОВ ОБ ИСКУССТВЕННОМ ИНТЕЛЛЕКТЕ И ИСКУССТВЕННОЙ СОЦИАЛЬНОСТИ

*РЕЗАЕВ Андрей Владимирович — доктор философских наук, профессор, руководитель Международной исследовательской лаборатории ТАНДЕМ, Санкт-Петербургский государственный университет, Санкт-Петербург, Россия
E-MAIL: rezaev@hotmail.com
<https://orcid.org/0000-0002-4245-0769>*

Аннотация. Тезисы формулируют фундаментальные теоретические и методологические проблемы, стоящие перед социальными науками в связи с вхождением технологий искусственного интеллекта (ИИ) в повседневную жизнь общества. Аргументы, развиваемые в тезисах, основаны на классических научных дискуссиях об ИИ и на современных исследованиях, касающихся человеко-ориентированного ИИ, искусственной социальности и онлайн-культуры. В статье предлагается рабочее определение искусственного интеллекта и обосновывается, что ИИ должен изучаться в связи с искусственной социальностью. Автор утверждает, что взаимозависимость «человек — машина» является новой реальностью искусственной социальности, и рассматривает исследования ИИ как междисциплинарную и потенциально антидисциплинарную научную область. В тезисах рассматриваются следующие вопросы: какие проблемы должны стать предметом внимания социальных ученых по мере развития искусственного интеллекта? Можно ли с помощью технологий ИИ решить проблемы современного общества и вывести его на новый уровень солидарности и материального благополучия?

чия? Могут ли в современном обществе существовать сферы, свободные от ИИ? Подвержены ли технологии искусственного интеллекта влиянию человеческих предубеждений и предрассудков? В заключение автор формулирует положение о том, что вызовы, обусловленные развитием искусственного интеллекта и искусственной социальности, необходимо рассматривать в контексте трех основных слагаемых капиталистического общества: частной собственности, прибыли и рыночных отношений.

Keywords: artificial intelligence, artificial sociality, human-machine interdependence, online culture, new social analytics, antidisiplinarity, human biases, capitalist society

Ключевые слова: искусственный интеллект, искусственная социальность, взаимозависимость «человек — машина», онлайн-культура, новая социальная аналитика, антидисциплинарность, человеческие предубеждения, капиталистическое общество

Widespread penetration of Artificial Intelligence (AI) in the everyday life of society, the exponential growth of research and publications about AI in computer sciences, natural sciences, and engineering contrasted with the sluggish and cautious efforts to examine AI in the social sciences and humanities is not a problem, today. It is a fact. What to do about it is a problem.

We proceed from an obvious premise — it is reasonable to think that AI will be part of the real world for human beings in the foreseeable future. It is also reasonable to postulate basic principles for reasoning on AI and formulate a rationale for scholarly examinations of the AI phenomena. To fulfill this objective, we develop our 12 theses on AI and Artificial Sociality (AS)¹.

Thesis One: AI appears in the mind of human beings as a combination of the three Ps — Phenomenon, Problem, and Phrase (or a concept)²

As a *phenomenon* of everyday life, AI cannot be revealed in any other way but in the form of materialized (tangible) products/devices. As such, these devices have a double determination — technical and social. On the one hand, AI appears as a technological device designed to solve a task that is impossible for a human. On the other hand, AI does not exist only in a technical environment. Technologies initially focused on

¹ [Malsch, 2005] validated AS as a notion for sociological studies.

² In more detail, we developed our position in [Rezaev, 2020].

instrumental tasks become the environment and participant in human interactions. We call this tendency ‘artificial sociality’.

As a research *problem*, artificial intelligence: (1) raises philosophical (worldview) questions; (2) in a new way characterizes social reality; sheds new light on the specifics and potentials of social reality; calls to reconsider the mainstream visions of social reality; (3) appears differently in different sciences. For natural and engineering sciences, problems of artificial intelligence are related to solving technical and instrumental problems. For sociology and other social sciences, the question of AI is essentially marginal, a side issue. In philosophy, as in humanities in general, AI is discussed in relation to worldview problems resolved differently in different historical epochs and different intellectual traditions.

As a *phrase* (concept) in scholarly literature, AI is not yet defined to have a generally accepted meaning³.

However, scholars must agree on defining AI to proceed in rational thinking about AI implementation into the reality of human lives and societal development.

Thesis Two: AI is not an objectivated thing but an ensemble of formalized rules

AI’s appearance in daily activities as a tangible product should not mislead scholars in determining its substantial nature.

At this juncture of scientific discussions, we have all the foundations to define AI as follows⁴:

Artificial Intelligence is an ensemble of rational, logical, and formalized instrumental rules developed and coded by human beings that organize the processes and activities to emulate rational/intellectual structures and fabricate and reproduce goal-oriented practices as well as the mechanisms for constructing further coding and decision making.

Five fundamental characteristics underline our definition:

- First, AI is an artifice, a product of human beings, it is not something transcendental or inherently a-social.
- Second, the substantial nature of AI is not something objectifiable as a material gimmick; AI’s essence manifests itself as a set of rational and logically formalized rules but not as an attained appliance per se.
- Third, AI is a set of instrumentally coded rules.
- Fourth, an instrumentally coded set of rules is oriented to generate/yield an appliance/product that imitates human beings’ intellectual activities.
- Fifth, emulated intellectual constructions make it possible for AI to further independently code and make intellectual or goal-oriented decisions, not necessarily with assistance or control from human beings.

³ For a vivid example of this statement, see the recent article on the definition of AI [Wang, 2019] and the subsequent discussion organized in the *Journal of Artificial General Intelligence* (2020) Vol. 11. No. 2. URL: <https://content.sciendo.com/view/journals/jagi/11/2/jagi.11.issue-2.xml> (accessed: 17.02.2021).

⁴ See also: [Rezaev, 2020]. In the attempt to present our working definition of AI we follow the paths proposed in [Russel, Norvig, 2016].

Thesis Three: AI is an inherently non-disciplinary project that brings anti-disciplinarity to academic research

In today's scholarly world, every discipline creates its objects for study and terminology to establish the borders. For example, although scholars use the common term 'nature' in biology and chemistry, the meaning is quite different. Indeed, 'nature' is a core notion for organizing discussion in science, but each discipline has constructed its view on explaining what nature is and how it functions. What biologists understand by nature does not necessarily fit what chemistry or physics define by nature. Disciplinarity turns out to be a real impediment when scholars from various departments meet to resolve a scientific problem, even if it is in the domain of so-called hard sciences. It also applies to 'cross-talk' in social and behavioral sciences.

Further, the sciences are usually distinguished from what is sometimes called the liberal arts, sometimes humanities. Three misconceptions generally accompany this distinction. One is that such division was always in the history of written culture. The second is that it will exist so far as scientific experiments' reality and the reality of words prevail. The third that it is evident from the methods employed in the two fields. However, historically the humanities signified the secular part of the curriculum taught in the medieval church schools (scholae) founded by Charlemagne in the eighth century.

Moreover, no separation of science was apparent until the Renaissance, when a temporary division was imposed due to somewhat superfluous circumstances.

The idea of AI as a technological innovation for social well-being was introduced in the mid-50s of the last century by a group of scholars who were not oriented in their vision toward one specific discipline⁵. As an object of scholarly interest AI from the very beginning was and could be considered only as a multidisciplinary and potentially a/anti-disciplinary object. The history of AI development shows that the more stress on interdisciplinarity in research and design, the better the outcomes.

AI must help scholars and humanists merge into one perspective to find new ways of seeing AI and what it is for human beings. It can be done only and if multi-cross-interdisciplinarity methodology evolves into a non-disciplinary structure. Further progress in AI research and design calls for an *a-disciplinary* approach.

Thesis Four: Human Centered AI calls for new Social Analytics

There is a sound and justifiable approach to the study and design of AI developed recently by scientists who are at the forefront of AI research. These are predominantly scholars with backgrounds in computer sciences, hard sciences, and technological engineering. Their vision for AI development in society is based on the understanding that AI's mission is to augment and enhance, but not to replace, human beings. The approach is oriented toward cross-multi-disciplinarity. It invites into conversation and co-work on AI representatives from the humanities and the social sciences. It is the most productive approach today. It is called Human Centered AI or Human Centric AI⁶.

As opposed to pure technology-driven modernization, the importance of human-centered AI surfaces as a useful tool to question, reflect on, and offer alternatives to nor-

⁵ For more details, see [Winograd, 1991] and [McCorduck, 2004].

⁶ For examples see URL: <https://hai.stanford.edu/> (accessed: 08.02.2021) and URL: <https://www.humane-ai.eu/> (accessed: 08.02.2021).

mative technology-based everyday practices. The quickest way to express what is at stake here is to say that the point of AI in society is not that humans have new tools but that they have new tools from someone, for particular purposes, and in a new set of relationships.

AI designers produce what interests them and makes them happy. In their minds, their highest-value product is a new version of their own values. They cannot but define value by what they value. However, designers' values may have nothing to do with the ideals of manufacturers or consumers [see Ford, 2018].

The social sciences and humanities representatives are not yet continually active in studying AI phenomena [Collins, 1992]. Again, the disciplinarity continues to dominate; and philosophers are not in accordance with anthropologists, cultural sociologists—with lawyers or economists.

We believe Social Analytics as a general approach that calls for a diversity of scientific methods and theoretical frameworks will achieve more productive results in studying AI at this juncture of scientific investigation of AI phenomena and problems.

Therefore, Social Analytics refers to a lens through which to examine social, legal, ethical, economic, political, engineering problems, intellectual dislocations, exclusions, and challenges for exploring the development and effects of AI phenomena on society.

Thesis Five: For Social Analytics it does not matter whether or not AI will be invented—the point is how to explore the interaction between humans and existing AIs

It does not matter for Social Analytics whether AI (or a General Artificial Intelligence (AGI), also called strong-AI [Goertzel, 2014]) will be invented or not. At this point, there are so many machines that are better at something than humans. Machines can and already do their jobs better than humans. The challenge for social analytics is how to explore the interaction between humans and machines, between machines when humans are off the interaction line.

The real problem for social analytics will be when AI is incorporated into AI research, into studying AI full scale. Now humans study AI in the framework of humans' cycles (time frame, imagination, emotions, etc.). AI as such will have no limits for studying AI. When Google can replace 50 000 engineers with ten AI systems, the time scale in science will be driven by machines but not by humans with their objective/subjective constraints.

Despite arguments that AI cannot be understood, in fact, we know more about what AI is than we know who humans are. Moreover, at the present stage of scientific knowledge development, we have more opportunities for understanding machines and algorithms than for understanding people. The challenge is to see that the human brain is not analogous to computer hardware, and the mind is not comparable to computer software [Dreyfus, 1992; Wolfe, 1993].

Social Analytics' immediate goal is to develop an outlook that goes beyond an 'old paradigm' of AI. It was presupposed for a long time that AI is and can function as an autonomous system or as a robot. This implied that the goal for AI is to replace human jobs. For example, the idea of artificial general intelligence (AGI) is based on the presupposition that machines will emulate human cognition, and there will come a time when AI—as a superintelligence—will be smarter than humans and take over the world [see Bostrom, 2012; Chalmers, 2010].

Social Analytics and Human-Centred approach to AI must formulate a new paradigm based on the partnership and co-existence of humans and AI. It cannot exist without human consciousness that is a hard riddle to resolve in the near future. It means that human intelligence cannot be replaced soon. Thus, humans and AI need to work together. The overall objective for AI is not to confront but to improve the reality of human life. The tasks for humans are to make AI technologies responsible and accountable. The human-machine interdependence is a new reality of artificial sociality. How to code the wider social context of human-human interaction?

Thesis Six: Human-machine interdependence is also a question of trust; it is not only about math

As a product of a human being, algorithm and data embed historical practices and social biases. You cannot rely on the machine because it is based on data and algorithms.

Recent research shows that there is an issue of algorithmic bias [O’Neil, 2016]. In the reality of human-machine interdependence, the algorithmic bias imposes systemic threats. It is revealed in problems of how human biases interact with the machines’ biases. How do biased humans interact with an algorithmic bias, and will this impose systemic threats for AI technologies’ implementation in everyday life? How will daily life be affected?

Thus, human-machine interdependence is contested by the value alignment problem.

There are two sides to the question here⁷: (1) technical — how to give AI (machines) an understanding of human values and goals; how will AI process what humans want, how they feel about certain things, and what is their worldview, and (2) whose values it should be? Society is divided; bankers and housekeepers have different values. The discriminatory codes and designs of AI technologies have the potential to hide, replace social divisions, and amplify racial hierarchies [Benjamin, 2019].

To be biased means to be human, to be a socially organized human being. Biases are due to social class, education, resources availability, culture, religion, gender, sexuality, age. Biases also have purposes and, in different situations, might be helpful or harmful. The objective is to minimize harmful and unintentional biases in the systems of interaction between humans and AI.

There is also the third side. The famous ‘Doctor Evil problem’: somebody does not care about humanity but cares only about power in the world. It raises fears s/he will control AI — this might be even worse when someone takes control over the nuclear bombs. The point is that ‘good guys’ might also misuse AI and damage humanity. The further development of AI technologies, the more pressure on ethical problems. There is no other way to develop AI in a society but to base it on moral norms and values. And specific technological issues like how to code human values are related to the fundamental question: what exactly does it mean to code human values?

Thesis Seven: Artificial sociality is contested by the problems of jobs replacement and leisure

The problem is not where people will work — the question is what people will do with their free time, how to organize their leisure in the age of AI and AS⁸.

⁷ For a detailed discussion, see [Russell, 2019].

⁸ AI technologies and online culture beget specific problems for the youth [see Twenge, 2017].

One of the main fears associated with the development of AI is the loss of jobs. Aristotle also wrote about machines that could replace human labor (in his time—slaves' labor). Why were such prospects not a matter of concern in Antiquity? Is not it because modern life is organized around professional activities that bring (a) an income, (b) self-respect and respect for others, and (c) a particular way of life? The ideal type of modern professional is a virtuoso—a narrow specialist (not a master, sage, or community member). The figure of a specialist is opposed by a social activist (volunteer) model who implements abstract principles of justice. The entry of AI into the labor markets leads to the fact that both the specialist and the activist are no longer needed: AI will be useful in solving narrow tasks and at the same time feed the hungry, protect the humiliated, and enlighten the illiterate. What are the remaining charges for people to go through with then?

Thesis Eight: Human sociality is contested by the lack of solitude in the world of algorithms

Online culture is both an intensive info-instrumental culture and the culture of existential atomization and loneliness. Human loneliness in online activities is intensified by offering the self as the prime source of enjoyment and the only source of agency worth validating⁹.

With the development of online culture, AI technologies penetrate the daily lives of people of all ages and different social statuses on all continents. Online culture changes reality and monetizes social interactions. It translates existential questions into a discussion of social problems and the release of collective emotions. Can a person be able to remain himself/herself in the situation when the machines regulate and mediate communication? Will human-machine interdependence give humans room for privacy?

The question is not in people's access to algorithms—the problem is whether a person can remain solitary and be herself/himself in privacy when the machines organize social interactions?

Thesis Nine: Are there 'no-AI areas' in society?

There is no sense in going back to a certain point in the history of science and technology, where society turned 'in the wrong direction' for some reason. We should not look for a 'Golden age' in history, but something that will allow society to move forward in entirely new conditions.

The point is not to replace 'imperfect' human activity with perfect AI. People, their life, health, and development are the meaning and ultimate goals of human coexistence in society. Replacing a person with a system of technological tools makes the existence of society meaningless. Only when a person is at the center of socio-political and cultural processes can the gap between people and technology be bridged in the center of society's attention. It is not that human-made technologies are developed to manipulate people, but people manage technology for society's benefit.

The real problem for Human Centered AI is how to define the areas, places where humans do not have to touch with artificial intelligence, of course, if there are those.

⁹ Spectacular examples you will find in [Turkle, 2011].

Suppose there are such areas that would coordinate and regulate, search and guide the further development of AI. So, the primary task is to find unquestionable taboo areas and how to enforce the taboos for AI.

Thesis Ten: AI progress will continue to change every aspect of life, and society will need to have ‘algorithmic auditing companies’

AI, AS, and online culture will not change human nature, at least in a historically short period of time. However, it is AI that will change the way people are organized. It is AI that will foster new possibilities for constructing and shaping society, affecting and simulating/mimicking the system of social relations. The forms and methods of organizing social relations will continue to change through information flows and modes and channels of communication. The intensity of human-machine interaction, its content, and, accordingly, the conditions for the formation of the prevailing values and ethical norms in society will change.

Thus, the point for scholars is to learn how to orient AI and AS to augment humans and enhance humane in social structures and institutions, to foresee, socially organize inevitable ‘human-machine interdependence’ and prevent its transformation into something a-societal, to advance society further based on social relations.

Therefore, such questions as ‘What will we have to change in the system of “human-algorithm interaction” in the future?’, ‘How to organize auditing of AI entrance in real life of human beings?’, ‘Who is the owner of AI technologies?’ are more essential than the question ‘What rewards AI brings to the society?’ [see also Etzioni, Etzioni, 2017].

AI technologies make the ownership of the machine and the ownership of intellectual property crucial. We have machines, algorithms, and an increasing number of patents, so the question of who controls the algorithms and owns the patents and the income flow associated with these properties becomes the first question for technology, economy, and politics. The public regulation of AI ownership becomes the first question for society.

Thesis Eleven: Philosophers and Scientists have tried to find similarities between Human Beings and AI, but the point is to see the differences between them

Humans are not Machines. Humans are not Computers, and Computers are not Humans. The primary deficiency of hitherto-existing views on AI is that scientists look for similarities between what computers are doing and what humans are doing. However, the real need is the opposite: science has to define what are the *differences* between humans and machines, what makes them unique, and how to enrich humanity via machines and AI.

Thesis Twelve: AI development is framed by the three Ps of capitalism: Private Property, Price, and Profit

The capitalism that we have today is the kind of capitalism that is empty of purpose, oriented toward the accumulation of capital in order to accumulate more capital. It does not have any articulated objective for societal well-being; it is still based on

the other three Ps: Private Property, Profit, Price (market¹⁰). And Artificial Sociality will increase capitalism's 'emptiness', or it might increase its social well-being if it is societal mission-driven capitalism¹¹.

Of course, due to current discourses developed in media and public at large, society might expect that capitalism with AI and AS will be oriented toward solving the challenging problems for humanity and eliminating the vices of climate change and preserving nature as well as fighting with possible pandemics. However, there is only one mission that current capitalism might have — it is self-preservation, preservation of its three Ps: Private Property, Price, and Profit. At this historical juncture, AI and AS can do nothing but supplement and enhance this mission¹². The question of what to do with this is both a question of theory and practice.

References

- Benjamin R. (2019) *Race After Technology: Abolitionist Tools for the New Jim Code*. Oxford (UK): Polity Press.
- Bostrom N. (2012) The Superintelligent Will: Motivation and Instrumental Rationality in Advanced Artificial Agents. *Minds & Machines*. Vol. 22. No. 2. P. 71—85. <https://doi.org/10.1007/s11023-012-9281-3>.
- Chalmers D. J. (2010) The Singularity: A Philosophical Analysis. *Journal of Consciousness Studies*. Vol. 17. No. 9—10. P. 7—65.
- Collins R. (1992) Can Sociology Create an Artificial Intelligence? In: R. Collins (ed.) *Sociological Insight: An Introduction to Non-Obvious Sociology* (2nd ed.). New York, NY: Oxford University Press. P. 155—184.
- Dreyfus H. (1992) *What Computers Still Can't Do: A Critique of Artificial Reason*. Cambridge, MA: MIT Press.
- Etzioni A., Etzioni O. (2017) Should Artificial Intelligence Be Regulated? *Issues in Science and Technology*. Vol. 33. No. 4. P. 32—36. URL: <https://issues.org/perspective-should-artificial-intelligence-be-regulated/> (accessed: 10.02.2021).
- Ford M. (2018) *Architects of Intelligence: The Truth about AI from the People Building it*. Birmingham (UK): Packt Publishing Ltd.
- Goertzel T. (2014) The Path to More General Artificial Intelligence. *Journal of Experimental & Theoretical Artificial Intelligence*. Vol. 26. No. 3. P. 343—354. <https://doi.org/10.1080/0952813X.2014.895106>.
- Malsch T. (2005) *Kommunikationsanschlüsse. Zur Soziologischen Differenz von Realer und Künstlicher Sozialität*. Wiesbaden: VS Verlag für Sozialwissenschaften.

¹⁰ Today, specifically after the pandemic, it is clear that a market cannot exist without a government to organise and enforce it. The critical question is to what purpose or whom the market has been organized to serve.

¹¹ Of course, if Human-Centred Approach to AI will be attained.

¹² For more details, see [Zuboff, 2019].

- McCorduck P. (2004) *Machines Who Think: A Personal Inquiry into the History and Prospects of Artificial Intelligence* (2nd ed.). Natick: A K Peters.
- O'Neil C. (2016) *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. New York, NY: Crown Publishing Group.
- Rezaev A. V. (ed.) (2020) *Artificial Intelligence on the Way to Artificial Sociality: New Research Agenda for Social Analytics*. Moscow: WSIOM; CrossRoads.
- Russell S., Norvig P. (2016) *Artificial Intelligence: A Modern Approach*. Harlow: Pearson Education Ltd.
- Russell S. (2019) *Human Compatible: AI and the Problem of Control*. London: Penguin Books Ltd.
- Turkle S. (2011) *Alone Together: Why We Expect More from Technology and Less from Each Other*. New York, NY: Basic Books.
- Twenge J. M. (2017) *iGen: Why Today's Super-Connected Kids Are Growing Up Less Rebellious, More Tolerant, Less Happy—and Completely Unprepared for Adulthood—and What That Means for the Rest of Us*. New York, NY: Atria Books.
- Wang P. (2019) On Defining Artificial Intelligence. *Journal of Artificial General Intelligence*. Vol. 10. No. 2. P. 1—37. <https://doi.org/10.2478/jagi-2019-0002>.
- Winograd T. A. (1991) Oral History Interview with Terry Allen Winograd. Stanford, CA: Charles Babbage Institute.
- Wolfe A. (1993) *The Human Difference: Animals, Computers, and the Necessity of Social Science*. Berkley, CA: University of California Press.
- Zuboff S. (2019) *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. New York, NY: Public Affairs.

DOI: [10.14515/monitoring.2021.1.1911](https://doi.org/10.14515/monitoring.2021.1.1911)



A. Karatzogianni

RESEARCH DESIGN FOR AN INTEGRATED ARTIFICIAL INTELLIGENCE ETHICAL FRAMEWORK

For citation:

Karatzogianni A. (2021) Research Design for an Integrated Artificial Intelligence Ethical Framework. *Monitoring of Public Opinion: Economic and Social Changes*. No. 1. P. 31–45. <https://doi.org/10.14515/monitoring.2021.1.1911>.

Правильная ссылка на статью:

Караджянни А. Интеграция этических оснований искусственного интеллекта: план исследования // Мониторинг общественного мнения: экономические и социальные перемены. 2021. № 1. С. 31—45. <https://doi.org/10.14515/monitoring.2021.1.1911>. (In Eng.)

RESEARCH DESIGN FOR AN INTEGRATED ARTIFICIAL INTELLIGENCE ETHICAL FRAMEWORK

*Athina KARATZOGLIANNI¹—Professor in Media and Communication
E-MAIL: athina.k@le.ac.uk
<https://orcid.org/0000-0002-6161-4423>*

¹ University of Leicester, Leicester, England

Abstract. Artificial Intelligence (AI) regulatory and other governance mechanisms have only started to emerge and consolidate. Therefore, AI regulation, legislation, frameworks, and guidelines are presently fragmented, isolated, or co-exist in an opaque space between national governments, international bodies, corporations, practitioners, think-tanks, and civil society organisations. This article proposes a research design set up to address this problem by directly collaborating with targeted actors to identify principles for AI that are trustworthy, accountable, safe, fair, non-discriminatory, and which puts human rights and the social good at the centre of its approach. It proposes 21 interlinked substudies, focusing on the ethical judgements, empirical statements, and practical guidelines, which manufacture ethicopolitical visions and AI policies across four domains: seven tech corporations, seven governments, seven civil society actors, together with the analysis of online public debates. The proposed research design uses multiple research techniques: extensive mapping and studies of AI ethics policy documents and 120 interviews of key individuals, as well as assorted analyses of public feedback discussion loops on AI, employing digital methods on online communities specialising in AI debates. It considers

ИНТЕГРАЦИЯ ЭТИЧЕСКИХ ОСНОВАНИЙ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА: ПЛАН ИССЛЕДОВАНИЯ

*КАРАДЖОЯННИ Афина — профессор, департамент медиа и коммуникации, Лестерский университет, Лестер, Англия
E-MAIL: athina.k@le.ac.uk
<https://orcid.org/0000-0002-6161-4423>*

Аннотация. Сегодня механизмы управления и регулирования для искусственного интеллекта (ИИ) только начинают формироваться. Принципы регулирования, правовые основы и «дорожные карты» развития ИИ фрагментарны и разрозненны, они существуют в теневом пространстве между национальными государствами, международными институциями, бизнес-корпорациями, сообществами практиков, аналитическими центрами и организациями гражданского общества. В статье предлагается план исследования, направленного на решение данной проблемы и предполагающего сотрудничество с ключевыми акторами с целью определения принципов развития ИИ, которые были бы надежными, понятными, безопасными, справедливыми, беспристрастными, ставили бы в центр права человека и общественное благо. Дизайн исследовательского проекта предполагает проведение взаимосвязанных мини-исследований (21 шт.), направленных на анализ этических суждений, эмпирических фактов и практических рекомендаций, формирующих этику и политику в отношении ИИ. План предполагает проведение исследований в семи технологических корпорациях, в административных органах семи национальных государств и в семи орг-

novel conceptual interactions communicated across the globe, expands the regulatory, ethics, and technological foresight, both at the individual level (autonomy, identity, dignity, privacy, and data protection) and the societal level (fairness/equality, responsibility, accountability and transparency, surveillance/datafication, democracy and trust, collective humanity and the common good). By producing an innovative, intercontinental, multidisciplinary research design for an Ethical AI Standard, this article offers a concrete plan to search for the Holy Grail of Artificial Intelligence: Its Ethics.

ганизациях гражданского общества, а также анализ публичных дебатов в интернете. Дизайн исследования включает несколько методов: подробное картирование и изучение политических и юридических документов, касающихся этики ИИ; 120 экспертиво; анализ циклов общественного обсуждения ИИ в специализированных онлайн-сообществах. Исследование направлено на анализ новых концептуальных взаимодействий между участниками процесса по всему миру, а также на расширение возможностей нормативного, этического и технологического прогнозирования как на индивидуальном уровне (вопросы автономии, идентичности, достоинства, конфиденциальности и защиты данных), так и на уровне общества (справедливость/равенство, ответственность, подотчетность и прозрачность, надзор/датафикация, демократия и доверие, общественный гуманизм и общее благо). Представляя дизайн инновационного, международного и междисциплинарного исследования этического стандарта ИИ, статья предлагает конкретный план поиска Святого Грааля искусственного интеллекта — его этических оснований.

Keywords: artificial intelligence, artificial intelligence regulation, ethical artificial intelligence standard, artificial intelligence policy, multidisciplinary research design, artificial intelligence ethics

Ключевые слова: искусственный интеллект, регулирование искусственного интеллекта, этический стандарт искусственного интеллекта, политика в отношении искусственного интеллекта, междисциплинарный исследовательский дизайн, этические основания искусственного интеллекта

Introduction

In Bristol, Artificial Intelligence (AI) has given us a ‘youth score’ computer programme, which combines crime data, housing information, and links them to others viewed as high-risk, together with information about the youth’s parents and domestic incidents. It also feeds school attendance records in. The police and social workers then surge resources towards high-risk cases and away from those that do not meet the indicators. In Philadelphia, face-to-face interviews with parole officers have been overtaken by predictive algorithms to set probation rules. In Amsterdam, algorithms identify welfare fraud risks and allocate credit. In February 2020, *The New York Times* observed that we have already entered an era when an algorithm grants freedom or takes it away¹.

Daily, advanced democratic societies are forced into operating more digitally by default. In this context, our social lives are increasingly governed by algorithms. AI software predicts who will commit a crime and making probation decisions, which demographics can have loans, who to provide healthcare to, who to hire or admit to university, even guiding sentences handed down by judges. Black box and unaccountable technologies offered by unregulated private companies have a profound effect on authority, trust, and transparency, with profound consequences for justice, education, and welfare in societies around the world.

And yet, there is a lack of a global ethical agreement on Artificial Intelligence (AI), although it poses the most significant moral challenge of our time. We are remarkably short of evidence-based social science research on how these systems are working now, how they are governed, and mainly how ethical standards are being practically applied, especially regarding social and economic inclusion [Jobin, lenca, Vayena, 2019; Redden, Dencik, Warne, 2020; Sanchez-Monedero, Dencik, Edwards, 2020]. Because this lack represents a severe test of humane values, it drives the central vision of this research design experiment: to propose a research design to develop an innovative, intercontinental, multidisciplinary integrated framework for an Ethical AI Standard. The most innovative aspect of this research design is a targeted programme to select, analyse, cross-examine, integrate and expand inputs and debates from twenty-one tech corporations, government organisations, civil society actors, and the analysis of debates generated on social media platforms by the general public, globally. This can be achieved by investigating in-depth ethicopolitical judgements, empirical statements, and practical guidelines produced in public AI policy documents, interviews with experts and practitioners, and debates circulating in the digital public domain. To create and implement ethical and legitimate AI governance, stakeholders need to be confronted with their own and others’ ethicopolitical visions and discourses. They must also be confident that the researchers understand the practicalities of delivering advanced AI technology and the concerns of individuals and organisations requiring privacy and transparency in government and corporate policy in this area. The overall objective is to investigate the ethical and political visions of corporate, governmental, and civil society organisations, and the general public and cross-examine these with the direct engagement of interview participants.

¹ Metz C., Satariano A. (2020) An Algorithm that Grants Freedom, or Takes It Away. *The New York Times*. Feb. 6. URL: <https://www.nytimes.com/2020/02/06/technology/predictive-algorithms-crime.html> (accessed: 27.02.2021).

With global, cross-sector, specialised, and general population input, this research design is set up to produce an integrated framework for Ethical Artificial Intelligence. Such a quest is the Holy Grail of technology ethics because of the high stakes involved in the use and abuse of Artificial Intelligence, which has critical consequences for humanity's future [Bostrom, 2014; Floridi, 2015; Harari, 2016]. The rapid development of AI and its application in fields as diverse as medical surgery, autonomous cars, and military robots, together with all-purpose use simulations of machine learning, has caused growing concerns about the unknown impact of AI in an anarchic world characterised by secretive commercial and nation-state competition [Kaplan, 2015; Acemoglu, Restrepo, 2018]. Artificial intelligent machines are advanced software systems. The questions are who designs and is in charge of these systems, who controls, regulates, and can have the data to intervene in time when AI is not serving the purpose with which it was designed [Garfinkel, Matthews, Shapiro, Smith, 2017]. Human values must be able to shape this future, and this future has to include everyone. By searching for AI, humanity is also searching for the best future for a human species capable of governing AI and developing an AI that displays the emotional and social intelligence to work with humans. Above all, we need an AI that compensates for rather than exploits human limitations because it understands blind spots in human cognition, memory, judgement, and attention, even empathy [van Dijk, 2014].

There are already remarkable visions of training AI to predict how a human would punish AI, when it ethically deviates. With rapid advancements in natural language translation, voice recognition, and a massive amount of computational time and space, whereby AI breathes, as it trains on human text, humans are often confused, feeling that the AI is human. Accordingly, there is concern that AI will bring a sense of loss: the uniqueness of being human. When this AI comes forward to interact with humans, it must go ahead with human values. We need to consider what value re-alignment is required in this partnership [Floridi, 2018]. What should humanity want from AI's future, in an era when machines will change human behaviour as never before?². Although humanity still understands little about how children are learning and have made little progress on the workings of human consciousness, there is nevertheless a pervasive use of AI that is unregulated, under little control and confronts legislation that is too slow for the accelerated sped up pace with which AI is evolving³. The symptoms of this unbounded acceleration are already in plain sight:

- Fake news and disinformation architectures which pose risks of populism, radicalism, violent extremism together with algorithmic interference⁴ [Sumpter, 2018];
- Gender, race, class, and other algorithmic bias [O'Neil, 2016; Chouldechova, 2017];
- Emerging issues in employment, health, education;

² Guillén M., Reddy S. (2018) We Know Ethics Should Inform AI. But Which Ethics? *World Economic Forum*. 26 July. URL: <https://www.weforum.org/agenda/2018/07/we-know-ethics-should-inform-ai-but-which-ethics-robotics/> (accessed: 27.02.2021).

³ The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation. (2018) URL: <https://www.repository.cam.ac.uk/bitstream/handle/1810/275332/1802.07228.pdf?sequence=1> (accessed: 27.02.2021).

⁴ Ong J. C., Cabañas J. V. A. (2017) Architects of Networked Disinformation: Behind the Scenes of Troll Accounts and Fake News Production in the Philippines (Public Report). URL: <https://newtontechfordev.com/wp-content/uploads/2018/02/ARCHITECTS-OF-NETWORKED-DISINFORMATION-FULL-REPORT.pdf> (accessed: 27.02.2021).

- Future of work, quantification, recruitment bias, digital labour and gig economy [Codagnone, Karatzogianni, Matthews, 2018];
 - Data justice, whistle-blowing, legal studies, digital rights, data inequality⁵ [Hintz, Dencik, Wahl-Jorgensen, 2018]; and
 - Superintelligence what Bostrom, Dafoe and Flynn⁶ call ‘mind crime prevention’, ensuring that advanced AI is governed in such a way that maltreatment of sentient digital minds is avoided or minimized.

Equally, we may see future potential resistance to Artificially Intelligent machines, which would predictably see future AI-resisting social movements and non-state actors taking digital activism and cyberconflict to unimaginable new heights.

AI is now a top research priority. In the past few years, there has been a proliferation of reports on AI from governmental and other organisations⁷. Let us consider the two most recent European commission responses to critical issues arising from AI with the publication of *Artificial Intelligence: A European Perspective*⁸, *A Draft Ethics Guidelines for Trustworthy AI*⁹; *A Definition of AI: Main Capabilities and Disciplines*¹⁰.

⁵ César J., Debussche J., van Asbroeck B. (2017) White Paper—Data Ownership in the Context of the European Data Economy: Proposal for a New Right. *Bird & Bird*. February. URL: <https://www.twobirds.com/en/news/articles/2017/global/data-ownership-in-the-context-of-the-european-data-economy> (accessed: 27.02.2021).

⁶ Bostrom N., Dafoe A., Flynn C. (2018) Public Policy for Superintelligent AI: A Vector Field Approach. URL: <https://nickbostrom.com/papers/aipolicy.pdf> (accessed: 27.02.2021).

⁷ Executive Office of the President National Science and Technology Council Committee on Technology (2016) Preparing for the Future of Artificial Intelligence. October. URL: <https://info.publicintelligence.net/WhiteHouse-ArtificialIntelligencePreparations.pdf> (accessed: 26.02.2021); UK Government Office for Science (2015) Artificial Intelligence: Opportunities and Implications for the Future of Decision Making. URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/566075/gs-16-19-artificial-intelligence-ai-report.pdf (accessed: 26.02.2021); UK House of Commons Science and Technology Committee (2016) Robotics and Artificial Intelligence. Fifth Report of Session 2016–17. URL: <https://publications.parliament.uk/pa/cm201617/cmselect/cmsctech/145/145.pdf> (accessed: 26.02.2021); European Economic and Social Committee (2017) Artificial Intelligence — The Consequences of Artificial Intelligence on the (Digital) Single Market, Production, Consumption, Employment and Society (Own-Initiative Opinion). URL: <https://www.eesc.europa.eu/en/our-work/opinions-information-reports/opinions/artificial-intelligence> (accessed: 27.02.2021); European Parliament Policy Department (2016) European Civil Law Rules in Robotics. URL: [EUROPEAN CIVIL LAW RULES IN ROBOTICS](http://europarl.europa.eu) (europarl.europa.eu) (accessed: 27.02.2021); Council of Europe Committee of experts on internet intermediaries (2017) Study on the Human Rights Dimensions of Automated Data Processing Techniques (in Particular Algorithms) and Possible Regulatory Implications. URL: <https://rm.coe.int/study-hr-dimension-of-automated-data-processing-incl-algorithms/168075b94a> (accessed: 27.02.2021); Ministry of Economic Affairs and Employment of Finland (2017) Finland's Age of Artificial Intelligence. Turning Finland into a leading country in the application of artificial intelligence. Objective and recommendations for measures. URL: http://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/160391/TEMrap_47_2017_verkkojulkaisu.pdf?sequence=1&isAllowed=y (accessed: 27.02.2021); France Intelligence Artificielle (2017) Rapport de Synthèse – France IA. URL: https://www.economie.gouv.fr/files/files/PDF/2017/Rapport_synthese_France_IA_.pdf (accessed: 27.02.2021); Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions (2018) Artificial Intelligence for Europe. URL: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2018%3A237%3AFIN> (accessed: 27.02.2021); European Commission's High-Level Expert Group on Artificial Intelligence (2018) Draft Ethics Guidelines for Trustworthy AI. 18 December. URL: [Draft Ethics guidelines for trustworthy AI | Shaping Europe's digital future \(euroopa.eu\)](https://ec.europa.eu/trustworthy-ai/-/shaping-europe-s-digital-future/(euroopa.eu)) (accessed: 27.02.2021); European Group on Ethics in Science and New Technologies (2018) Statement on Artificial Intelligence, Robotics and 'Autonomous' Systems. URL: https://ec.europa.eu/research/ege/pdf/ege_ai_statement_2018.pdf (accessed: 27.02.2021); Deloitte (2017) Study on Emerging Issues of Data Ownership, Interoperability, (Re-)Usability and Access to Data, and Liability. URL: http://ec.europa.eu/newsroom/dae/document.cfm?doc_id=51486 (accessed: 27.02.2021).

⁸ European Commission Science Hub (2018) Artificial Intelligence: A European Perspective. URL: <https://ec.europa.eu/jrc/en/publication/artificial-intelligence-european-perspective> (accessed: 27.02.2021).

⁹ European Commission's High-Level Expert Group on Artificial Intelligence (2018) Draft Ethics Guidelines for Trustworthy AI. 18 December. URL: [Draft Ethics guidelines for trustworthy AI | Shaping Europe's digital future \(europa.eu\)](https://ec.europa.eu/digital-single-market/en/news/draft-ethics-guidelines-trustworthy-ai-shaping-europe-s-digital-future) (accessed: 27.02.2021).

¹⁰ European Commission's High-Level Expert Group on Artificial Intelligence (2019) A Definition of AI: Main Capabilities and Disciplines. URL: <https://www.aepd.es/sites/default/files/2019-12/ai-definition.pdf> (accessed: 27.02.2021).

The AI European perspective report differentiates between individual and collective implications of AI: autonomy, identity, dignity, privacy, and data protection at the individual level. Further, it recognises that AI dramatically affects the societal level: fairness and equity, responsibility, accountability and transparency, privacy in terms of surveillance/datafication, democracy and trust, and collective identity and good life. In reviewing key ethical and social issues in AI, it identifies two new rights:

- (1) The right to meaningful human contact, whereby every person may refuse to be cared for by a robot, and robots should respect humans' autonomy in decision-making;
- (2) The right to refuse being profiled, tracked, measured, analysed, coached, or manipulated.

The authors also prescribe responsible AI design, which engages critically with civil society, establishes multi-stakeholder fora to promote such public debate translating outcomes to strategies for AI enforcing ethical and social values, and the design practice to address potential sources of the AI system from selection of team, to labelling and training data evaluation of outputs, and assessments of outcomes¹¹. Given the significance of these findings, it is not unreasonable to anticipate these issues making up additional protocols to the European Convention of Human Rights in the near future.

Where are we with AI ethical governance at present? Arguably we are at an elementary stage, and this area of research requires urgent advancement. Accordingly, the 'Draft Ethics Guidelines for Trustworthy AI' aimed at maximising the benefits of AI while minimising its risks, advocating that a human-centric approach to AI is needed to create 'Trustworthy AI', respecting fundamental human rights, ensuring an 'ethical purpose', and asserting that it should be technically robust. The Guidelines operationalise the requirements of ethical purpose and robustness, providing us with a concrete assessment list. This list includes the incorporation of Trustworthy AI from the earliest design phase: accountability, data governance, design for all, governance of AI autonomy (human oversight), non-discrimination, respect for human autonomy, respect for privacy, robustness, safety, and transparency. The key guidance embraces technical and non-technical methods to implement while keeping in mind ethical considerations when recruiting the team building the system, the system itself, the testing environment, and potential applications. Additionally, stakeholders (customers, employees) should have transparent and proactive information regarding the AI's capabilities and limitations, ensuring traceability.

Suppose we want to promote ethical AI policies and practices AI to the level of strategic goals for leading organisations. In that case, AI should be part of the organisation's culture, embedded in deontological chapters or codes of conduct, ensuring stakeholders' inclusion in the AI development and diversity in the team producing it. This enables us to foresee training and education in Trustworthy AI and ensure a specific process for accountability governance¹². Assessing Trustworthy AI includes

¹¹ European Commission Science Hub (2018) Artificial Intelligence: A European Perspective. URL: <https://ec.europa.eu/jrc/en/publication/artificial-intelligence-european-perspective> (accessed: 27.02.2021).

¹² European Commission's High-Level Expert Group on Artificial Intelligence (2018) Draft Ethics Guidelines for Trustworthy AI. 18 December. URL: [Draft Ethics guidelines for trustworthy AI | Shaping Europe's digital future \(europa.eu\)](https://ec.europa.eu/jrc/en/publication/draft-ethics-guidelines-for-trustworthy-ai-shaping-europe-s-digital-future) (accessed: 27.02.2021).

accountability, data governance, design for all, governing AI autonomy, non-discrimination, respect for privacy, robustness, reliability, reproducibility, and accuracy through data usage and control fall-back plan, safety, transparency, and traceability¹³.

There are important choices to be made. Geopolitically, this approach differentiates European AI Ethics from the unsustainable and undemocratic development of AI involving massive surveillance and control of populations in the Middle East and North Africa, and East Asia. Also, in relation to oligopolistic algorithmic governance by tech companies without significant governmental regulatory commitment to democracy and trust in North America. The World Intellectual Property Organisation (WIPO) Technology Trends 2019 identifies deep learning as the fastest growing technique with an increase of 175 % between 2013 and 2016¹⁴. Crucially, deep learning is ‘black box’ AI, which relies on neural networks, in contrast to ‘white box’ AI, where all the code lines are explicit. Companies represent twenty six of the top thirty applicants, and it is striking that just four are university or public research organisation. IBM tops the list, followed by Microsoft, and out of the top twenty, twelve are based in Japan, three are from the US, and two are from China. The report identifies the geographical origin of the university and public research organisations in the top 500 as China, US, Korea, Taiwan, Europe, Japan, Russian Federation, Saudi Arabia, in that order with China clearly by far dominating patents in that field¹⁵.

A Global Ethical Problem: In Europe, Ethics and AI involve significant ethical judgments, empirical statements, and practical guidelines, which rely heavily on the direct adaptation of what we could call ‘social-democratic humanism’. To the West, the United States is a prime example of drawing from a ‘neo-liberal humanism’, whereby the individual knows best, the customer will decide, and the company will create a product which the customer does not even know that they already want. Here, dataism and techno-utopianism are the resulting principles. Humans will accept all, as long as they can stay in the data flow and take advantages of AI leaps. Those particular humans that can enhance their body and life with AI will evolve as a new elite of superhumans, and those that are rendered obsolete and useless by AI will be left behind and out-evolved. The era of the masses is over. These alarming Darwinist ethics are drawing from ‘evolutionary humanism’ ideologies of the past (eugenics is a prime example). Harari [2016] inspired this line of argumentation in *Homo Deus*, where he traces some continuities and discontinuities in the AI ethicopolitical visions. These have become the philosophical departure point of this project.

To integrate and produce an Ethical AI Standard, this research design is set up to answer the following key research question: What are the competing AI ethicopolitical visions of key actors in the field of AI?

Designing for AI Ethics Research

With this research question in mind, this research design experiment is set up to analyse, trace, evaluate, select, integrate and expand diverse and fragmented ethico-

¹³ Ibidem.

¹⁴ World Intellectual Property Organisation (2019) WIPO Technology Trends 2019 — Artificial Intelligence. URL: <https://www.wipo.int/publications/en/details.jsp?id=4386> (accessed: 27.02.2021).

¹⁵ Ibid.: 32.

political visions of AI, considering the proposals envisaged by the European AI ethicists above. The project utilises directly European recommendations because they are crucial for any Ethical AI Standard. Still, it does so by engaging multistakeholder fora by interviewing key actors, engaging with civil society, and promoting public debate beyond the European Union countries. Furthermore, it discusses the European AI ethical framework with key stakeholders, governments, corporations, civil society actors, and the global public. It anticipates an open and reflexive critique that will take a potential project forward. The research design offered here relies on the following research techniques and objectives:

Project Objective 1 (PO1): Maps key AI ethicopolitical frameworks in circulation by the 21 key actors. This will involve the collection of AI policy documents produced by three sets of key players: seven tech corporations (Google, Amazon, Facebook, Apple, Microsoft, Tesla, and Alibaba), seven governmental organisations (China, Japan, United States, European Union, Australia, India, and South Africa), seven civil society actors (The Partnership on AI, Open AI, Association for the Advancement of Artificial Intelligence, European Association for Artificial Intelligence, Future of Life Institute, Society for the Study of Artificial Intelligence and the Simulation of Behaviour, and the Machine Intelligence Research Institute).

(PO2): Investigates ethicopolitical visions on AI across seven tech corporations. Examines ethicopolitical visions by seven tech corporations and juxtaposes these with the findings from PO1, asking interview participants to compare their views concerning PO1. The specific tech corporations (dubbed the internet oligopoly with the acronym GAFAM: Google, Amazon, Facebook, Apple, Microsoft) are chosen because they have been recently embattled in ethical issues publicly and extensively. Alibaba is selected because of the sheer scale of AI application involved in its trade. Tesla is examined because it has ranked as the world's best-selling plug-in passenger car manufacturer and works across several technological innovation domains applying AI. Crucially, Tesla founder Elon Musk has repeatedly advocated strong AI regulation in public.

The GAFAM tech corporations have been recently embattled in ethical issues publicly and extensively. Examples of why Google is chosen involves recent reports of the crowd workers outsourced to support a contract the company had with the US military on drones and the extensive ethical issues this brought up with employees with the company, resigning and demanding adherence to the company motto 'Do No Evil'. Google owns YouTube, which has also been controversial in terms of content moderation in relation to online radicalisation videos appearing next to advertisements, with companies and governments withdrawing advertising from the platform. Facebook has been embroiled in the Cambridge Analytica scandal, disinformation and potential impact during elections around the world. Subsequently, the slow and inadequate response the company rendered against its public critics, in relation to privacy, together with a ramification of its advertising practices, and alleged interference on its platform, potentially influencing the election and referendum results in several countries (e.g., the US 2016 Elections, Brexit), together with the ultimate 'hacking' of democratic institutions. The closer integration of WhatsApp and Instagram, which the company acquired, has caused widespread public criticism and an array of ethical issues relating to children and youth's use of their platforms in particular.

Amazon, Apple, Microsoft are tech companies that have also involved much in the development of advanced software systems and are considered influential players in mobile and desktop applications and hardware. Alibaba and Tesla are included here for their significant record in this domain. Alibaba is chosen because of the sheer scale of AI applications involved in its trade. During 'Singles Day' Alibaba processed 325,000 orders per second through pop-up stores selling products fitted with Virtual Reality mirrors, using an AI fashion consultant matching items. One day, it sold 25 billion dollars' worth of goods¹⁶. Tesla is chosen for this specific reason: Tesla founder Elon Musk has repeatedly advocated strong AI regulation in public for the past decade. 'It needs to be a public body that has insight and then oversight to confirm that everyone is developing AI safely. This is extremely important. I think the danger of AI is much greater than the danger of nuclear warheads by a lot and nobody would suggest that we allow anyone to build nuclear warheads if they want. That would be insane'¹⁷.

(PO3): *Investigates ethicopolitical visions on AI across seven governments.* Examines ethicopolitical visions by seven governmental organisations and asks interview participants to compare their views in relation to PO2. Seven governmental organisations are investigated (China, Japan, the United States; the European Union — focus on Germany France and the Nordic-Baltic Eight (NB8)¹⁸; Australia, India and South Africa). AI is ultra-nationalised and governments are pressured into the impossible position to develop AI policies that are competitive while protecting citizen rights (transparency, accountability, privacy, equal treatment, non-discrimination, mitigation of harmful impacts). In recent years, these actors have released AI visions. Interviews with policy-makers in government will involve direct questions about the AI ethicopolitical visions expressed by actors in PO2.

Although the first AI patent filings were made in Japan in the 1980s, the field has been overtaken by China and the United States. Since 2014, China has been the leader in a number of first patents filed. In 2017, the State Council announced the 'Next Generation AI Development Plan' with the ambition of becoming the world's primary innovation centre by 2030, followed up by a 'Three Year Plan to Promote the Development of the New-Generation AI Industry'¹⁹. In the United States, three reports were released in 2016: 'Artificial Intelligence, Automation, and the Economy'; 'Preparing for the Future of Artificial Intelligence'; and 'The National Artificial Intelligence Research and Development Plan', while in 2018, a Select Committee on Artificial Intelligence was announced²⁰. The European Union is going to be researched as an intergovernmental organisation, however, with the understanding the specific countries are going to be investigated in more depth, Germany, France, particularly the Nordic-Baltic Eight (NB8), because they made a joint statement in May 2018 to enhance

¹⁶ European Commission Science Hub (2018) Artificial Intelligence: A European Perspective. P.60. URL: <https://ec.europa.eu/jrc/en/publication/artificial-intelligence-european-perspective> (accessed: 27.02.2021).

¹⁷ Young A. (2018) Musk says AI 'More Dangerous Than Nukes' — Expert Stays Optimistic *SecurityBrief.eu*. 13 March. URL: [Musk says AI 'more dangerous than nukes'-expert stays optimistic \(securitybrief.eu\)](https://www.securitybrief.eu/musk-says-ai-more-dangerous-than-nukes-expert-stays-optimistic.html) (accessed: 27.02.2021).

¹⁸ Denmark, Estonia, Finland, Iceland, Latvia, Lithuania, Norway, and Sweden.

¹⁹ World Intellectual Property Organisation. (2019) WIPO Technology Trends 2019 — Artificial Intelligence. P. 127. URL: <https://www.wipo.int/publications/en/details.jsp?id=4386> (accessed: 27.02.2021).

²⁰ Ibid.: 126.

access for data for AI, stating that they want to ‘avoid unnecessary regulation that could get in the way of this fast-developing field’²¹.

Apart from these top players, Australia, India, and South Africa are chosen to provide a more intercontinental perspective. With headlines such as ‘Australia lags on AI, automation’²² and ‘Australian needs to embrace automation or risk missing a 2.2 trillion-dollar boom’²³, Australia is a case worth studying further. India is chosen because of their #AIforAll approach. In their ‘National Strategy for Artificial Intelligence #AIforAll’, whereby #AIforAll will focus on harnessing collaborations and partnerships, and aspires to ensure prosperity for all. Thus, #AIforAll means technology leadership in AI for achieving the greater good’²⁴. In Africa, we will focus on South Africa, together with a broader interest in understanding AI for development and organisations such as Machine Intelligence Institute of Africa²⁵. In addition, events such as AI for Good Global Summit and United Nations AI conference are the type of events where actors’ interplay can be observed, and potential fieldwork interviews can be conducted.

(P04): *Investigates ethicopolitical visions on AI across seven AI-specialised civil society organisations.* Examines ethicopolitical visions by seven AI-specialised civil society organisations and asks interview participants to compare their views concerning P02 and P03. The current sample includes the following organisations: The Partnership on AI, Open AI, Association for the Advancement of Artificial Intelligence, European Association for Artificial Intelligence, Future of Life Institute, Society for the Study of Artificial Intelligence and the Simulation of Behaviour, and the Machine Intelligence Research Institute. As P01 kicks off mapping the AI policy environment and during fieldwork for P02 and P03, we will acquire further insights on which are the most relevant specialised organisations to investigate. The reason we are interested in civil society organisations specialised on AI and not generally, for example, in privacy, transparency, or digital rights organisations is because there are several areas of technical and policy expertise involved in this area, and we do require a sufficient level of specialisation to integrate insights and principles from.

(P05): *Investigates public perceptions of ethicopolitical visions identified in P01–P04.* Examines ethicopolitical visions of the 21 actors, and how they are received by the general population on social media platform debates. We will choose to collect public debates across Facebook (groups such as ‘Artificial General Intelligence’; ‘Artificial Gods’, ‘Real AGI’, ‘Artificial Intelligence and Deep Learning’), Twitter, Sina Weibo, as the most dominant globally, and online forums specialising on AI debates, such as ‘The Artificial Intelligence Forum’²⁶, ‘Ai dreams’²⁷, on ‘Reddit’²⁸, ‘Quora’s Artificial General

²¹ Ibid.: 127.

²² Australia lags on AI, automation. (2019) *Erpinnews*. October 24. URL: <https://erpinnews.com/australia-lags-on-ai-automation> (accessed: 27.02.2021).

²³ Dunn M. (2018) Australian needs to embrace automation or risk missing a 2.2 trillion boom. News.com.au. June 1. URL: <https://www.news.com.au/technology/innovation/inventions/australia-needs-to-embrace-automation-or-risk-missing-a-22-trillion-boom/news-story/23b2608dec515e3749601d46bac6143d> (accessed: 27.02.2021).

²⁴ Future of Life Institute. AI Policy — India. URL: <https://futureoflife.org/ai-policy-india/?cn-reloaded=1> (accessed: 27.02.2021).

²⁵ URL: <http://machineintelligenceafrica.org/> (accessed: 26.02.2021).

²⁶ For more details, see URL: <https://ai-forum.com/> (accessed: 27.02.2021).

²⁷ For more details, see URL: <https://aidreams.co.uk/> (accessed: 27.02.2021).

²⁸ For more details, see URL: <https://www.reddit.com/r/machinelearning> (accessed: 27.02.2021).

Intelligence'²⁹. The purpose is to conduct social network analysis and semantic analysis of what are the dominant actors, relations and debates in the digital public domain, what are the ethical judgements and empirical statements in circulation and particularly how the 21 actors we are interested in are received in those circles.

(PO6): *Produces an Integrated framework for an Ethical AI Standard (iExIST)*. The final work package will first select, integrate and expand the AI policy mapping from PO1. It will establish and synthesise the themes and principles informing the 21 actors plus public circulation ethical judgements, juxtapose these to empirical statements accordingly (PO2, PO3, PO4, PO5), integrate best practical guidelines, and then disseminate these findings to the actors interviewed to create a feedback loop of best approaches to produce the final framework for an Ethical AI standard.

Research Techniques

AI policy document analysis for PO1, using NVIVO, will enable us to collect, organise and analyse content from interviews, social media data, YouTube videos and web pages. In this way, we can describe and document data in a highly organised fashion, which will help both during critical multimodal discourse analysis and when data are shared after the research projects end. Collection and analysis of primary (reports, documents, legislation, policy assessments) and secondary (academic/other) materials, focusing on process-tracing each stakeholder's role in the evolving system of AI implementation. We will concentrate on three departments in each country: typically ministries of health/social security, education, and justice. In some countries, AI responsibilities are more dispersed and include departments dealing with media/culture, technical assurance, policing, and security.

Semi-structured interviews for PO2, PO3, PO4: Research interviews will be conducted with the primary stakeholders and their attitudes and beliefs regarding AI ethical and social issues. The type of questions will include beliefs (what people believe to be the case); attitudes (what people would prefer to be the case); behaviour (examples from their own experience as practitioners, policymakers, and activists). Interviewee attributes will be recorded and anonymised when this is necessary during consent and ethical issues emerging. Interviews are essential in establishing what our key actors think about the changing nature of AI in society and their more general attitudes towards current practice and procedure. While it is possible to obtain some of this information from policy documents, our emphasis on the interview will allow us to draw out the respondents at length regarding their thoughts on real world issues. It will enable us to ask open-ended questions and permits the respondent to talk more freely. It is appropriate for a project in which we wish to gather rich ethnographic about working with AI.

We also emphasise interviews because of our desire to undertake a degree of the process- tracing. This is to establish the decisions and attitudes that underpinned existing protocols and responses to particular AI issues in the respective countries. Within this focused inquiry, we will be able to reconstruct specific practical episodes based on the interview testimony and then compare accounts to give us a sophisti-

²⁹ For more details, see URL: <https://www.quora.com/What-is-artificial-general-intelligence-AGI> (accessed: 27.02.2021).

cated picture of particular regulatory phenomena. Expert interviews are especially appropriate to study regulation-building since they can illuminate hidden elements of social action that are not clear from analysis of political outcomes using documentary materials. As others have argued, the existing literature on AI is fragmented and predominantly focused on the formal, legal, and informational rather than social aspects of regulation. Our emphasis on semi-structured interviews distinguishes between formal and informal processes and seeks to unpick some everyday activity around AI implementation. This approach will also encourage the co-production of knowledge during the project and beyond.

Social Network Analysis (SNA) and semantic analysis for P05–P06: The examination of transnational debates surrounding the 21 actors, and debates on AI ethics relations to the offline world. Wasserman and Faust [1994] explain that SNA ‘provides a precise way to define important social concepts, a theoretical alternative to the assumption of independent social actors, and a framework for testing theories about structured social relationships’ [ibid.: 17]. SNA is appropriate for the analysis and the investigation of ‘kinship patterns, community structure, interlocking directorships and so forth’ [Scott, 2000: 2]. SNA supports examining different social entities or social units, including individual, corporate, or collective social units [Wasserman, Faust, 1994: 16—21]. Key social media platforms, such as Facebook, will be examined to understand the formation of online networks and coalitions, dominant actors and structural characteristics, and Twitter to investigate the evolution of discourses and real-time reactions to various discrete events and processes. For the collection, analysis, and visualisation of networks, actors, and debates on AI ethics, the project will deploy the following tools: NodeXL visual representations and analytics for Twitter [Hansen, Schneiderman, Smith, 2010]; Netvizz, for data collection and extraction, an application tool allowing the export of data in file formats from different sections of the Facebook social networking service [Rieder, 2013], and Gephi open-source network graph and analysis tool [Cherven, 2015] for analysis and visualisation. These digital methods efficiently support the research objectives of the study with no need for engaging more complicated statistical and analytical tools which often require researchers with a rare social data science skillset (social science disciplinary background with the ability to use Python, R, UCINet).

Conclusion

The selection of a universal ethical standard on AI is the Holy Grail in this research area. Offering a holistic usable answer to this problem will benefit several disciplines, and stakeholders in global interdisciplinary academic and transnational practitioners fora. This research design proposed here has its high gains and high risks. First, the intercontinental scope of the study. It involves the collection, integration and expansion of material for a new framework for AI ethics. Although it might not produce the ultimate answer of the absolute standard, it will generate new data, findings, and advanced detailed recommendations taking AI ethics to new lines of inquiry. Second, data collection. It is a recognised possibility that we will not have access to specific individuals to interview in the actors already identified. This risk could be addressed by relying on researchers’ networks to locate interview participants. Third, it involves

fieldwork around the world, which contains a certain level of risk for researchers; however, the countries specified are not presently volatile to political developments and social unrest. Interviews can also be conducted online to mitigate travel risk in relation to the COVID-19 pandemic developments. Last, the fieldwork's operational success relies on solid research management to co-ordinate the five work packages reflecting the POs, so that they feed seamlessly into each other, enhanced by frequent reviews, peer assessment, and reports, as well as mentoring, collegiality, and sensitivity in managing the research teams involved.

References (Список литературы)

- Acemoglu D., Restrepo R. (2018) The Race between Man and Machine: Implications of Technology for Growth, Factor Shares, and Employment. *American Economic Review*. Vol. 108. No. 6. P. 1488—1542. <https://doi.org/10.1257/aer.20160696>.
- Bostrom N. (2014) Superintelligence: Paths, Dangers, Strategies. Oxford: Oxford University Press.
- Cherven K. (2015) Mastering Gephi Network Visualization. Birmingham (UK): Pack Publishing Limited.
- Chouldechova A. (2017) Fair Prediction with Disparate Impact: A Study of Bias in Recidivism Prediction Instruments. *Big Data*. Vol. 5. No. 2. P. 153—163. <https://doi.org/10.1089/big.2016.0047>.
- Codagnone C., Karatzogianni A., Matthews J. (2018) Platform Economics: Rhetoric and Reality in the “Sharing Economy”. Bingley: Emerald Publishing Limited.
- Floridi L. (ed.) (2015) The Online Manifesto: Being Human in a Hyperconnected Era. Cham: Springer.
- Floridi L. (2018) Soft Ethics and the Governance of the Digital. *Philosophy & Technology*. No. 31. P. 1—8. <https://doi.org/10.1007/s13347-018-0303-9>.
- Garfinkel S., Matthews J., Shapiro S. S., Smith J. M. (2017) Toward Algorithmic Transparency and Accountability. *Communications of the ACM*. Vol. 60. No. 9. P. 5. <https://doi.org/10.1145/3125780>.
- Hansen D., Schneiderman B., Smith M. A. (2010) Analyzing Social Media Networks with NodeXL: Insights from a Connected World. San Francisco (US): Elsevier Science & Technology.
- Harari Y. N. (2016) Homo Deus: A Brief History of Tomorrow. London: Vintage Publishing.
- Hintz A., Dencik L., Wahl-Jorgensen K. (2018) Digital Citizenship in a Datafied Society. Oxford: Polity Press.
- Jobin A., Ienca M., Vayena E. (2019) The Global Landscape of AI Ethics Guidelines. *Nature Machine Intelligence*. No. 1. P. 389—399. <https://doi.org/10.1038/s42256-019-0088-2>.

- Kaplan J. (2015) Humans Need Not Apply: A Guide to Wealth and Work in the Age of Artificial Intelligence. New Haven: Yale University Press.
- O'Neil C. (2016) Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. London: Penguin Books Ltd.
- Redden J., Dencik L., Warne H. (2020) Datafied Child Welfare Services: Unpacking Politics, Economics and Power. *Policy Studies*. Vol. 41. No. 5. P. 507—526. <https://doi.org/10.1080/01442872.2020.1724928>.
- Rieder B. (2013) Studying Facebook via Data Extraction: The Netvizz Application. In: *Proceedings of the 5th Annual ACM Web Science Conference (WebSci'13)*. May. P. 346—355. <https://doi.org/10.1145/2464464.2464475> (accessed: 27.02.2021).
- Sanchez-Monedero J., Dencik L., Edwards L. (2020) What Does it Mean to 'Solve' the Problem of Discrimination in Hiring? Social, Technical and Legal Perspectives from the UK on Automated Hiring Systems. Paper presented at Conference on Fairness, Accountability, and Transparency (FAT*'20). Barcelona, Spain. 27—30 January. URL: <https://arxiv.org/abs/1910.06144> (accessed: 27.02.2021).
- Scott J. (2000) Social Network Analysis: A Handbook (2nd ed.). London: Sage.
- Sumpter D. (2018) Outnumbered: From Facebook and Google to Fake News and Filter-bubbles — The Algorithms that Control our Lives. London: Bloomsbury Publishing PLC.
- Wasserman S., Faust K. (1994) Social Network Analysis: Methods and Applications. Cambridge: Cambridge University Press.
- van Dijck J. (2014) Datafication, Dataism and Dataveillance: Big Data between Scientific Paradigm and Ideology. *Surveillance & Society*. Vol. 12. No. 2. P. 197—208. <https://doi.org/10.24908/ss.v12i2.4776>.

DOI: [10.14515/monitoring.2021.1.1793](https://doi.org/10.14515/monitoring.2021.1.1793)



N. Klowait, M.A. Erofeeva

THE RISE OF INTERACTIONAL MULTIMODALITY IN HUMAN-COMPUTER INTERACTION

For citation:

Klowait N., Erofeeva M.A. (2021) The Rise of Interactional Multimodality in Human-Computer Interaction. *Monitoring of Public Opinion: Economic and Social Changes*. No. 1. P. 46–70. <https://doi.org/10.14515/monitoring.2021.1.1793>.

Правильная ссылка на статью:

Кловайт Н., Ерофеева М. А. Мультимодальный поворот в исследованиях взаимодействия человека и компьютера // Мониторинг общественного мнения: экономические и социальные перемены. 2021. № 1. С. 46—70. <https://doi.org/10.14515/monitoring.2021.1.1793>. (In Eng.)

THE RISE OF INTERACTIONAL MULTIMODALITY IN HUMAN-COMPUTER INTERACTION

*Nils KLOWAIT^{1,2,3} — Research Fellow at the Center for Innovative Social Research; Senior Research Fellow at the International Center for Contemporary Sociological Theory; Researcher
E-MAIL: nils.klowait@gmail.com
<https://orcid.org/0000-0002-7347-099X>*

*Maria A. EROFEEVA^{1,2,3} — Cand. Sci. (Soc.), Researcher at the Center for Sociological Research; Senior Research Fellow at the International Center for Contemporary Sociological Theory; Researcher
E-MAIL: erofeeva-ma@universitas.ru
<https://orcid.org/0000-0002-0874-5272>*

МУЛЬТИМОДАЛЬНЫЙ ПОВОРОТ В ИССЛЕДОВАНИЯХ ВЗАИМОДЕЙСТВИЯ ЧЕЛОВЕКА И КОМПЬЮТЕРА

*КЛОВАЙТ Нильс — научный сотрудник Центра перспективных социальных исследований, Российская академия народного хозяйства и государственной службы при Президенте РФ, Москва, Россия; старший научный сотрудник Международного центра современной социологической теории, Московская высшая школа социальных и экономических наук, Москва, Россия; эксперт, Лаборатория геймификации Сбера, Москва, Россия
E-MAIL: nils.klowait@gmail.com
<https://orcid.org/0000-0002-7347-099X>*

*ЕРОФЕЕВА Мария Александровна — кандидат социологических наук, научный сотрудник Центра социологических исследований, Российская академия народного хозяйства и государственной службы при Президенте РФ, Москва, Россия; старший научный сотрудник Международного центра современной социологической теории, Московская высшая школа социальных и экономических наук, Москва, Россия; эксперт, Лаборатория геймификации Сбера, Москва, Россия
E-MAIL: erofeeva-ma@universitas.ru
<https://orcid.org/0000-0002-0874-5272>*

¹ The Russian Presidential Academy of National Economy and Public Administration, Moscow, Russia
² The Moscow School of Social and Economic Sciences, Moscow, Russia
³ The Sber Gamification Lab, Moscow, Russia

Abstract. The field of human-computer interaction (HCI) investigates the intersection between the design of devices and user practices. From an early focus on interaction modeling based on psychological experiments, the field has since experienced a shift towards the study of how actual humans interact with autonomous devices. The field became conducive to ethnographic,

Аннотация. Предметом исследований взаимодействия человека и компьютера является взаимосвязь между дизайном устройств и пользовательскими практиками. Изначально взаимодействие «человек — компьютер» моделировалось на основании психологических экспериментов, однако со временем в этой области произошел сдвиг в сторону изучения того, как

observational and videographic studies of human-device interaction. Conversation-analytic HCI became possible. That said, this new wave of researchers was never truly able to dethrone the psychological common sense of the field. With recent developments in both the technical-sensorial capabilities and outward actuational range of embodied virtual agents, the field of HCI has once again returned to the question of the sequential unfolding of the interaction between users and intelligent agents, and the multimodal interactional repertoire that is deployed throughout. This review will highlight the situational orientation of high-impact research in the field, and relate it to the cotemporaneous development of ethnomethodological and conversation analytic frameworks.

Keywords: human-computer interaction, embodied conversational agents, conversation analysis, multimodality, interactive resources

Acknowledgments. The article was prepared in the framework of a research grant funded by the Ministry of Science and Higher Education of the Russian Federation (grant ID: 075-15-2020-908). The article was prepared in cooperation with the Sber (ex.-Sberbank) Gamification Lab.

Introduction

The field of human-computer interaction (HCI) is a strongly multidisciplinary endeavor that focuses on questions surrounding the engagement of humans with various kinds of interfaces. HCI has historically been the domain of engineering and computer science. As such, it has always had an eminently practical concern with the design of useable, functional, properly integrated technologies. That said, over its brief history, it has seen a number of theoretical interventions by various disciplines.

люди используют автономные устройства в реальных условиях. Начали проводиться качественные исследования (этнография, включенное наблюдение, видеоанализ), основанные в том числе на методологии конверсационного анализа. Благодаря недавним разработкам в области сенсорных и коммуникативных способностей воплощенных виртуальных агентов (аватаров) исследователи вернулись к вопросу о последовательном развертывании взаимодействия между пользователями и аватарами, а также о доступном мультимодальном интерактивном репертуаре последних. В статье анализируются актуальные направления исследований взаимодействия человека и компьютера в контексте развития конверсационного анализа.

Ключевые слова: взаимодействие человека и компьютера, воплощенные разговорные агенты, конверсационный анализ, мультимодальность, интерактивные ресурсы

Благодарность. Статья подготовлена в рамках гранта, предоставленного Министерством науки и высшего образования Российской Федерации (№ соглашения о предоставлении гранта: 075-15-2020-908). Статья подготовлена в сотрудничестве с Лабораторией геймификации Сбера.

One of these interventions occurred in the mid-1990s: the discipline of *conversation analysis* then attempted, with only moderate success, to establish itself as a resource for the study of naturally-occurring device use, and as a toolset for the development of more humanlike interactional technologies [see Klowait, 2018b]. This review will investigate both the reasons for the lack of universal adoption of conversation analysis as the go-to conceptual toolset for HCI, as well as the emerging trends that point to an upcoming conversation-analytic renaissance in the field.

Conversation Analysis

In the middle of the 20th century, a new approach to the analysis of human interaction emerged: conversation analysis (CA). Spearheaded by thinkers like H. Sacks, G. Jefferson, and E. A. Schegloff, it expanded its progenitors — E. Goffman and H. Garfinkel — in two substantial ways. On the one hand, it made the profound — yet unsystematic — insights of Goffman commensurable with a methodologically robust framework of multi-generational analytic practice. It moved pointedly away from newspaper vignettes, implicitly known cultural codes, and ‘illustrative examples’, instead focusing on gathering and analyzing naturalistic (i.e. unstaged, unrehearsed, unscripted) data in the form of audio/video recordings.

CA equally moved away from Garfinkel’s early focus on the irreducible indexicality of the single situation, making first attempts to gain generalized systematic insights about typical structures employed by interactants for practical purposes: how hotline operators accomplish the smooth conclusion of a call, how speakers project epistemic authority, how interactants refuse birthday party invitations without major disruptions to conversational flow [for an introduction see Stokoe, 2018]. Regardless of the concrete interaction type (face-to-face dyads, teleconferencing, multi-party lecture environments, etc.) or modality (talk, prosody, non-verbal, etc.), CA’s focus remained squarely on the participants’ methods of getting things done, and the multitude of resources employed for this purpose.

The resources, meanwhile, turned out to be unexpectedly minute and subtle: for example, we habitually display an orientation to notable silences during our interactions with others. Imagine a two-person conversation, where the first speaker asks, ‘Will you come to my birthday party?’, followed by a response by the second speaker. If there is even a one-second silence between question and response, the first speaker may already expect a rejection. The notable silence, in other words, is something all parties orient to — we know that everybody knows that a long silence is problematic here. Moreover, there is something more problematic about rejecting a birthday invitation than accepting it, and all parties tend to be aware of this, making interactional elements like silences, coughs, gaze shifts, interruptions, key elements of analytic focus for CA. Imagine a situation where a birthday invitation is followed by an immediate ‘no’, without flourish or elaboration; the general fact that this hypothetical is extraordinary hints at the implicit normative structuring going on here.

A classic example from early conversation analysis illustrates how speakers appropriate the implicit norms of conversation to achieve their practical goals all the while orienting to these very norms [Sacks, Jefferson, 1992: 3]:

(2)¹

A: This is Mr. Smith may I help you

B: Yes, this is Mr. Brown

(3)

A: This is Mr. Smith may I help you

B: I can't hear you.

A: This is Mr. Smith.

B: Smith.

The above are transcripts from different telephone conversations between a client and an emergency psychiatric hospital; A is the staff member, B is the client. Suppose that (2) and (3) are conversational sequences that frequently appear in the very beginning of these kinds of phone calls. Suppose further that this is the first interaction between A and B. If one were to create an interpretational gloss of (2) and (3), respectively, it would likely be something akin to 'In (2), A and B introduce themselves to each other' and 'In (3), there seems to be a problem with the connection that prevents B from hearing A's introduction'. With such a gloss, nothing truly remarkable seems to occur.

However, if — upon noticing that, in (3), B's half of the introduction does not occur — it is possible to view (3) as an example of an **intentional** interactional device employed by B to avoid a name-based introduction on their part. We can ask ourselves: *apart from its explicit meaning, what does 'I can't hear you' achieve, interactionally?* We could then ask ourselves *how* this sequence is employed and what it says about the nature of conversation in general. Sacks [Sacks, Jefferson, 1992] suggests that (3) is an example of a sequence where B gracefully avoids giving up his name, whilst not being outright confrontational (e.g. 'I don't want to tell you my name', silence, etc.). A repeats their introduction, yet B's previously expected complementary contribution can now be less problematically skipped through the following confirmation sequence.

The field of conversation analysis is incredibly rich, and this brief paragraph cannot do Sacks' analysis justice. For the purposes of this introduction, what matters is that Sacks, through this shift of interpretation, laid the foundation for the analysis of interaction of hitherto unheard-of granularity.

Compared with earlier microsociological theories, CA's shift from the unique to the typical, repeatable, and comparable made a number of advancements possible. First of all, scholars could now specialize on a class of interactional phenomena. For example, following Goffman's [1981] writing on response cries, Heritage [1984; 2016] spent the last three decades investigating 'oh' as a change-of-state token, i. e. a way whereby interactants publicly produce a purportedly internal change of knowledge (e.g. 'Oh, I see'). The relative stability of this analytic focus, in turn, made it possible to attempt a longitudinal investigation of changes in language use. Couper-Kuhlen [2019], for instance, recently started to investigate how 'oh' has changed in its use over time.

Secondly, repeatable interactions invite the possibility of formalization and application. If we, for instance, know how people manage to hang up politely, why not use that insight to provide better customer service? If a particular turn of phrase contributes

¹ Transcript numbering preserved from the original.

ammunition for adversarial conversational developments, why not discourage its use in institutional settings? If name-based introductions invite ‘I can’t hear you’-type exchanges, perhaps these introductions can be dropped for certain types of calls? Thus, the comparative rigor and systematicity of contemporary CA has contributed to a number of voyages into multidisciplinary fields of application.

The following section will introduce the first such intervention into the field of HCI, starting with an account of the assumptions common to the field prior to this intervention.

The Ethnographic Turn in HCI: The Two Waves of Conversational Intervention

It would not be correct to call HCI a traditional disciplinary field since it focused on a specific object from its very inception. However, engineering and computer science were traditionally allied disciplines, to which psychology soon joined. The latter managed to cement itself as a key player in most in-field discussions. While there are many reasons for this, one key reason is arguably the introduction of the *Model Human Processor* as a ready-made metaphor for the study of human-interface interactions [Card, Moran, Newell, 1983].

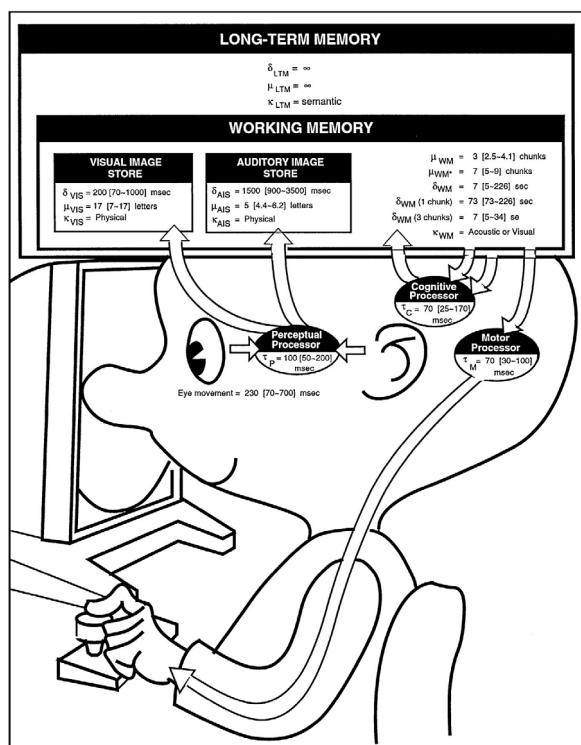


Fig. 1. The Model Human Processor²

² Source [Card et al., 1983: 26].

In the Model Human Processor, one could conceive of human-computer interaction as an integrated computational system, where the human's capacity to (re-)act, process and memorize could be quantitatively assessed in terms of its enmeshedness with a specific interface (see fig. 1). While this general approach may appear somewhat simplistic, its potential applications proved to be quite productive for the field. This is primarily because it, itself, functioned as an interface between the myriad findings of contemporary cognitive psychology and matters of interface design. This intersection could be readily investigated in psychological research laboratories all over the world.

The Ethnographic Turn

Over time, however, a number of researchers of a more sociological persuasion made their forays into this field. Contrary to the psychological take on HCI, sociologists, most notably L. Suchman, argued that there is a key difference between a model of human-interface interaction and the way it unfolded in actuality [Suchman, 1987]. The issue was not so much that the model of human interaction was flawed — and should therefore be refined or replaced with a better model — but that the very idea of modeling human action as a processing sequence did not account for the inherently procedural nature of human action. Adopting insights from ethnomethodology [Garfinkel, 1967], thinkers like Suchman stressed how humans interact through a process of procedural reinterpretation and reassessment of what is going on. These processes of actual in-situ interactions could not be expressed in sequential models; they had to be studied as they unfolded.

This shift in focus contributed to what may be called the ethnographic turn in HCI [for review see Carroll, 2010; Klowait, 2018a]. While laboratory-based studies of human-interface interactions still had their place, they were now more readily expanded by observation-based investigations of the actually unfolding interaction, and the radically contingent systematicities to be found therein.

This ethnographic turn, notably, was not an ethnomethodological turn. Although Suchman [1987] imported some conversation-analytic insights, these insights came at the price of simplification. The distinction between *plans* and *situated actions* allowed future researchers to argue that we *have to see how interaction **actually** unfolds*, but did not entrench ethnomethodology specifically as the methodology of choice.

Sociological Bugfixing

Moreover, Suchman's work, while certainly making a good case for ethnographic investigation of the device's actual-situated-use-practices, is not incompatible with a fairly regular iterative design cycle (see fig. 2).

A device was created based on plans, prototyped in a certain number of expected scenarios, and deployed. Upon evaluating the performance of the device in the field, changes to the design were implemented in the plan for the next cycle of deployment.

In other words, the ethnographic turn could very easily be integrated as a kind-of social bugfixing: sociologists would iron out quirks that emerged at the messy and difficult-to-predict stage of human-device interaction, by observing and cataloguing 'errors' to be fixed at later stages. Thus, the ethnographic turn partially became palatable to a much less radical paradigm of human-computer interaction. While it gave

ample job opportunities to aspiring ethnographers of technology, it remained a far cry from a wholesale adoption of conversation-analytic principles. An attempt to go beyond such first forays was attempted in the early 1990s.

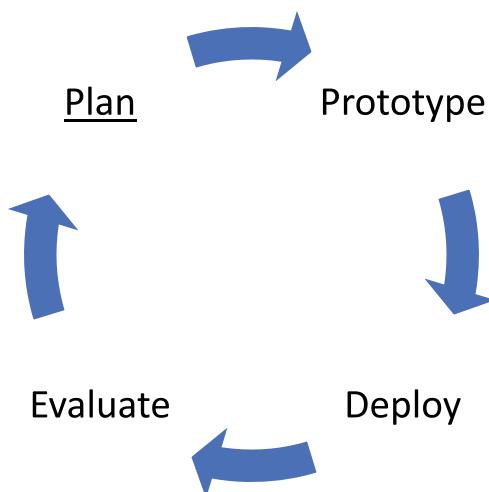


Fig. 2. Iterative design cycle³

The Second Wave of Conversation Analysis in HCI

In 1990, Luff et. al published a collection of articles in a volume entitled *Computers and Conversation* [Luff, Gilbert, Frohlich, 1990] detailing their vision for a conversation-analytic school of human-computer interaction [Frohlich, Luff, 1990: 187—188, emphasis ours]:

*Sacks suggests a programme of analysis of the details of individual sequences of talk, directed towards the discovery of what he calls a technology of conversation; a kind of common machinery and know-how for the manufacture of conversational contributions within a particular culture. **The direction of the process here is from phenomena to technology.***

*Our own programme has been one in which we have attempted to apply Sack's programme of analysis in reverse. That is, we have taken various aspects of the technology of conversation as it has been described in the CA literature and tried to build them into the operation of a computer in such a way as to generate and support orderly sequences of talk. **The direction of the process here is from technology to phenomena.***

In other words: conversation analysis reveals the rules of interaction. These rules can now be integrated into our non-human systems.

³ For a formal treatment of iterative design, see [Boehm, 1988].

If successful, this attempt would establish a firm stronghold of conversation analysts in HCI. After all, the field of CA was still relatively young and would go on to consistently generate novel findings about the structures of social action on the microscale: Sacks' legacy made it possible to see order in the most mundane snippets of recorded interaction, which in turn could generate insights about how to set up more effective computer interfaces. CA was, at least on its promotional leaflets, a powerful new microscope that could reveal a hitherto unavailable actual social reality. It could make a case for its systematicity and compatibility with the kind of systematicities handled by traditional HCI researchers.

Button's Rebuttal

Sacks' [1984] own writing seemed to generate support for the kind of project envisioned by most of the authors of *Computers and Conversation* [Luff et al., 1990]. He argues [Sacks, 1984: 413, emphasis ours]:

*The gross aim of the work I am doing is to see how finely the details of actual, naturally occurring conversation can be subjected to analysis that will yield the technology of conversation. The idea is to take singular sequences of conversation and tear them apart in such a way as to find rules, techniques, procedures, methods, maxims... **that can be used to generate** the orderly features we find in the conversations we examine.*

Not only is he talking about a technology of conversation, but he explicitly talks about how these technologies can actively generate observed social order. As such, one might get the impression that CA would generally welcome this newfound field where its findings could find practical application. As it turned out, this was far from the case.

At the forefront of the resistance against this kind of understanding of CA was, ironically enough, a chapter of the very volume that proposed this epistemic intervention in the first place [Button, 1990: 84, emphasis ours].

Conversation analysis displays: that the rules involved in the organisation of turn-taking for conversation are not part of a mental machinery of rules that stand outside of actual occasions of activity; that they are not algorithms; that they are not sets of instructions; that they are not programs of human thought; that they are not the cause of human action. Rules are oriented to features of action; they are contextual, situated practices of use.

Button [ibidem] reminded his colleagues that Sacks explicitly moves away from a view that attributes causal efficacy to rules. Much in line with the rich psychological tradition, a causal view of rules would 'mine' CA research for 'social rules' that people 'follow'. Since these rules function as a kind-of set of instructions for people, they may as well have the same function for interactive devices. Against this causal conception of rules, Button highlights the ethnomethodological tradition of viewing rules as a local achievement. In that view, rules are not something that is part of an internal set of instructions of how to interact, but function as locally, mutually, publicly orientable, noticeable, follow-able, break-able and account-able objects in an actual unfolding interaction.

Sacks' famous analysis of the first few seconds of talk during suicide prevention calls [Sacks, Jefferson, 1992: 3—11] highlights how ordinary speakers do not reproduce subconscious scripts but rely on a playful, actionable and thoughtful orientation to conversational expectations that can be used to further the distributions of conversational agendas at any given time.

In other words, there is not simply a rule like 'To ordinary speakers of the English language, when a speaker includes their name in their spoken introduction, the other speaker should include their name in the response'. If that were so, we could theoretically formalize sequences of social interaction into definite ritual-like processions steered by a long-yet-finite list of conversational rules. Instead, the rule itself is something that is available to the interactants. Not only that, its mutual availability is, in turn, also mutually available. This moves the relevance of rules from a causal—and therefore easily computerizable and instructable—to an instrumental dimension. Rules, therefore, feature as local hinges for the accomplishment of contingent actions.

This rule-shift is important, as it re-classifies the focus of sociology from the question of social order as a function of normative integration to the question of social order as a situated, second-by-second achievement. Computers are therefore principally unable to digest the kinds of rules introduced by CA, at least so long as their engineers implemented these rules at the level of pre-situational instructions. Thus, a conversation-analytic HCI (CA-HCI) is deeply problematic, as it glosses over the fact that CA's programme represents a shift in the analysis and granularity of action. Until computers become sufficiently advanced to be able to rule-orient rather than rule-follow, the machinery of conversation would, as a computer script, only result in a simulacrum of conversation. It can be argued that more advanced AI may actually be capable of moving beyond the rule-following paradigm described as being fundamental to Button-era AI [see Fordham, Gilbert, 1995; Button, Lee, Coulter, Sharrock, 1995; Wooffitt, Fraser, Gilbert, McGlashan, 1997]. The subsequent sections will discuss the technological advancements that have cast doubt on overly simplistic assessments of the capabilities of AI as a basis for understanding human-computer interaction.

Theoretical Shifts in HCI

The Sociological Status Quo: The Media Equation

While CA theorists engaged in a prolonged discussion of the fundamentals of action theory, and its significance for the application of CA to HCI, the dominant sociological contribution to HCI, apart from the purely methodological niche occupied by researchers like Suchman, remained the media equation paradigm. Its persistent foothold in HCI—despite the untimely recent death of one of its founders, C. Nass—can be partially attributed to the ease with which the media equation paradigm blended insights about human interaction and the dominant computational logic of HCI.

Nass, along with Reeves, in their Computers are Social Actors (CASA), media equation, or social response theory [Reeves, Nass, 1996; Nass, Steuer, Tauber, 1994] posited that humans, when using technology, regularly (and unconsciously) attribute humanlike characteristics to non-human material objects. More specifically, the presence of a number of subconscious cues (such as language use, voices, facial features) in a non-human artifact will result in an automatic application of social rules to that

object. It means that interfaces imbued with minimal anthropomorphic cues activate social behavioral patterns in users; they unconsciously start to apply stereotypes, display politeness and treat the interface's conduct as though it was a human [Nass, Moon, Green, 1997; Nass, Moon, 2000].

Thus, for example, an embodied conversational avatar with a female face may be met with suspicion when placed in the context of a hardware store. In other cases, users may give more favorable reviews to a product in its presence (provided the abovementioned caveat of minimal cues), adhering to the principle of politeness.

Nass' approach, in short, has a peculiar understanding of what 'social rules' are, and how they are applied: social rules are treated as autonomous, involuntary, oftentimes non-reflexive, responses. These rules may be part of some genetic imprint or they may have been 'installed' by cultural forces over time; in either case, they stand in a causal relation to human action.

In sum, Nass' minimalist theory created a powerfully simple algorithm for designing interactive agents: on the condition of a very small number of anthropomorphic features being present, the machine's behavior could be directly informed by social-psychological insights on human behavior. Best practices for interface design could thus be directly mined from human-focused research, without much consideration for the specific configuration of the agent. The question would become 'what kind of human would you like in place of the avatar?', and the relevant characteristics could then be assembled by adding specific behavioral patterns. This paradigm is still prominently represented across a broad range of research on interface design and robotics [see for example Aeschlimann, Bleiker, Wechner, Gampe, 2020; Cameron et al., 2021; Lawson et al., 2021]. Yet, current theoretical trends in HCI are gradually changing the sociological status quo of the field.

We are currently observing a convergence of two theoretical developments. Firstly, the debates surrounding the second conversation-analytic intervention into HCI are turning to pragmatic — rather than paradigmatic — concerns. At the same time, the field of HCI in general is currently trending towards a need for a greater understanding of the situation of device use. More specifically, HCI is becoming increasingly interested in understanding the interactional resources available to people during the actually-unfolding interaction. The following sections will deal with these shifts in greater detail.

From Paradigmatic to Pragmatic — On Useful Simulacra

Over time, CA-inspired contributions to HCI have shifted away from the radical position of Button [1990]. They have done so by shifting the question from the domain of rule-following vs. rule-orientation to questions of practical accomplishments and sensemaking [Jones, Mitchell, 1994: 528]:

While true 'conversation' between human and computer is arguably not possible in principle [e.g. Button, 1990], nevertheless, as a metaphor for human-computer interaction, guidelines for effective human interaction are applicable to the design of intelligent support systems.

In other words, while a 'conversation' is not possible, *something* is possible. After all, a *something* between human and machine still displays moment-by-moment sen-

semaking on the part of the human, and still has an outcome. Wooffitt et al. [1997: 166] make the point more forcefully:

The basis of Button's distinction between human-human interaction and human-computer interaction is that, in the former case, rules are embodied in interaction. People orient to rules, whereas computers are determined by them. Thus, computers cannot converse because they cannot register or display any sensitivity to procedures for producing intelligible interaction. Therefore, it makes no sense to talk of interaction between humans and computers. There are good reasons, however, for assuming that this position may be simply incorrect, because in the case of human-computer exchanges, there is always one party that does possess the range of sense-making procedures which, according to Button, demarcate human-human interaction from human-computer interaction: the human participant will still be doing the things that humans do when they interact. That is, the full range of culturally available sense-making procedures will be brought to bear on any occasion, even if the other party to the interaction is a computer.

In short, the question about the nature of conversational interaction was adjusted to the question of the practical accomplishment of a human-machine encounter. This move is not entirely fair to Button [1990], since his arguments concerned not the impossibility of humans interacting with, say, Wilson the volleyball, but the methodological pitfalls that may come with an overly enthusiastic equation of that type of encounter with a human-human interaction: we cannot make the volleyball do sensemaking by stuffing a CA textbook inside of it.

In practice, however, this move allowed CA-HCI researchers to digest Button's [1990] challenge by investigating (and contributing to the production of) *useful simulacra*.

Moore and Arar [2019], for example, while conceding that ‘the “rules” and models of natural conversation provided by CA are not the same kind of rules as those found in a programmed system’ and are ‘not deterministic, but rather are representations of resources that human speakers use in repeated by nondeterministic ways’, argue that this ‘detailed picture of how human conversation works a speech-exchange system’, allowing UX⁴ designers to ‘create interaction patterns that emulate features of human talk, although certainly with limitations and approximations’ [ibid.: 88]. Moore and Arar call these systems ‘conversation games’, representing a ‘distinctive form of interaction, which borrows interaction patterns from natural human conversation but also exhibits its own mechanics’ [ibid.: 5].

In sum, by accepting the metaphorical character of ‘human machine conversation’, contemporary CA researchers sidestep the impossibility of investigating ‘human machine conversation, but without quotes’ by incorporating Button’s methodological cautions without being defeatist about the entire project⁵. Shifting the research to *outcomes* over essences makes it possible to get on with practice-oriented CA-inspired research.

⁴ User Experience (UX) design is a design paradigm that focuses on improving the user's experience with a product throughout its entire life cycle.

⁵ Another theoretical shift in conversation-analytic approaches to technologies occupies the high-ontology segment of the conceptualization of artefacts: objects as sets of affordances or autonomous interactants. For a discussion of the conceptual and methodological implications of this shift, see [Erofeeva, 2019; Klowait, 2019].

As a consequence, HCI—CA cemented its ability to produce meaningful insights to the *in-situ* interaction of humans and computers.

The Situational Turn in HCI

This pragmatic direction of HCI—CA coincided with the field of HCI on the whole. The potential for increased computational performance and representational fidelity contributed to a move towards designing more complex interactional systems. This, in turn, has contributed to a renewed interest in understanding the interactional resources used by humans. This section will discuss the paradigmatic shifts that have occurred within the more practice-and design-oriented spaces of HCI, and link it to the conceptual issues discussed previously. As it turns out, the largely CA-specific discussions that took place more than two decades ago were prescient regarding technological developments that occurred in more recent times. In other words, we may say that the technological capabilities have caught up with more theoretical concerns, and made the latter pragmatically relevant to the field as a whole.

Recently, Realism Maximization Theory (RMT) has emerged, arguing that ‘minimal resemblance to human beings is not enough to improve the interaction, but that care must be taken to maximize realism, defined as the virtual character’s degree of visual and/or behavioural resemblance to a human being’ [Chérif, Lemoine, 2019: 30; Kang, Watt, 2013].

The consequences of RMT, with its inclusion of considerations of behavioural resemblance (rather than the presence of anthropomorphic cues) has contributed to an increased attention to *multimodality*, i.e. the consideration of how human behavior is accomplished with a body-in-space, and how ‘different verbal and non-verbal cues such as presence/absence of audio, pitch, prosody, backchannel (BC), turn-taking, body posture/gesture (upper-torso, arms, hands, legs, etc.), facial expression, gaze, etc.’ [Norouzi et al., 2018: 19] would need to be considered in interface design.

More abstractly, the shift to behavioral realism marked a shift in emphasis: interactional modalities such as gaze-behavior or gesticulation needed to be investigated in terms of their contribution to human social interaction, which, in turn, would inform human-avatar interaction. For example, the fact that *Clippy* has eyeballs becomes more than just a cue for social scripts; instead, RMT calls for a consideration of what the eyes actually do during interaction: what do they do when the avatar is listening, speaking, thinking, observing movement, and how is that behavior related to prosody, body posture, gesticulation and interactional context?

It is perhaps no coincidence that RMT emerged at a time where the technological capabilities of artificial intelligence increased dramatically. Contemporary conversational agents⁶ can be imbued with complex language parsers, sentiment analysis systems, emotion detection, and image and video classifiers; at the same time, they are able to do so whilst rendering a responsive and fully embodied avatar. The central concern of the field—and how to design well-functioning embodied conversational agents (ECAs)—consequently becomes a matter of **congruence**.

⁶ In keeping with the relevant research tradition, in what follows these systems will be called ECAs—Embodied Conversational Agents.

The Rise of Congruence

Contemporary evidence-based ECA design can be characterized as an attempt to find the correct balance between the avatar's self-presentation and its actual capabilities (see fig. 3).

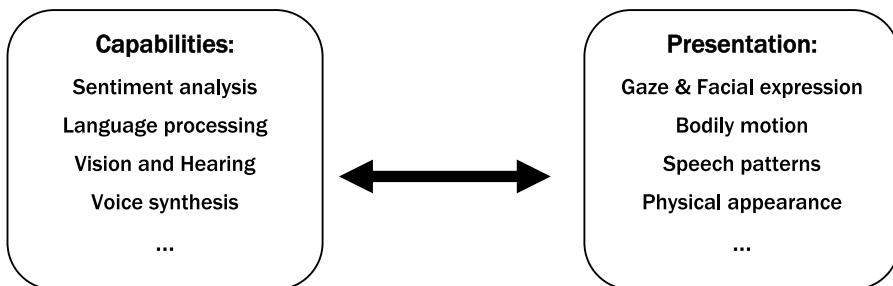


Fig. 3. The tension between capabilities and presentation

Since 'people are more comfortable with a robot clearly identified as non-human than with a human-looking robot that still has imperfections' [Chérif, Lemoine, 2019: 32], the design directive can be bi-directional, where the specific demands for the ECA can dictate the correct trade-off between realism and capability. If, for instance, only minimal behavioural humanlike-ness is required, the avatar may be **scaled back** in realism to adhere to the congruence principle. If, conversely, the avatar is hyper-real, steps must either be taken to imbue it with the correspondingly high level of capabilities, **or** make sure that the limited capabilities of the agent are adequately circumscribed (see Table 1).

Table 1. **Paradigms mapped to the interaction-realism nexus**

	Low-realism	High-realism
Low-interaction capabilities	CASA paradigm	Incongruence
High-interaction capabilities	Limited congruence	RMT paradigm

Evidence-based Insights on Avatar Creation

ECAs, being conversational, are inherently interactive systems. Based on the aforementioned congruence principle, interaction should cohere with social rules, and the modalities present within the expressive range of the ECA should cohere with one another [Krämer, 2008]. As Tan and Liew [2020] argue, 'the mismatch between the seemingly sophisticated embodiment of virtual agents and the agents' lack of functional performance can cause the expectation gap in users, leading to disappointment and frustration with the embodied virtual agents' [ibid.: 1]. Samsung's STAR Labs recently debuted Neon at the International Consumer Electronics Show 2020. The way it was presented is a convenient means of illustrating the issue in practical terms (see fig. 4):

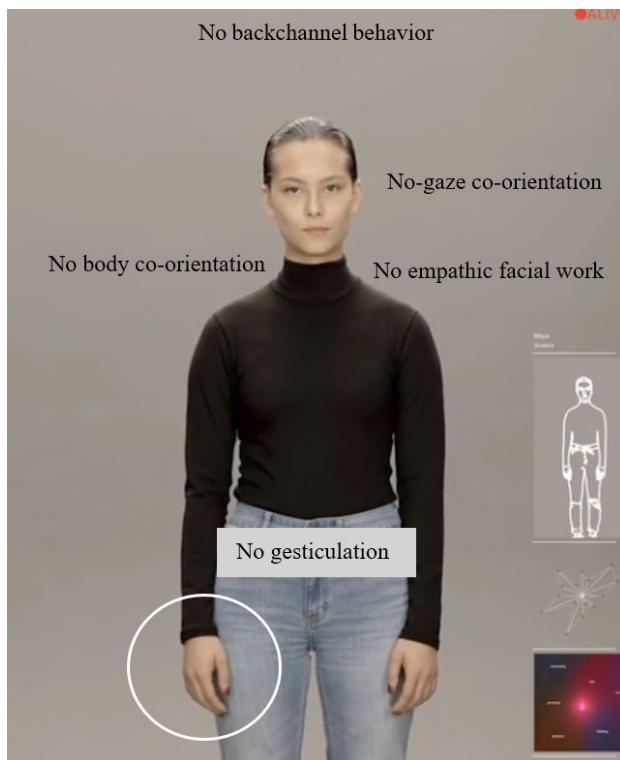


Fig. 4. Neon Avatar⁷

The above graphic is adapted to illustrate the CES demo of the current prototype⁸. It is a stark contrast to the humanlike idle animations that were demonstrated *prior* to the interactive demo: no alignment was demonstrated between human and ECA; no alignment was demonstrated between utterances and body/gestural work, either. While the system is still in development, some of the negative reactions can be explained through this incongruence.

In contrast to this, Soul Machines, arguably one of the industry leaders of realistic ECAs, demonstrates sophisticated backchannel behavior in even its less recent products, such as the Australia and New Zealand Banking Group's digital assistant, Jamie⁹. To reiterate: the focus is not realism at all costs, but the *evidence-based* application of humanlike interactional resources in appropriate contexts.

Verbal Behavior and Beyond

Verbal behavior, unsurprisingly, plays a prominent role in the design of conversational systems. While ECAs can remain 'mute' and talk through a purely text-based system,

⁷ Source: URL: www.neon.life (accessed: 03.02.2021); text added by the present authors.

⁸ For more details see: URL: <https://youtu.be/4IPNOWiDOK> (accessed: 03.02.2021).

⁹ See: URL: <https://youtu.be/eyoBgNY1KA0> (accessed: 03.02.2021).

they are typically imbued with both a voice and the corresponding capability to move the lips. This is firstly due to the observation that embodied verbal behaviour — i. e. the match between visible embodied conduct and the audiostream — improves engagement, realism, trust [Bos, Olson, Gergle, Olson, Wright, 2002; Greenspan, Goldberg, Weimer, Basso, 2000], and, curiously, aids comprehension: in their seminal paper, Walker Sproull and Subramani [1994] ‘investigated subjects’ responses to a synthesized talking face displayed on a computer screen [...]. Compared to subjects who answered questions presented via text display on a screen, subjects who answered the same questions spoken by a talking face spent more time, made fewer mistakes, and wrote more comments’ [ibid.: 85].

Voice characteristics

Voice — far from being just a means of transmitting information between speakers — is a vehicle for personality, status, identity and relationship. As such, the construction of the ECA’s voice is of key importance to the subsequent interaction with the user.

The lowest level of consideration is the question of what kind of voice technology to use. Currently, the choice is made between synthetic text-to-speech (TTS) generators and pre-recorded human voice that is played at the appropriate times. For practical purposes, this means a trade-off between conversational flexibility and expressiveness. Synthesized speech, be it concatenative (i. e. assembled out of fragments of recorded human speech) or parametric (i. e. newly generated through the creation of a linguistic phonetic speech model) has the benefit that it can produce context-aware responses from scratch. The downside, however, is that the speech itself is typically lifeless and robotic (especially in the case of parametric speech synthesis), with locally incoherent prosodic expressiveness. The alternative, using ‘canned’ pre-recorded (and therefore maximally human-sounding) responses greatly limits interactional flexibility, yet contributes to higher levels of trust when voice source is the only varied factor.

More recently, neural network based systems like Google’s WaveNet proposed a ‘parametric synthesizer on steroids’ [van den Oord et al., 2016], i. e. a means of leveraging machine learning to create a neo-parametric model that is *both* lifelike and locally flexible in terms of what it can say. However, it and similar neural TTS solutions are arguably more difficult to implement and represent a potentially high overhead. Moreover, since systems like WaveNet learn to synthesize speech in a holistic manner — with the linguistic model becoming an implicit and entangled component of the learning process — ‘prosody realization is randomly chosen and cannot be easily altered’ [Shechtman, Sorin, 2019: 275]. In other words, the system becomes more lifelike in *sound*, but loses the ability to alter intonation to the context of the spoken word¹⁰. Thus, ‘congratulations, you won’ and ‘my apologies, you failed’ would sound the same without additional intervention, and would consequently be contextually incongruent. This shortcoming can be overcome [Shechtman, Sorin, 2019] but requires additional resources compared to simply recording two messages with appropriate prosodic contours.

In sum, as is the case with the aspect of avatar presentation (presence and embodiment), good ECA design becomes a matter of *leveraging existing capabilities to ensure*

¹⁰ For an example of how a complex artificial intelligence system’s use of prosody prolongs the closing of a service encounter see Egorova, Klowait, this issue.

congruence. That is, if the conversational system is designed with a focus on pre-determined sentences, human voice recordings would be the obvious direction to head towards. Conversely, if conversational flexibility is desired — this is especially true for ECAs that try to maximize realism and generalize its range of applications — the more flexible synthetic voice is preferable, especially since conversational interactivity can moderate the de-anthropomorphizing effect of robotic voice. A negative example of the latter case is described in Klowait [2017], where an automated computer-telephone interviewing system was created with realistic pre-recorded human voices, yet integrated with a contextually inflexible conversational system. The resulting negative user experiences — a considerable number of which never reached understanding that they were talking to a limited AI rather than a rude/incapable human — could have been avoided by opting for an obviously synthetic voice.

Turn-taking behavior

Beyond the choice of voice, the things done with the voice are of key importance to realism. In that respect, the field has moved beyond the informational model of conversation, whereby speech is a kind of message exchange [Quarteroni, 2018; Schröder et al., 2012; Bernard, 2017]. Instead, the field has embraced the socio-linguistic, anthropological and sociological insight that *talk is action* [Goodwin, 2000] whereby humans typically accomplish disparate goals, moment-by-moment. A key element of this action is *turn-taking* behavior, i.e. the organization of participants' talk-action in a recognizable turn-by-turn basis, and the demarcation of interactional phenomena such as backchanneling ('listener' activity during another speaker's turn at speech), interruptions, anticipations of transitions between speakers and overlap.

Some of those behaviors are currently beyond the limits of AI, such as the human ability to exploit the implicit rules of conversational organization in order to change a topic, take over as a speaker or gracefully avoid answering a direct question. However, especially 'attentive listening behavior', where a 'combination of head nods, vocalizations and facial feedback [shows] agreement and acknowledgment' [Yalçın, 2020: 124] is possible to implement, and is argued to increase realism and rapport.

For example, Yalcin's [Yalçın, 2020; Yalçın, DiPaola, 2019] work on empathic ECAs includes an interruptible state-based model of affective listening [Yalçın, 2020: 125—126]:

According to the state of the dialogue, the behavior of the agent can change and adapt to the user. While the interaction partner is speaking, the agent enters the listening state. Listening mode will be activated via the speech and video input from the agent [...]. In this state, the agent is expected to provide proper backchanneling behavior as well as the emotional feedback. After the speech of the interaction partner is completed, the agent will enter the thinking state. In this state, the agent will be finished gathering information from the perceptual model and start processing the speech input for generating a response. This response generation process will make use of the context of the dialogue as well as the emotional content of the message. Lastly, the agent will enter the speaking state, where it executes the prepared response via its output channels including voice, facial expression and body gestures.

In other words, such a simplified system can interface with pre-existing machine states, such as information retrieval, idling, processing and behavior production. Moreover, as this framework is multimodal, i. e. incorporates synchronous and overlapping modalities (e. g. gaze, head and body movement, prosody, speech), it represents another way of transforming unrelated idling mechanisms (such as randomized avatar body shifts) into context-sensitive tools to improve presence. Lee, Badler and Badler's [2002] work on turn-taking-sensitive saccades represents evidence that this is a general trend: gaze that is interactionally purposeful enhances user experience. This observation is notably true for both random eye movement and constant direct eye contact, which has a tendency to become unnerving to the human participants.

Lastly, a system that is mutually integrated—i. e. where e.g. gaze is coordinated with body posture and speech—makes it possible to flexibly integrate higher-order multimodal behavior such as affective matching and emotional mimicry [Yalçın, DiPaola, 2019], both of which further aid the humanlike-ness of interaction.

Specialized Subsystems

Beyond the more modest strategies of implementing at least some listening behavior, research has increasingly shifted focus to implementations of higher-order interactional features, such as social status-specific multimodal conduct [Nixon, DiPaola, Bernardet, 2018] and emotion monitoring, classifying and mirroring [Yalçın, 2020]. The general trend can be represented by a hierarchy (see fig. 5):

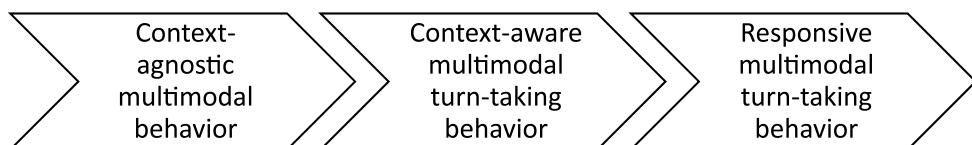


Fig. 5. Transition towards high-order interactional specificity

In other words, incongruent and random behavior (or the absence thereof) is rated worst, followed by behavior that displays a distinct (albeit teleologically fuzzy) orientation to ongoing interaction, followed by systems that attempt to respond to this ongoing interaction in a meaningful way (e. g. through empathy displays at the appropriate moments).

Towards a Multimodal Framework in HCI Research

As can be seen, current avatar design best practices are moving towards realism maximization. One consequence of this movement is the shifting focus towards the multimodality of communication when developing avatars. Despite this trend, ECA evaluations tend to focus on non-situational, non-ethnographic research methods [Kramer, ter Stal, Mulder, de Vet, van Velsen, 2020]. This methodological preference does not make it possible to investigate the actual use of multimodal resources for the design of ECAs, since experimental research does not rely on the study of what functions different modalities (gesture, gaze, etc.) perform in human communication.

For instance, a number of studies in the field of multimodal human-computer interaction show that people not only perceive ECA behavior in a certain way, but also adapt

to it in real-time [Norouzi et al., 2018; Zhang et al., 2010; Bohus, Horvitz, 2010]. If the agent is set up to track an object at the center of the user's attention 90 % of the time, the user will pay less attention to the virtual agent. If this condition is changed to 50 % or 10 %, the user will begin to focus on the ECA [Zhang et al., 2010]. In other words, the level of involvement inscribed in the system affects the level of user involvement in said interaction: if an ECA has a single point of gaze-focus, users will tend to expend little resources for tracking the ECA's gaze; if, conversely, the ECA visibly distributes their attention to other objects in the world, users attend to that fact by tracking the ECA's gaze more attentively. However, this result in itself does not say anything about the need for this level of user involvement. In order to evaluate it, we need to understand what function attention management serves in interaction.

When it comes to complex multimodal interaction scenarios for state-of-the-art ECAs, it appears especially important to incorporate interaction-focused research methodologies at all stages of the design process. More specifically, there is a lack of a true multimodal framework for the development and final evaluation of the interactional outcomes of ECAs.

Given the multimodal complexity of the multiparty coordination of speech, gaze and other interactive resources, it would seem reasonable to assume that the research and evaluation included insights from state-of-the-art multimodality research. Yet, only classical gaze studies [Kendon, 1967] seem to find widespread purchase in the field.

As was elaborated in a previous section of this review, CA's current response to Button's critique (useful simulacra) goes hand in hand with technical developments in HCI on the whole. As such, the field is now at a stage where CA may make a third serious attempt at a methodological intervention¹¹. This time, however, with the aid of multimodal conversation analysis, specifically.

Multimodal conversation analysis makes it possible to analyze the practices and resources used by people to interact with one another. Such resources are not limited to the domain of language but also include their embodied actions—such as gestures, gaze directions, and body movement [Goodwin, 2000; Mondada, 2019]. In contrast to the multimodal approach in psychology, communicative resources are thought to be applied situationally, i. e. within a specific context and for a specific task.

Research of human interaction with ECAs, carried out in line with multimodal conversation analysis, follows the logic proposed by Suchman [1987]: on the one hand, there is the object's design into which a particular mode of interaction is inscribed; on the other, there are the ways this object is interacted with *insitu*. In the process of analyzing video recordings of real-life interaction, discrepancies are found between what the system 'sees' (its interactive abilities) and what the user expects [Ewa, Abigail, 2016]. In particular, studies of interaction with conversational agents reveal problems with the turn-taking system (transfer of speakership from an ECA to a person and vice versa): people either start talking too early or 'skip' the allotted space for an answer, as a result of which their utterance is ignored, forcing participants repeat the sequence of interaction over and over [Arend, Sunnen, 2017; Pelikan, Broth, 2016; Pitsch, Gehle, Dankert, Wrede, 2017]. Furthermore, conversational agents may overlook the constant

¹¹ See Saul Albert's recent keynote: Albert S. (2020, June 29) Three Meeting Points between CA and AI. URL: <https://saaulbert.net/blog/three-meeting-points-between-ca-and-ai/> (accessed: 03.02.2021).

repair sequences performed by human speakers, with similar consequences [Klowait, 2017; Trott, Rossano, 2017].

The main practical conclusion of such studies is the idea that—for a smooth interaction with an ECA—the user needs to somehow be ‘explained’ how the virtual agent works, and the communicative abilities the system possesses. That said, the aforementioned learning processes occur within the interaction itself. From a theoretical point of view, this means that the social expectations of users dynamically change based on their reaction to the sequentially unfolding actions of the virtual agent.

Multimodality, understood in the conversation-analytic tradition, challenges both the media equation paradigm and the idea of realism maximization. Unlike the media equation, people interact with an ECA on the basis of a situational adaptation to the interface, and not on the basis of the unconscious illusion that we are interacting with a person rather than a computer (these theoretical assumptions were conceptualized by Klowait as a distinction between pragmatic and ontological anthropomorphism [see Klowait, 2018a]). As for the realism maximization theory, it merely manages to touch the tip of the iceberg of real engagement with an ECA.

Conclusion

This review traced a convergence of developments. On the one hand, we demonstrated the path of CA-HCI from an initial low-stakes intervention [Suchman, 1987], to a high-stake overenthusiasm [Frohlich, Luff, 1990] and, finally, to a marked shift towards pragmatic in-field concern [Moore, Arar, 2019]. On the other, we demonstrated how contemporary studies of ECAs have converged on the need to understand the multimodality of human interaction to facilitate the development of novel forms of human-computer encounters. We showed how there is currently a growing need for a congruence between an ECAs representational fidelity and interactional competence. We believe that multimodal conversation analysis can provide much needed answers, both in terms of insights and in terms of the appropriate ways to generate them.

References (Список литературы)

- Aeschlimann S., Bleiker M., Wechner M., Gampe A. (2020) Communicative and Social Consequences of Interactions with Voice Assistants. *Computers in Human Behavior*. Vol. 112. 106466. <https://doi.org/10.1016/j.chb.2020.106466>.
- Arend B., Sunnen P. (2017) Coping with Turn-Taking: Investigating Breakdowns in Human-Robot Interaction from a Conversation Analysis (CA) Perspective. In: N. Callaos, B. Sanchez, M. Savoie, F. Welsch, J. V. Carrasquero (eds.) *Proceedings: The 8th International Conference on Society and Information Technologies (ICSIT 2017)*. Orlando (FL): International Institute of Informatics and Systemics (IIIS). P. 149—154. URL: <http://hdl.handle.net/10993/30952> (accessed: 05.02.2021).
- Bernard D. (2017) Cognitive Interaction: Towards “Cognitivitiy” Requirements for the Design of Virtual Assistants (Working paper). In: *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. Banff (CA): IEEE. P. 210—215. <https://doi.org/10.1109/SMC.2017.8122604>.

- Boehm B. W. (1988) A Spiral Model of Software Development and Enhancement. *Computer*. Vol. 21. No. 5. P. 61—72. <https://doi.org/10.1109/2.59>.
- Bohus D., Horvitz E. (2010) Facilitating Multiparty Dialog with Gaze, Gesture, and Speech. In: *Proceedings of the 12th International Conference on Multimodal Interfaces and the Workshop on Machine Learning for Multimodal Interaction (ICMI—MLMI'2010)*. New York (NY): Association for Computing Machinery. Article No. 5. P. 1—8. <https://doi.org/10.1145/1891903.1891910>.
- Bos N., Olson J., Gergle D., Olson G., Wright Z. (2002) Effects of Four Computer-Mediated Communications Channels on Trust Development. In D. Wixon (ed.) *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'02)*. New York, NY: ACM Press. P. 135—140. <https://doi.org/10.1145/503376.503401>.
- Button G. (1990) Going Up a Blind Alley. Conflating Conversation Analysis and Computational Modelling. In: P. Luff, G. N. Gilbert, D. Frohlich (eds.) *Computers and conversation*. London: Academic Press. P. 67—90.
- Button G., Lee J. R., Coulter J., Sharrock W. (1995) Computers, Minds and Conduct. Cambridge: Polity Press.
- Cameron D., de Saille S., Collins E., Aitken J., Cheung H., Chua A., Loh E. J., Law J. (2021) The Effect of Social-Cognitive Recovery Strategies on Likability, Capability and Trust in Social Robots. *Computers in Human Behavior*. Vol. 114. P. 106561. <https://doi.org/10.1016/j.chb.2020.106561>.
- Card S. K., Moran T. P., Newell A. (1983) *The Psychology of Human-Computer Interaction*. Hillsdale, NJ: Erlbaum.
- Carroll J. M. (2010) Conceptualizing a Possible Discipline of Human—Computer Interaction. *Interacting with Computers*. Vol. 22. No. 1. P. 3—12. <https://doi.org/10.1016/j.intcom.2009.11.008>.
- Chérif E., Lemoine J.-F. (2019) Anthropomorphic Virtual Assistants and the Reactions of Internet Users: An Experiment on the Assistant's Voice. *Recherche et Applications en Marketing (English Edition)*. Vol. 34. No. 1. P. 28—47. <https://doi.org/10.1177/2051570719829432>.
- Couper-Kuhlen E. (2019) American English OKAY over Time: Challenge and Chance for Interactional Linguistics. Lecture given at the *Ninth Meeting of the Language and Social Interaction Working Group (LANSI)*. New York, NY: Teachers College, Columbia University.
- Erofeeva M. (2019) On Multiple Agencies: When do Things Matter? *Information, Communication & Society*. Vol. 22. No. 5. P. 590—604. <https://doi.org/10.1080/1369118X.2019.1566486>.
- Ewa L., Abigail S. (2016) “Like Having a Really Bad PA”: The Gulf between User Expectation and Experience of Conversational Agents. In: *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI'16)*. San Jose, CA: Association for Computing Machinery. P. 5286—5297. <https://doi.org/10.1145/2858036.2858288>.

Fordham A., Gilbert N. (1995) On the Nature of Rules and Conversation. *AI & SOCIETY*. No. 9. P. 356—372. <https://doi.org/10.1007/BF01210587>.

Frohlich D., Luff P. (1990) Applying the Technology of Conversation to the Technology for Conversation. In: P. Luff, G. N. Gilbert, D. Frohlich (eds.) *Computers and conversation*. London: Academic Press. P. 187—220.

Garfinkel H. (1967) Studies in Ethnomethodology. Englewood Cliffs, NJ: Prentice Hall.

Goffman E. (1981) Forms of Talk. Philadelphia, PA: University of Pennsylvania Press.

Goodwin C. (2000) Action and Embodiment within Situated Human Interaction. *Journal of Pragmatics*. Vol. 32. No. 10. P. 1489—1522. [https://doi.org/10.1016/S0378-2166\(99\)00096-X](https://doi.org/10.1016/S0378-2166(99)00096-X).

Greenspan S., Goldberg D., Weimer D., Basso A. (2000) Interpersonal Trust and Common Ground in Electronically Mediated Communication. In: W. Kellogg, S. Whittaker (eds.) *Proceedings of the 2000 ACM Conference on Computer Supported Cooperative Work (CSCW'00)*. New York (NY): ACM Press. P. 251—260. <https://doi.org/10.1145/358916.358996>.

Heritage J. (1984) A Change-of-State Token and Aspects of its Sequential Placement. In: J. M. Atkinson, J. Heritage (eds.) *Structures of social action: Studies in conversation analysis*. Cambridge: Cambridge University Press. P. 299—345.

Heritage J. (2016) On the Diversity of ‘Changes of State’ and their Indices. *Journal of Pragmatics*. Vol. 104. P. 207—210. <https://doi.org/10.1016/j.pragma.2016.09.007>.

Jones P.M., Mitchell C. M. (1994) Model-Based Communicative Acts: Human-Computer Collaboration in Supervisory Control. *International Journal of Human-Computer Studies*. Vol. 41. No. 4. P. 527—551. <https://doi.org/10.1006/ijhc.1994.1072>.

Kang S.-H., Watt J. H. (2013) The Impact of Avatar Realism and Anonymity on Effective Communication via Mobile Devices. *Computers in Human Behavior*. Vol. 29. No. 3. P. 1169—1181. <https://doi.org/10.1016/j.chb.2012.10.010>.

Pitsch K., Gehle R., Dankert T., Wrede S. (2017) Interactional Dynamics in User Groups: Answering a Robot’s Question in Adult-Child Constellations. In *Proceedings of the 5th International Conference on Human Agent Interaction (HAI’17)*. Bielefeld (Germany): Association for Computing Machinery. P. 393—397. <https://doi.org/10.1145/3125739.3132604>.

Kendon A. (1967) Some Functions of Gaze-Direction in Social Interaction. *Acta Psychologica*. Vol. 26. P. 22—63. [https://doi.org/10.1016/0001-6918\(67\)90005-4](https://doi.org/10.1016/0001-6918(67)90005-4).

Klowait N. (2017) A Conceptual Framework for Researching Emergent Social Orderings in Encounters with Automated Computer-Telephone Interviewing Agents. *International Journal of Communication and Linguistic Studies*. Vol. 15. No. 1. P. 19—37.

Klowait N. (2018a) Reflexive Anthropomorphism: Ontological Ignorance, or Ignorant Ontology? *Sotsiologicheskii Zhurnal*. Vol. 24. No. 1. P. 8—33. [\(In Russ.\)](https://doi.org/10.19181/socjour.2018.24.1.5711)

Кловайт Н. Рефлексивный антропоморфизм: неведение онтологии или невежественная онтология? *Социологический журнал*. 2018. Т. 24. № 1. С. 8—33. <https://doi.org/10.19181/socjour.2018.24.1.5711>.

Klowait N. (2018b) The Quest for Appropriate Models of Human-Likeness: Anthropomorphism in Media Equation Research. *AI & SOCIETY*. Vol. 33. No. 4. P. 527—536. <https://doi.org/10.1007/s00146-017-0746-z>.

Klowait N. (2019). Interactionism in the Age of Ubiquitous Telecommunication. *Information, Communication & Society*. Vol. 22. No. 5. P. 605—621. <https://doi.org/10.1080/1369118X.2019.1566487>.

Kramer L., ter Stal S., Mulder B. C., de Vet E., van Velsen L. (2020) Developing Embodied Conversational Agents for Coaching People in a Healthy Lifestyle: Scoping Review. *Journal of Medical Internet Research*. Vol. 22. No. 2. e14058. <https://doi.org/10.2196/14058>.

Krämer N. C. (2008) Social Effects of Virtual Assistants. A Review of Empirical Results with Regard to Communication. In: H. Prendinger, M. Ishizuka, J. Lester (eds.) *Intelligent Virtual Agents. IVA 2008. Lecture Notes in Computer Science*. Vol. 5208. Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg. P. 507—508. https://doi.org/10.1007/978-3-540-85483-8_63.

Lawson A. P., Mayer R. E., Adamo-Villani N., Benes B., Lei X., Cheng J. (2021) Recognizing the Emotional State of Human and Virtual Instructors. *Computers in Human Behavior*. Vol. 114. 106554. <https://doi.org/10.1016/j.chb.2020.106554>.

Lee S. P., Badler J. B., Badler N. I. (2002) Eyes Alive. In: *Proceedings of the 29th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH'02)*. New York (NY): Association for Computer Machinery. P. 637—644. <https://doi.org/10.1145/566570.566629>.

Luff P., Gilbert G. N., Frohlich D. (eds.) (1990) Computers and Conversation. London: Academic Press.

Mondada L. (2019) Contemporary Issues in Conversation Analysis: Embodiment and Materiality, Multimodality and Multisensoriality in Social Interaction. *Journal of Pragmatics*. Vol. 145. P. 47—62. <https://doi.org/10.1016/j.pragma.2019.01.016>.

Moore R. J., Arar R. (2019) Conversational UX design: A Practitioner's Guide to the Natural Conversation Framework (ACM Books). New York, NY: ACM Books.

Nass C., Moon Y. (2000) Machines and Mindlessness: Social Responses to Computers. *Journal of Social Issues*. Vol. 56. No. 1. P. 81—103. <https://doi.org/10.1111/0022-4537.00153>.

Nass C., Moon Y., Green N. (1997) Are Machines Gender Neutral? Gender-Stereotypic Responses to Computers with Voices. *Journal of Applied Social Psychology*. Vol. 27. No. 10. P. 864—876. <https://doi.org/10.1111/j.1559-1816.1997.tb00275.x>.

Nass C., Steuer J., Tauber E. R. (1994) Computers Are Social Actors. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'94)*. New York, NY: ACM. P. 72—78. <https://doi.org/10.1145/191666.191703>.

Nixon M., DiPaola S., Bernardet U. (2018) An Eye Gaze Model for Controlling the Display of Social Status in Believable Virtual Humans. In: *2018 IEEE Conference on Computational Intelligence and Games (CIG)*. Maastricht: IEEE. P. 1—8. <https://doi.org/10.1109/CIG.2018.8490373>.

Norouzi N., Kim K., Hochreiter J., Lee M., Daher S., Bruder G., Welch G. (2018) A Systematic Survey of 15 Years of User Studies Published in the Intelligent Virtual Agents Conference. In: *Proceedings of the 18th International Conference on Intelligent Virtual Agents*. Sydney (AUS): Association for Computing Machinery. P. 17—22. <https://doi.org/10.1145/3267851.3267901>.

Pelikan H. R. M., Broth M. (2016) Why That Nao? How Humans Adapt to a Conventional Humanoid Robot in Taking Turns-at-Talk. In: *Proceedings of the 34th Annual CHI Conference on Human Factors in Computing Systems (CHI'16)*. San Jose (CA): Association for Computing Machinery. P. 4921—4932.

Quarteroni S. (2018) Natural Language Processing for Industry. *Informatik-Spektrum*. Vol. 41. No. 2. P. 105—112. <https://doi.org/10.1007/s00287-018-1094-1>.

Reeves B., Nass C. I. (1996) The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places. CSLI Publications. Stanford (CA): Cambridge University Press.

Sacks H. (1984) On Doing “Being Ordinary”. In J. M. Atkinson, J. Heritage (eds.) *Structures of Social Action: Studies in Conversation Analysis*. Cambridge (UK): Cambridge University Press. P. 413—429. <https://doi.org/10.1017/CBO9780511665868.024>.

Sacks H., Jefferson G. (1992) Lectures on Conversation. Oxford (UK): Blackwell.

Schröder M., Bevacqua E., Cowie R., Eyben F., Gunes H., Heylen D., ter Maat M., McKeown G., Pammi S., Pantic M., Pelachaud C., Schuller B., de Sevin E., Valstar M., Wöllmer M. (2012) Building Autonomous Sensitive Artificial Listeners. *IEEE Transactions on Affective Computing*. Vol. 3. No. 2. P. 165—183. <https://doi.org/10.1109/T-AFFC.2011.34>.

Shechtman S., Sorin A. (2019) Sequence to Sequence Neural Speech Synthesis with Prosody Modification Capabilities. In: *Proceedings of the 10th ISCA Speech Synthesis Workshop*. ISCA. P. 275—280. <https://doi.org/10.21437/SSW.2019-49>.

Stokoe E. (2018) Talk: The Science of Conversation. London: Robinson.

Suchman L. A. (1987) Plans and Situated Actions: The Problem of Human-Machine Communication. Cambridge: Cambridge University Press.

Tan S.-M., Liew T. W. (2020) Designing Embodied Virtual Agents as Product Specialists in a Multi-Product Category E-Commerce: The Roles of Source Credibility and Social Presence. *International Journal of Human–Computer Interaction*. Vol. 36. No. 12. P. 1136—1149. <https://doi.org/10.1080/10447318.2020.1722399>.

- Trott S., Rossano F. (2017) Theoretical Concerns for the Integration of Repair. In: *Artificial Intelligence for Human Robot Interaction*. Arlington, VA: AAAI. P. 118—122.
- van den Oord A., Dieleman S., Zen H., Simonyan K., Vinyals O., Graves A., Kalchbrenner N., Senior A., Kavukcuoglu K. (2016) WaveNet: A Generative Model for Raw Audio. URL: <https://arxiv.org/abs/1609.03499> (accessed: 06.02.2021).
- Walker J. H., Sproull L., Subramani R. (1994) Using a Human Face in an Interface. In: B. Adelson, S. Dumais, J. Olson (eds.) *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems Celebrating Interdependence (CHI'94)* New York, NY: ACM Press. P. 85—91. <https://doi.org/10.1145/191666.191708>.
- Wooffitt R., Fraser N. M., Gilbert N., McGlashan S. (1997) Humans, Computers and Wizards: Analysing Human (Simulated) Computer Interactions. London: Routledge.
- Yalçın Ö. N. (2020) Empathy Framework for Embodied Conversational Agents. *Cognitive Systems Research*. Vol. 59. P. 123—132. <https://doi.org/10.1016/j.cogsys.2019.09.016>.
- Yalçın Ö. N., DiPaola S. (2019) Evaluating Levels of Emotional Contagion with an Embodied Conversational Agent. In: A. K. Goel, C. M. Seifert, C. Freska (eds.) *Proceedings of the 41st Annual Meeting of the Cognitive Science Society (CogSci'19)*. P. 3143—3149. URL: <https://ivizlab.org/wp-content/uploads/sites/2/2019/09/0528.pdf> (accessed: 06.02.2021).
- Zhang H., Fricker D., Smith T. G., Yu C. (2010) Real-Time Adaptive Behaviors in Multimodal Human-Avatar Interactions. In: *International Conference on Multimodal Interfaces and the Workshop on Machine Learning for Multimodal Interaction (IVA'10)*. Beijing: Association for Computing Machinery. Article 4. URL: <https://dl.acm.org/paper/10.1145/1807000.1807004.pdf> (accessed: 06.02.2021).

DOI: [10.14515/monitoring.2021.1.1840](https://doi.org/10.14515/monitoring.2021.1.1840)



N. Hosogaya

DEVELOPMENT OF TELEWORK AND ITS EFFECTS ON JAPAN'S HUMAN RESOURCE MANAGEMENT: ARTIFICIAL SOCIALITY AS A FOCUS

For citation:

Hosogaya N. (2021) Development of Telework and its Effects on Japan's Human Resource Management: Artificial Sociality as a Focus. *Monitoring of Public Opinion: Economic and Social Changes*. No. 1. P. 71–90. <https://doi.org/10.14515/monitoring.2021.1.1840>.

Правильная ссылка на статью:

Хосогая Н. Развитие удаленной работы и ее влияние на управление человеческими ресурсами в Японии в фокусе искусственной социальности // Мониторинг общественно-го мнения: экономические и социальные перемены. 2021. № 1. С. 71—90. [\(In Eng.\)](https://doi.org/10.14515/monitoring.2021.1.1840)

DEVELOPMENT OF TELEWORK AND ITS EFFECTS ON JAPAN'S HUMAN RESOURCE MANAGEMENT: ARTIFICIAL SOCIALITY AS A FOCUS

*Nobuko HOSOGAYA¹—Associate Professor, Faculty of Economics
E-MAIL: n-hosoga@sophia.ac.jp
<https://orcid.org/0000-0002-6899-0093>*

¹ Sophia University, Tokyo, Japan

Abstract. This study explores social and economic activities that promote ‘telework’ and its impact on human resource management in Japan. Telework had not been common in Japan’s workplaces, but the situation suddenly changed when the COVID-19 pandemic began to spread in 2020. The Japanese government issued the declaration of emergency and encouraged the commerce, industry, and other enterprises associations to introduce telework as a ‘new way of working’. To emphasize the characteristics of telework in Japan, the article observes research outcomes published recently. We argue that there is a difference between companies’ size and location and the level of their employees’ skills: telework has already been implemented in some major companies, while there is a marginal development in online work in rural areas and in small and medium-sized enterprises. A hybrid type of face-to-face and online communications emerged as the result. The article finds out the domains where Japanese corporations adopted telework and shows some changes and effects caused by telework in work styles and human resource management. In conclusion, we systematize the data in a respective research realm and classify the elements associated with artificial sociality.

РАЗВИТИЕ УДАЛЕННОЙ РАБОТЫ И ЕЕ ВЛИЯНИЕ НА УПРАВЛЕНИЕ ЧЕЛОВЕЧЕСКИМИ РЕСУРСАМИ В ЯПОНИИ В ФОКУСЕ ИСКУССТВЕННОЙ СОЦИАЛЬНОСТИ

*ХОСОГАЯ Нобуко — доцент, факультет экономики, Софийский университет, Токио, Япония
E-MAIL: n-hosoga@sophia.ac.jp
<https://orcid.org/0000-0002-6899-0093>*

Аннотация. В статье исследуются социальные и экономические условия, способствующие становлению и развитию удаленной работы (телефоры), и ее влияние на управление человеческими ресурсами в Японии. Телефорта никогда не была обычной трудовой практикой в Японии, однако ситуация внезапно изменилась в 2020 г. в связи с распространением пандемии COVID-19. Японское правительство объявило чрезвычайное положение и призвало торговые, промышленные и прочие бизнес-ассоциации ввести удаленную работу как «новый способ трудовой деятельности». Последствия распространения телефонии оказались настолько существенными, что вызвали изменения в традиционном для Японии стиле трудовой деятельности. С целью выделения характеристик удаленной работы в Японии в статье анализируются релевантные теме результаты национальных эмпирических исследований. Мы отмечаем, что между размером и месторасположением компаний, с одной стороны, и уровнем квалификации их сотрудников, с другой, существуют различия: если в ряде крупных японских корпораций телефония уже успешно внедрена, в сельской местности и на малых и средних

предприятиях страны наблюдается лишь незначительное ее развитие. На основании этого в статье делается вывод о становлении гибридного типа личных (face-to-face) и онлайн-деловых коммуникаций. Подводя итог, мы выделяем сектора занятости, в которых удаленная работа уже внедрена, а также описываем обусловленные ею изменения в стиле трудовой деятельности и управления человеческими ресурсами. В заключение мы систематизируем данные в соответствующей исследовательской области и классифицируем элементы, ассоциируемые с искусственной социальностью.

Keywords: telework, artificial sociality, Japan, COVID-19, human resource management

Ключевые слова: удаленная работа, телеработа, искусственная социальность, Япония, COVID-19, управление человеческими ресурсами

Introduction

This study explores the social and economic activities that promote telework¹ and its impact on human resource management (HRM) in Japan. Telework is a technological innovation and is now closely associated with efforts to advance artificial intelligence (AI). Nonetheless, it refers to a business practice that was not widespread in Japan until 2020. However, following the COVID-19 crisis, on April 13th, 2020, Japanese Minister of Economy, Trade and Industry, requested the heads of the Japan Chamber of Commerce and Industry and other small- and medium-sized enterprises associations encourage to promote telework and other actions from the standpoint of preventing expansion of the novel coronavirus disease. Companies should encourage office workers to shift to telework at their homes in principle and reduce the number of workers showing up at offices by at least 70 % even if such workers need to go to the offices for unavoidable reasons, except those who need to continue working to maintain social functions².

After this request companies had seriously to take on two social responsibilities which include providing their employees with a safe working environment and simul-

¹ The term ‘telework’ is equivalent to such expressions as ‘telecommuting’ or ‘work from home’. Blount [2015] explains the reason for the existence of multiple expressions by the diversity of socio-cultural backgrounds when the newly phenomenon occurs. It is essential to consider the labour and employment system, the degree of ICT development, and the social and historical background in which the phenomenon developed. However, in terms of telework style, even though some different expressions exist, common factors are location and work-related travel. In Japan, the term ‘telework’ is generally used.

² Ministry of Economy, Trade and Industry. METI Requested Related Associations to Promote Teleworking of Workers and Other Actions as Responses to Government’s Declaration of State of Emergency Involving Novel Coronavirus Disease. 2020. 13 April. URL: https://www.meti.go.jp/english/press/2020/0413_002.html (accessed 11.02.2021).

taneously keeping their business operations going. Consequently, various companies that had been reluctant suddenly begun to introduce the system.

Such online communication was also uncommon in Japanese HRM field. However, the impact of ‘telework’ is currently considered a powerful means of changing traditional work styles. It is thought to alter relationships seen in the conventional office setting at various stages of the labor process: communication, leadership, and the exchange of authority and obedience.

Although AI has already been introduced in some fields such as the digital economy and society, artificial sociality (AS) has not yet been widely adopted. Further, the concept itself is not clearly defined. This article aims to contribute to the formulation of an AS concept from a Japanese perspective with foci on HRM studies. We will first briefly describe AI’s progress in Japan, and the development of work processes in places where telework has been implemented. In order to emboss distinctive character of Japan’s telework, we refer some relevant research in countries other than Japan. We will thereafter discuss recent workplace changes that have occurred with the manifestation of the COVID-19 crisis, by referring to some recent research outcome. To put in precisely, we will ascertain the levels and parts of society the telework has been introduced. In other words, we have classified which domain of Japanese corporations adopted telework. Then we attempt to systematize data in these research contents to classify the elements which associate with artificial sociality.

Japan’s Concern of Artificial Intelligence and Artificial Sociality

In fact, with the ongoing development of the information and communication technology (ICT) over recent decades, AI is already an essential aspect of the Government’s public policy. This can be seen in its proposed concept of ‘Society 5.0’³:

The Japanese government has proposed Society 5.0 as a vision of the future society the country is aiming for. Society 5.0 is a society that follows the hunter-gatherer society (Society 1.0), agrarian society (Society 2.0), industrial society (Society 3.0), and information society (Society 4.0). Society 5.0 refers to a human-centered society that achieves both economic development and solutions to social challenges by means of systems that fuse cyber spaces with the real world in a highly integrated manner.

AI has been expanding in various areas. However, in comparison with other industrialized societies in relation to the introduction of AI, Japan is seen to have fallen behind. According to a survey conducted by Japan’s Ministry of Internal Affairs and Communications in 2016, only 5% of Japanese workers answered that artificial intelligence had been introduced in their places of work (the choices were: ‘it has already been used’ and ‘it has been introduced but has not used yet’)⁴. By contrast, 13.7% of workers in the United States answered that AI had been introduced in their workplac-

³ Ministry of Internal Affairs and Communications of Japan. The White Paper on Information and Communications in Japan. 2019. URL: <https://www.soumu.go.jp/johotsusintoeki/whitepaper/eng/WP2019/2019-index.html> (accessed: 30.01.2021).

⁴ Ministry of Internal Affairs and Communications of Japan. The White Paper on Information and Communications in Japan. 2016. URL: <https://www.soumu.go.jp/johotsusintoeki/whitepaper/eng/WP2016/2016-index.html> (accessed: 30.01.2021).

es⁵. Besides, only 5.6 % of Japanese workers responded that they planned (including having an actual timetable as well as being under consideration) to introduce AI in the future, while 16.5 % of the workers in the United States gave a positive answer⁶. This is about three times higher. Moreover, depending on future efforts, the difference may widen further. There is also a similar tendency in studies concerning telework. In formal situations in Japan, telework has not been widely used. Besides, where the proportion of telework is small, and where AI has been widely discussed and already introduced in some areas of Japan, the concept of AS has not widely penetrated in any Japanese social scientific discourse.

Rezaev and Ivanova [2018] show the following three dimensions of AS: the first one is human–human interactions that proceed through machines, for instance, in social networks. The second is human-machine interactions — from programmers to computer service, and from users at work to playing games children. The third is a computer-computer (screen-to-screen) communication. Therefore, the social nature of AI could be conceived in two ways: ‘strong’ artificial sociality does not yet exist; it would be in the ability of AI to interact spontaneously and be emotionally involved in communication; ‘weak’ artificial sociality is an empirical fact of the participation of AI in various social interactions, in this sense, it is expanding more and more today.

As previously mentioned, in order to elucidate the AS concept and its usage, it is essential to accumulate case study data from various fields. Thus, we attempt to provide some current ideas regarding ICT effects on workplace communication in Japan’s HRM studies that relate to AS, especially as an example of the first type,— human–human interactions of AS.

The Characteristics of the Japanese-Style Employment System and New Trends in Communication

The term ‘telework’ as often used comes from the terms ‘tele’ which refers to a distant place, and ‘work’. Therefore, we derive a concept of a flexible work style that utilizes information and communication technology, regardless of location or time, from a distance. The Ministry of Internal Affairs and Communications indicates the new work style⁷:

In the realm of how people work, working arrangements have drawn attention in which people do not belong to a company or organization but instead work in a freelance capacity via the Internet on one-off projects or short-term assignments.

In addition to economic activities, a change is seen in lifestyles⁸:

It is now possible to elicit emotional responses from a large number of people transcending real-world human relationships and the geographical bounds of daily life by posting the

⁵ Ibidem.

⁶ Ibidem.

⁷ Ministry of Internal Affairs and Communications of Japan. The White Paper on Information and Communications in Japan. 2019. URL: <https://www.soumu.go.jp/johotsusintoeki/whitepaper/eng/WP2019/2019-index.html> (accessed: 30.01.2021).

⁸ Ibidem.

videos or music, pictures, stories, or messages that you make, on various sharing sites or social media channels.

A person that performs telework is called a teleworker. Further, according to the definition of the Ministry of Land, Infrastructure, Transport and Tourism of Japan, a teleworker is a person that works using ICT outside their workplace⁹. According to the ministry's survey, telework is divided into three types¹⁰:

- (a) working from workers' home residences,
- (b) working from satellite offices, which refers to teleworking for those who work in a fixed place such as a satellite office, shared office place,
- (c) mobile works, which refers for those who work outside of their office places, such as their clients' offices, hotels during business trip, and coffee shops during their travel.

Telework has not always been popular in Japan as it is now. The introduction of telework in Japan began in the 1980s. This was mainly to support female office workers that found it a little difficult to commute to their office following childcare and other family reasons. In the 1990s, the government supported the telework measures for reducing costs by decentralizing offices from metropolitan areas, developing local economies, and encouraging social participation of people with disabilities and the elderly. Many companies have introduced telework for peripheral workers to have flexible working hours and locations, rather than for core male employees. Until this stage, the promotion policy of telework was positioned to balance between people's economic activities and their lives.

In 2013, however, Japanese government released the 'World's Most Advanced IT Nation Creation Declaration' where they regarded and advocated telework as a part of the work style reform which was necessary for creating an IT nation. Thus, it is a part of pressing for improvement in Japanese companies' traditional employment management.

Although telework has been encouraged for many years in Japan, it has not yet penetrated to workplaces of Japanese. Even though the government encouraged telework and media has focused on telework, the actual implementation has not followed.

The Relationship between 'Membership-type' Employment and HRM

Introduction of telework has delivered a heated debate on Japan's employment system. The study on characteristics of Japan's employment has been conducted not only by Japanese researchers, but also foreign experts. It has long been mentioned that although the feature of Japan's employment system is characterized by its history of industrial development, it requires environmental adaptation.

Supposedly, the primary characteristic of Japanese companies was a professional system in which work is allocated to people and their wages determined by reference to qualifications and promotion based on the performance of the duties assigned to

⁹ Ministry of Land, Infrastructure, Transport and Tourism of Japan. Heisei 29nendo terewa-ku jinko jittaichosa chosakekka gaiyo [Survey on the Telework Population and Work Condition in 2017]. 2018. URL: https://www.mlit.go.jp/crd/daisei/telework/docs/29telework_jinko_jittai_gaiyo.pdf (accessed: 30.01.2021). (In Jap.)

¹⁰ Ibidem.

them. Nihon Keizai Dantai Rengokai (Japan Business Federation) indicates that the character of the ‘Japanese-style employment system’ consists of adherence to the three leading characteristics [Nihon Keizai Dantai Rengokai, 2020: 11]:

- (1) batch hiring of new graduates from school,
- (2) long-term and lifetime employment until retirement age, and
- (3) seniority-type wage system.

These three characteristics have been a feature of Japan’s employment system since the end of World War II [Sugayama, 2011].

Further, it is quite common in the Japanese society to assume that the long-term employment system basically continues until retirement age. Academic specialists in this area tend to choose the term ‘long-term,’ rather than ‘lifetime’. This issue tends to be discussed in terms of the method of work assignment and the education and training of workers’ skills. Therefore, it is indivisible from the condition of the labor market and the education system in that country. Marsden [1999] proposed to set up the criteria for comparing employment systems, when examining the relationship between managerial authority and worker obligations. He takes account of these two characteristics for identifying Japan’s employment character by connecting jobs to works and organizing tasks into jobs. It is essential for Japanese companies to train their employees, as it prioritizes allocation of employees’ workplace according to companies’ needs, rather than each employee’s ability at the time of the contract. The point is to assign people to the workplace (function) rather than to a specific job. Although this method is limited to apply to large corporations, it is regarded as an absolute Japanese model. Since then, discussions on similar ideas have been delivered in Japan.

This employment system is defined as the ‘membership-type employment’. The job demarcation and wage determination under the ‘membership-type employment’ contrasts with what is found in many Western companies in which job-based wage system is present¹¹ [see also: Hamaguchi, 2014]. In the Western ‘job-based type’ employment system, employees’ salary is defined by the job content and its value in the local labor market. Since the contract presupposes the job existence, theoretically the contract will not be renewed when the job disappears, and the employee loses employment. In Japanese corporations, under the ‘membership-type’ employment, the job descriptions are not prepared for the majority of employees. This is also the case for new graduates’ batch hiring. They are hired without specifying certain job demarcations. Employees are in charge of various duties when they get ordered. Therefore, even if their current jobs disappear, they are relocated to another work and thus the employments continue. As a result, a long-term stable employment is established as a corporate system. Employees are theoretically required to take charge of any job at anywhere, according to companies’ orders. Since the job content changes at the companies’ orders, the necessary training program is executed by the company’s responsibility during business hours.

‘Membership-type employment’ begins with the bulk recruitment or collective adoption of new graduates from universities and high schools. When hired, a member (employee) considers themselves part of a company. In other words, freshmen would

¹¹ Hamaguchi K. Addressing the Problems with Japan’s Peculiar Employment System. *Nippon.com*. July 24. 2013. URL: Addressing the Problems with Japan’s Peculiar Employment System | *Nippon.com* (accessed: 30.01.2021).

enter a certain company and work there until retirement. Further, students start hunting for jobs while still in school, and their potential duties are not limited; what follows is a human resources development process for promotion during which the employee will take charge of various kinds of jobs within the company.

Concerning the batch hiring of new school graduates, job skills are transmitted through inhouse training. The inhouse training is basically conducted by face-to-face communication, as freshmen and junior workers learn by watching the seniors' performance. This type of training is referred to as 'on the job training (OJT)'. Through this process, direct supervisor observes and evaluates how workers carry out their duties and what attained out of their work. This supervisor's observation is important to evaluate employees' abilities and ensure the effectiveness of training. At the same time, that observation is also important in evaluating employees' achievement.

Consequently, the direct supervisors' evaluation of employees has large effects upon these employees' promotion. In other words, the results of the evaluation are the major factors that are considered in the decision to upgrade their employees. This is the fundamental norm and form in developing human resource management. Hence, in this seniority principle at the work system, the length of one's work is related to both an employee's wage rises and the record of promotion within the company.

The salary is rather linked to the employee's grade on the status ladder in the company. Therefore, remuneration involves seniority-type wages, in that their wages are raised every year based on the progress in job performance ability, which is reflected in the upgrading of the ladder. This seniority-wage rise suggests that the number of years of service and work experience will eventually improve the job performance ability.

The Japanese-style employment system, which focuses on hiring and also training, designated as 'membership-type employees', has various merits. However, as the business environment changes, challenges are expected to direct orientation towards job-type. Nihon Keizai Dantai Rengokai [2020: 12—13] has formulated an understanding of these challenges in these terms:

While many companies, especially large ones, place importance on the batch recruitment of new graduates, mid-career recruitment has been relatively restrained. This makes it difficult for those that could not get the desired job during the difficult employment environment to re-challenge themselves¹². It is also difficult to develop high employability outside the company because the company's internal measure system promotes human resources suitable for that particular company. Consequently, the labor market's fluidization was hindered. Further, the external labor market was insufficiently developed. This is considered one of the reasons why job-type employment did not spread in Japan.

However, when telework increases, it is likely that the traditional work allocation and seniority system will have become less effective, and a job-based employment system

¹² This is considered to be the case for a large-scale school graduates in the economic downturn period of the 1990's, who are called 'the ice age generation'. Those generation who experienced this employment difficulty and could not get an appropriate job at the initial employment stage did not go well in their subsequent job change. Since they were placed at a disadvantage in Japanese labor market, there have been deliberation as it is regarded that the society was unfair between generations.

would replace it. Well-known companies such as KDDI, Hitachi¹³, Shiseido, and Fujitsu have announced that they will shift to a ‘job-type’ personnel system¹⁴.

Previous Research on Factors in the Features of Telework/Telecommuting

The trends of the abovementioned large corporations may have some influence on the Japanese companies’ employment system. However, each company’s actual choice can be conditioned on multiple factors. In this section, by reviewing previous research on telework and human resources management, we will systematize the components for the social relationship at these online workplaces. In this process, we can find possible references in the new communication style’s deliberation, so-called AS. This will also provide us with essential points for considering possible changes in Japanese places of work after the rapid prevalence of telework under the COVID-19 crisis.

Telework is defined as a work organization that implies remote working (distant from the company’s main premises and without the physical presence of any person in charge of controlling the teleworker’s job execution) and resorts to ICTs. The lack of the physical presence of a manager or supervisor raises issues that differ from the traditional management systems in workplaces [Dambrin, 2004]. In telework, it is highly possible for managers, supervisors, and workers not to share workplaces and time. Therefore, telework brings about various issues that differ with the conventional management style in workplaces. The issues include elements that influence the work environment, information exchange, management process (especially managers’ roles and skills), and time or timing in the management process.

First, there are personal factors that are related to the workplace and work environment. The work environment is a significant source for productivity; therefore, the management is expected to be responsible for the preparation and provision of a suitable one. However, in the case of telework or telecommuting, a necessary working environment tends to be prepared and set up by the individual and family initiatives. Therefore, environmental related issues appear to be closely connected to personal and family features. Additionally, productivity or measured work outcomes are evidently influenced by personal and family factors. These personal characteristics include the workers’ gender [Maruyama, Tietze, 2012; Pigini, Staffolani, 2019]; personality, educational background, family features [Smith, Patmos, Pitts, 2018]; occupation [Maruyama, Tietze, 2012], as well as the quality of the home-office environment that they can provide [Morganson et al., 2010; Nakrošienė, Bučiūnienė, Goštautaitė, 2019]. Research on the remote work environment in normal times can consider the dispatch

¹³ Hitachi is a Japanese company that Ronald Dore analyzed in his well-known book ‘British Factory, Japanese Factory’ published in 1973. Since then, Hitachi has remained one of the most famous and influential companies in Japan. Currently, the chairman of Hitachi is the chairman of Nihon Keizai Dantai Rengokai, who has a great influence on the statements from the business community on the issue of the Japanese employment system. On the other hand, for several years, Hitachi had positioned the senior management class as a global key position to prepare the introduction of Global Talent Management (GTM) system. In line with the introduction of GTM which is often used in multinational companies, it is reported that Hitachi started to introduce job descriptions for their global key positions. See Owada N. Sekai 30 man nin wo Jobu gata ni tankan, Hitachi ga sodaina jinjikaikaku ni idomu honto no ryū! [The Real Reason Why Hitachi is Taking on a Grand Personnel Reform by Converting 300,000 People around the World to a Job Type]. XTECH. 2020. August 19. URL: <https://xtech.nikkei.com/atcl/nxt/column/18/01400/081300001/> (accessed: 11.02.2021) (In Jap.).

¹⁴ Nihon Keizai Sangyo shinbun Indigoburū kaichō shibata reiji — jobu-gata koyō e no gekiyaku (SmartTimes) [Indigo Blue Chairman Koji Shibata — Powerful Drug for Job-Type Employment (SmartTimes)]. 2020. August 26. URL: <https://www.nikkei.com/article/DGXKZ063032520V20C20A8XY0000> (accessed: 11.02.2021). (In Jap.).

at the customer's office [Morganson et al., 2010]. Further, personal attributes such as gender and childcare were significantly reflected on the workers outcomes [Feng, Savani, 2020].

The second element is the organizational and technological factors that influence the work environment. Although telework places are often provided by the workers, organizational, ICT equipment and technology support is often needed, as these are usually crucial resources leading to a favorable work outcome. The lack of or insufficient ICT support may serve well to reduce the viability of working from home for many people. In fact, it may be one of the reasons why telework has not hitherto made inroads in Japan. Support should be offered when the requirements for the work to be done are not present in the immediate work environment. This is described as the absence of a teleworker-environment fit [Bentley et al., 2016]. The teleworker-environment fit is necessary to secure an appropriate workspace [Nakrošienė et al., 2019], work-life balance [Morganson et al., 2010] and reduce workplace isolation [Golden, Veiga, Dino, 2008]. This outcome is related to personal productivity and satisfaction, overcoming stress and anxiety about one's career [Nakrošienė et al., 2019], and experiencing a sense of inclusion [Morganson et al., 2010].

The third element is exemplified as information exchange. Difference in styles of information exchange will influence the outcome of work, that is, the increased use of mobile technology, which causes a sense of autonomy [Fujimoto et al., 2016] and reduced communication with colleagues, which induces anxiety [Golden, Fromen, 2011; Nakrošienė et al., 2019]. They further bring about a difference in the means of communication, change the knowledge acquisition, and affect the satisfaction level [Lee, Shin, Higa, 2007; Smith et al., 2018]. Additionally, the workers' perception and choice of suitable media has changed [Lee et al., 2007] as new technologies and tools are created in response to the needs at tele-workplaces. The use of these new tools will reduce loneliness and improve communication and work outcomes.

The fourth element is the management process for telework. Although telework does not share a place, it is a process of organized labor. Further, in order to achieve telework, organizational management and coordination are required [Dambrin, 2004]. Studies on telework have examined the two factors fundamental to coordination and management; technology and the person involved. The former is related to media and technology as tools, and the latter reveals the relationship between the manager and their subordinates. These include managers' skills, the effects of remote managers' supervision on associates [Golden, Fromen, 2011], and changes in the employees' perception of management. These factors evidently influence the workers' output. Further, the frequency or duration of telework are the factors that control a manager's influence over the output.

From such previous research, it is inferred that in the management process of telework, the manager's role in providing detailed instructions diminishes. In fact, it is now considered outdated. This is because employees have increasingly had to make autonomous decisions in the execution of their work. Thus, the managers' skills are related to providing directional support similar to that provided by a coach in sports [Groen, van Triest, Coers, Wtenweerde, 2018]. Over time, employees will also focus on output. Supervisor support will also increase employee satisfaction if they contribute

to employees' progress towards achieving these goals. Thus, from the employees' point of view, the manager's role has increasingly become unconvincing or redundant [Masuda, Holtschlag, Nicklin, 2017].

Evidently, time would have some effects on the management process. Steward [2000] showed that gradual changes in employee perceptions and expectations for the management were affected by the passage of time. By experiencing simultaneous communication, management's effectiveness is either strongly realized [Dambrin, 2004], or the personal sense of time and working style changes [Fujimoto et al., 2016]. The use of mobile technology has greatly overturned the conventional management premise that companies and factories are established by workers' gathering in a common workplace and simultaneously sharing time and space.

The State of Emergency and Dramatic Expansion of Telework

In May 2020, the implementation of telework in Japan expanded significantly in response to the COVID-19 crisis, especially after the proclamation of the state of emergency. The first COVID-19 infection in Japan was confirmed in mid-January 2020. Since then, the number of infected people has spread rapidly throughout the country, reaching 1,000 in mid-March, then over 2,000 at the end of March, and over 3,000 in early April. On April 7th, the government issued an 'emergency declaration' for seven prefectures: Tokyo, Kanagawa, Saitama, Chiba, Osaka, Hyogo, and Fukuoka. It was based on the provisions of art. 32, para. 1 of the Revised Special Measures for New Influenza Measures Law [Nihon Seisansei Honbu, 2020: 1]. After that, and by mid-April, the 'emergency declaration' was extended to all prefectures. On the 13th April 2020, Japanese Minister of Economy, Trade and Industry, issued the request for the promoting telework.

Japan's life and economy have changed in an attempt to prevent the spread of COVID-19: there are refrains from going out, government requests to suspend business commuting, and school closures. In response to this, telework has suddenly been promoted among Japanese companies. Further, many universities started to conduct online classes. Under this state of emergency, telework has immediately become a significant issue for management in response to the government's request, though it had been a support measure for employees who had commuting-inconvenience.

JILPT¹⁵ survey on the Application of Telework/Work from Home

In reference to the JILPT survey, we can see how telework was established as a 'new way of working' after the spread of the COVID-19 pandemic and the cancellation of the declaration of emergency by the government [Takami, 2020]. All the research samples included those that applied telework during the period of the state of emergency. In this section, we focus on full-time workers and confirm the relationship between the applicability of telework.

Taking advantage of work-life change by the COVID-19 pandemic, the 'working from home/telework' has expanded rapidly. Prior to the state of emergency in early April, three-quarters of the companies had not even introduced telecommuting as an

¹⁵ JILPT — The Japan Institute for Labour Policy and Training.

optional form of working (see fig. 1). During the state of emergency, telecommuting spread throughout the nation. Particularly, under the state of emergency, telework has adapted to almost every employee, including male core members in the company to protect their safety and economic activity. It was very different from four decades ago, when telework was applied only for the company's flexible or peripheral employees.

Figure 1 The number of days of 'working from home / teleworking'

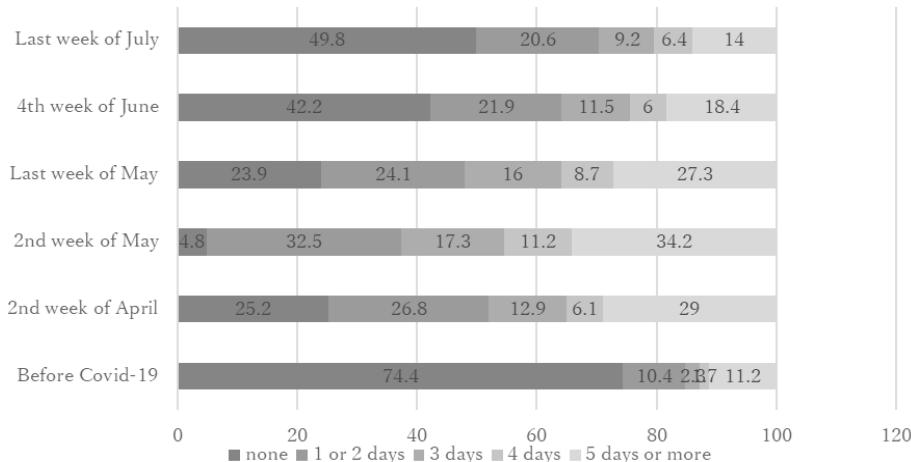


Fig. 1. The number of days of 'working from home/teleworking'¹⁶

This can be confirmed by data for the period before the declaration of the state of emergency. In the second week of April three-quarters of companies introduced more than one telework day per week. During this time, 29% of the companies telecommuted for five days or more per week. Over the 'second week of May' the percentage rose significantly. More than 95% of the companies introduced telework for more than one day, and over one-third of the companies applied telework for five days or more per week. However, the proportion of 'not implementing' began to increase when the government lifted the state of emergency on May 25, and afterwards, that is the last week of May onwards. In the last week of May, approximately a quarter of the companies reported that they had completely stopped telework. In the 4th week of June, this percentage further increased to 42.2%. Moreover, by the 4th week of July, about half of the companies had stopped telework (see fig. 1).

By referring to the JILPT survey, we identify some characteristics in the applicability of telework during the period of the state of emergency. Takami [2020] selected full-time employees from the JILPT May survey data and found clear trends in the applicability of telework.

Table 1 shows the condition of the applicants' working from home under the state of emergency. First, this table indicates that telecommuting did not apply to all kinds of workers. Evidently, from the 'total' column, the larger ratio of applicants is categorized into

¹⁶ Source: Takami [2020].

'did not work from home' while less than 30% worked for 'three days or more a week'. The number of people that worked from home three days or more a week is characterized by being male, highly educated, working for an information industry, service industry, manufacturing industry, large companies, and in urban areas such as Tokyo.

**Table 1. Applicants working from home during the state of emergency
(Full-time work before the COVID-19 pandemic)***

Telework per week		0 (%)	1 or 2 days (%)	3 days or more (%)	N
Total		49.8	20.6	29.6	652
Gender	Male	46.7	21.5	31.7	460
	Female	57.3	18.2	24.5	192
Academic backgrounds	High School Graduates	65.6	14.6	19.8	96
	Vocational school/Junior college Graduates	59.1	17.0	23.9	88
	Undergraduates/Postgraduates	44.9	22.4	32.7	468
Industry type	Construction industry	84.8	6.1	9.1	33
	Manufacturing industry	48.5	20.8	30.7	231
	Electricity, gas, heat supply, water supply	60.0	20.0	20.0	10
	Information and communication industry	29.6	14.8	55.6	108
	Transportation industry	48.3	31.0	20.7	29
	Wholesale/retail	62.3	17.0	20.8	53
	Finance/insurance business	53.1	31.3	15.6	64
	Real estate business	42.1	47.4	10.5	19
	Education and learning support business	54.5	31.8	13.6	22
	Service industry	55.4	14.5	30.1	83
Company size	Less than 29 people	53.8	17.9	28.2	39
	30–299 people	57.1	22.1	20.9	163
	300–999 people	52.2	20.0	27.8	115
	More than 1,000 people	45.1	20.3	34.6	335
Personal annual income before COVID-19	Less than 3 million yen	68.0	9.3	22.7	75
	3–5 million yen	58.3	20.6	21.1	204
	5–7 million yen	43.9	23.9	32.3	155
	7 million yen or more	39.9	22.0	38.1	218
Residential areas	Greater Tokyo area (1 metropolitan area and 3 prefectures)	35.1	21.4	43.5	271
	Kansai area (3 prefectures)	57.0	20.0	23.0	100
	Other areas	61.6	19.9	18.5	281

* Source: Takami [2020]. The table is made with some selected items.

Results of Working from Home

There has been some research on the actual penetration of telework ever since the government lifted the state of emergency. In terms of teleworkers satisfaction and productivity, as a general assessment, workers appear to appreciate telework. Although the survey outcome indicates that telework is viewed as inefficient, their reported satisfaction level is not negative. Approximately 60 % of the respondents answered that they were satisfied, where 18.8 % of them answered 'satisfied' and 38.2 % answered 'somewhat satisfied' [Nihon Seisansei Honbu, 2020]. After this survey, Nikkei HR conducted a similar survey from July 30th to August 7th, 2020 that involved the registered members of 'Nikkei Career NET' (the number of valid responses was 735). According to this survey, 76 % of the respondents experienced telework (working from home) during a state of emergency (April to May), while approximately 90 % of them answered that they would like to continue telework in the future. Enatsu et al. [2020] report on teleworking and employees' problems and their psychological conditions from 3,073 respondents from questionnaire surveys conducted in April and August 2020. They found that those who have experienced telework report that they give a positive evaluation to the continuation of telework.

As for the productivity, the Japan Productivity Center conducted a survey on the concerns of the workers regarding the spread of the new coronavirus infection, in May 2020. The survey targeted 1,100 employees aged 20 and over, excluding self-employed and family workers. More than 30 % of the respondents claimed to have realized that efficiency had improved in terms of the telework style of work raised in the survey. Meanwhile, 7.2 % answered that 'efficiency has improved' and 26.6 % answered that 'efficiency has slightly improved'. However, 66 % replied that efficiency had declined, of which 41.4 % had indicated that efficiency had 'slightly reduced' and 24.8 % answered 'decreased' [Nihon Seisansei Honbu, 2020].

Enatsu et al. [2020] provide a similar outcome. In particular, they indicate that people might develop proficiency in remote work (online literacy) which leads to workers' productivity growth. In other words, people who engage in a lot of remote work spend less time in meetings, and their skill grows through work, and think that remote work is comfortable. Those who have experience in remote work respond that they do not have communication difficulties in this time of remote work introduction.

Enatsu et al. [2020] identify several factors which hinder telework penetration into Japanese workplaces. They point out that the fit between telework and work assignment influence the employees' working style and their psychology. People who find it difficult to bring their work online tend to report the negative feeling on telework. They have a hard time in securing space, especially setting furniture. Experienced individuals can occupy online devices by their own, but there are problems with furniture properties. Other survey has indicated that although companies allow home use of company-owned PCs necessary for teleworking, they have not offered any other allowance. If it goes onto full swing in the future, these problems need to be fixed¹⁷.

¹⁷ Mitsubishi Research and Consulting. (16 Nov) Tele wa-ku no roumukanri nadonikannsuru jittaichosa [Survey on Labor Management of Telework]. Ministry of Health, Labour and Welfare. 2020. November 16. URL: https://www.mhlw.go.jp/stf/newpage_14849.html (accessed: 30.01.2021). (In Jap.)

In addition, when the management of working hours is the same as that of normal conditions, and many companies manage their attendance online¹⁸, workers are under the stress due to the fact that the time constraints are the same as usual. Furthermore, those who have experienced remote work prior to the COVID-19 situation finds the improvement and growth of their own skills through remote work. This can be understood that the skills of remote work made it easier for individuals to conduct autonomous learning, but it is not clear for individuals who do not have experiences to have problems in their learning.

Looking at the overall trend from these survey results, telework is not recognized as a new working style for all workers. It is not presently an established option and might only be a temporary expansion during the state of emergency. However, on this occasion, core employees with higher educational backgrounds, those commuting for larger companies in Tokyo, and those that are the typical core members of the traditional Japanese employment system, will also be aware of the possibilities of various work styles.

In particular, the improvement of online literacy has brought about the possibility of replacing the complex communication that was actually experienced in the Japan's workplace, leading to their sense of growth, challenges, and their comfortable feelings [Enatsu et al., 2020]. It is expected that further research can more precisely identify online literacy, skills and learning abilities individuals acquire through repeated teleworking.

Expected Changes in the Work and Life in Japan

Following the progress of online communication, changes in various situations are expected to occur. Particularly, the media is paying attention to various changes, including lifestyles and employment systems requirements. In this section, we address how the changes are introduced in the mass media, and further examine whether these changes occur nationwide.

Remote work has influenced their sense of psychological distance from the company and this has become an opportunity to review the relationship between work and life. Over the last six months, people's thoughts what takes to deepen and establish their own personality and to enrich their lives outside of their work have changed [Enatsu et al., 2020].

In terms of lifestyle changes, suburban areas have become popular as there is no need to commute to the city center offices. There are some articles in newspapers and journals on moving to a rural area and how 'going to work' online occurs in such circumstances. The term 'workation' has also been used domestically. The term comes from work and vacation. It means to telework while on a break, especially at a travel destination. Some real estate companies are promoting villas by encouraging reforms the work style that involves a shift to holiday resorts.

Changes in the forms of employment have also been discussed on mass media. There is a view that job-type works likely to expand. The characteristic of Japanese-style employment has been the collective recruitment of new graduates, which coincides with the structured seniority and wage systems. Contrastingly, in job-type employment

¹⁸ Ibidem.

experienced personnel in sales and technology have an advantage in searching and landing a job. Their salary increases when they change jobs by being employed by a company that offers a higher salary or so. There is a caution in job type employment in terms of easy dismissal. For instance, ‘employees are more likely to be dismissed’ following the shift of employment in the job-based employment system, developed by telework¹⁹. Some will be dismissed following poor-quality work and duties, or their inability to perform as required.

Nevertheless, human resource management must be influenced by changed human–human interactions. The COVID-19 impact has brought about changes in people’s communication at workplaces. In the case of COVID-19 and lockdown, the workers who have been assigned to the telework, must stay and work at home. A big problem has been raised about human-to-human communication from our encounter with the sudden barrage of telework. What was previously considered vital is no longer viewed as essential to the productivity of a workplace. Consequently, the primary communication style among Japanese companies is deemed to be utterly changed.

ICT-based work style has, to some extent, become a new stream in Japan as they have adapted to Japanese customs and workplaces. Whether or not this workstyle coexists with the old conventional system for the future is something to be examined by further research. Regarding communication in the management process, the perception of appropriate management and the shared value behind it is emphasized. Communication is essential for maintaining coaching. Evaluation by skill development on the employee and supervision process is also an important factor for the employees’ promotion and wage increase. Therefore, the supervisor’s daily supervision and evaluation is a significant subject in Japanese management. However, as telework progresses, the appropriate support and evaluation style will change. Even in Japan’s case, especially among large companies, job-based employment comes to be seen as appropriate because it is easier to measure work output.

Concluding Remarks: Possible Link to Artificial Sociality

The extent of telework penetration influences the utility of human resources, more broadly, the Japanese corporation’s characteristics. Consequently, a hybrid type of face-to-face with online communications can also develop. These must greatly influence the day-to-day details of human resource management in the future. It must also have a decisive effect on the job training and patterns of recruitment. Although several articles have indicated that problems exist with new employees’ education, there have not yet been made any large-scale research in this matter.

This issue, in relation to hiring younger employees, can have a more powerful impact on communication style and defusing corporate value into employees. Survey results show experienced workers who have assigned to telework have not reported serious problems in online workplace communication. Newly hired employees are trained and, therefore, organizationally socialized through detailed supervision and close communication. The prospective issue is how communication and value sharing will be carried out via ICT.

¹⁹ Kuhara Y. Senmon Rinrai ‘jobu-gata koyō’ Kaua’i [Expansion of Professional Human Resources ‘Job-Type Employment’]. *Tokyo Shinbun*. 2020. September 28. P. 1, 3. (In Jap.)

However, even though many workers that have experienced telework prefer this work style, the penetration of telework is currently obviously limited. The difference exists between the size and location of companies and the level of employees' skills. As previous research review shows, there are still severe problems for the company side, such as the cost of introducing ICT and supporting or preparing the home office for the employee. Large corporations may implement it in their workforces. Further, it may be appropriate for some highly skilled workers. Telework has already been implemented to some extent in some major companies. There is also a marginal development in online work from rural areas and in small and medium-sized enterprises. The further evolution of technology and tools is vital for the broader acceptance of telework. Particularly, if the problem of the COVID-19 rises again in the future, it is possible that the next telework request will come out, and the situation may gradually progress in the future.

On the other hand, for well-funded companies, it is clear from previous research outcomes that telework and online work can break the physical and spatial barriers of the office. Spatial issue supports the further globalization of companies. The fact that large companies like Hitachi are trying to reduce the infinitive number of jobs in membership type and shift to job types, which is one of the characteristics of Japan's employment system. This shift is also largely related to the problem of globalization.

Hofstede [2018] discusses AS based on cross-cultural issues, which come from his original 'international business idea'. Under the COVID-19 crisis, people had to pay attention to domestic telework to prevent the virus from spreading nationwide. Ironically, whether domestic telework will characterize employment nationwide is currently a topic of interest. However, it is presently not that widespread. Nonetheless, there is a strong possibility that telecommunication-oriented work will develop. Further, this comes into focus when a company's international workforce that is variously located must communicate with each other, make decisions, and work co-operatively. In these times of restricted travel, Japan's headquarters of many multinational corporations are enthusiastic about telework, and are more likely to collaborate with their subsidiary offices which are operating overseas. In these companies, a complex mix of various values, norms, beliefs, and rituals exists. In current circumstances, this becomes even more complex for multinational corporation. Misunderstandings that lead to malfunctioning systems are very likely to occur [Hofstede, 2018: 5] when people differ in their values, norms, beliefs, and rituals, but have to collaborate in a single technical environment, or using a single formal organizational structure.

A multi-level analysis is required when we view AS in relation to Japan's cultural components. We shall further deal with the changes in existing communication. Changes in personnel deployment and the utilization of telecommunications are expected to encourage AS, in which both formal social structures and cultural 'unwritten rules' matter. Cross-cultural issues will turn out to be important in many socio-technical systems especially in terms of the AS conception. In terms of human resource management, it is necessary to incorporate corporate values and ethics for their code of activities, even in foreign lands.

In Japan, face-to-face communication still plays a crucial role in human resource management. However, since other types of communication such as online networks have been steadily introduced, they confront the need to implement a new evaluation

system, other than evaluation by observing employees in the company, or by face-to-face communication. This is among the reasons to introduce ‘job-type’ employment system, which clarifies employees’ duties in the context of changes to workplace communication. These changes may be the earliest stages of the emergence of a new international work standard or communication style. The role of AS in how the members of a work community communicate within the multinational corporation’s network is now a matter for further research and analysis.

References

- Bentley T. A., Teo S. T. T., McLeod L., Tan F., Bosua R., Gloet M. (2016) The Role of Organisational Support in Teleworker Wellbeing: A Socio-Technical Systems Approach. *Applied Ergonomics*. No. 52. P. 207—215. <https://doi.org/10.1016/j.apergo.2015.07.019>.
- Blount Y. (2015) Pondering the Fault Lines of Anywhere Working (Telework, Telecommuting): A Literature Review. *Foundations and Trends in Information Systems*. Vol. 1. No. 3. P. 163—276. <https://doi.org/10.1561/2900000001>.
- Dambrin C. (2004) How Does Telework Influence the Manager-Employee Relationship? *International Journal of Human Resources Development and Management*. Vol. 4. No. 4. P. 358—374. <https://doi.org/10.1016/j.apergo.2015.07.019>.
- Enatsu I., Kanki N., Takao Y., Hattori Y., Fumoto H., Yatera A. (2020) Shingata koro-nauirusu ryūkō-ka de no shūrō-sha no seikatsu gyōmu kankyō to shinri kōdō [Living/Working Environment and Psychology/Behavior of Workers under the COVID-19] (Discussion Paper No. 33). URL: https://www.works-i.com/research/paper/discussionpaper/2020/DP_0033.html (accessed: 11.02.2021). (In Jap.)
- Feng Z., Savani K. (2020) COVID-19 Created a Gender Gap in Perceived Work Productivity and Job Satisfaction: Implications for Dual-Career Parents Working from Home. *Gender in Management*. Vol. 35. No. 7/8. P. 719—736. <https://doi.org/10.1108/GM-07-2020-0202>.
- Fujimoto Y., Ferdous A. S., Sekiguchi T., Sugianto L.-F. (2016) The Effect of Mobile Technology Usage on Work Engagement and Emotional Exhaustion in Japan. *Journal of Business Research*. Vol. 69. No. 9. P. 3315—33230. <https://doi.org/10.1016/j.jbusres.2016.02.013>.
- Golden T. D., Fromen A. (2011) Does it Matter Where your Manager Works? Comparing Managerial Work Mode (Traditional, Telework, Virtual) across Subordinate Work Experiences and Outcomes. *Human Relations*. Vol. 64. No. 11. P. 1451—1475. <https://doi.org/10.1177/0018726711418387>.
- Golden, T. D., Veiga J. F., Dino R. N. (2008) The Impact of Professional Isolation on Teleworker Job Performance and Turnover Intentions: Does Time Spent Teleworking, Interacting Face-to-Face, or Having Access to Communication-Enhancing Technology Matter? *Journal of Applied Psychology*. Vol. 93. No. 6. P. 1412—1421. <https://doi.org/10.1037/a0012722>.

- Groen B. A. C., van Triest S. P., Coers M., Wtenweerde N. (2018) Managing Flexible Work Arrangements: Teleworking and Output Controls. *European Management Journal*. Vol. 36. No. 6. P. 727—735. <https://doi.org/10.1016/j.emj.2018.01.007>.
- Hamaguchi K. (2014) Wakamono to Rodo [The Youth and Labour]. Tokyo: Chuko Shinsho. (In Jap.)
- Hofstede G. J. (2018) Artificial Sociality. URL: <https://geerthofstede.com/wp-content/uploads/2018/05/Artificial-sociality-2018-05-22-Web-version.pdf> (accessed: 30.01.2021).
- Lee H., Shin B., Higa K. (2007) Telework vs. Central Work: A Comparative View of Knowledge Accessibility. *Decision Support Systems*. Vol. 43. No. 3. P. 687—700. <https://doi.org/10.1016/j.dss.2006.11.007>.
- Marsden D. (1999) A Theory of Employment Systems: Micro-Foundations of Diversity. Oxford: Oxford University Press.
- Maruyama T., Tietze S. (2012) From Anxiety to Assurance: Concerns and Outcomes of Telework. *Personnel Review*. Vol. 41. No. 4. P. 450—469. <https://doi.org/10.1108/00483481211229375>.
- Masuda A. D., Holtschlag C., Nicklin J. M. (2017) Why the Availability of Telework Matters: The Effect of Telecommuting on Engagement via Goal Pursuit. *Career Development International*. Vol. 22. No. 2. P. 200—219. <https://doi.org/10.1108/CDI-05-2016-0064>.
- Morganson V. J., Major D. A., Oborn K. L., Verive J. M., Heelan M. P. (2010) Comparing Telework Locations and Traditional Work Arrangements: Differences in Work-Life Balance Support, Job Satisfaction, and Inclusion. *Journal of Managerial Psychology*. Vol. 25. No. 6. P. 578—595. <https://doi.org/10.1108/02683941011056941>.
- Nakrošienė A., Bučiūnienė I., Goštaitė B. (2019) Working from Home: Characteristics and Outcomes of Telework. *International Journal of Manpower*. Vol. 40. No. 1. P. 87—101. <https://doi.org/10.1108/IJM-07-2017-0172>.
- Nihon Keizai Dantai Rengokai [Japan Business Federation] (2020) Keiei rodo seisaku tokubetsu iinkai hokoku 2020-nenban [Report of Special Committee on Management and Labor Policy 2020 Edition]. Tokyo: Keidanren Publisher. (In Jap.)
- Nihon Seisansei Honbu [Japan Productivity Center] (2020) Shingata koronavirusu no kansen kakudai ga hatarakuhito no ishiki ni oyobosu chōsa [Survey on How the Spread of Coronavirus Infection Affects the Workers' Consciousness]. Tokyo: Japan Productivity Center. (In Jap.)
- Pigini C., Staffolani S. (2019) Teleworkers in Italy: Who are They? Do They Make More? *International Journal of Manpower*. Vol. 40. No. 2. P. 265—285. <https://doi.org/10.1108/IJM-07-2017-0154>.
- Rezaev A. V., Ivanova A. A. (2018) Studying Artificial Intelligence and Artificial Sociality in Natural Sciences, Engineering, and Social Sciences: Possibility and Reality. In: N. Callaos

et al. (eds.) *Proceedings of the 22nd World Multi-Conference on Systemics, Cybernetics and Informatics (WMSCI 2018)*. Orlando, FL: International Institute of Informatics and Systematics. P. 68—71.

Smith S. A., Patmos A., Pitts M. J. (2018) Communication and Teleworking: A Study of Communication Channel Satisfaction, Personality, and Job Satisfaction for Teleworking Employees. *International Journal of Business Communication*. Vol. 55. No. 1. P. 44—68. <https://doi.org/10.1177/2329488415589101>.

Steward B. (2000) Changing Times: The Meaning, Measurement and Use of Time in Teleworking. *Time & Society*. Vol. 9. No. 1. P. 57—74. <https://doi.org/10.1177/0961463X00009001004>.

Sugayama S. (2011) Shusha Shakai no Tanjo [The Advent of Shusha Employment System in Japan]. Nagoya: Nagoya Daigaku Shuppankai. (In Jap.)

Takami T. (2020) Zaitaku inmu wa dare ni teichaku shite iru no ka — ‘Kinkyu-ji’ o heta henka o yomu [Who is Working from Home? To Read the Changes that Have Passed through ‘Emergency’] (JILPT Research Eye No. 46). The Japan Institute for Labor Policy and Training. URL: https://www.jil.go.jp/researcheye/bn/046_200916.html (accessed: 30.01.2021). (In Jap.)

DOI: [10.14515/monitoring.2021.1.1895](https://doi.org/10.14515/monitoring.2021.1.1895)



A. Tadesse, W. R. Allen, C. Mitchell-Kernan

INTEGRATING EDUCATIONAL TECHNOLOGY IN EAST AFRICA: ONE SIZE DOES NOT FIT ALL

For citation:

Tadesse A., Allen W.R., Mitchell-Kernan C. (2021) Integrating Educational Technology in East Africa: One Size Does Not Fit All . *Monitoring of Public Opinion: Economic and Social Changes*. No. 1. P. 91–108. <https://doi.org/10.14515/monitoring.2021.1.1895>.

Правильная ссылка на статью:

Тадесс А., Аллен У.Р., Митчелл-Кернан К. Интеграция образовательных технологий в Восточной Африке: один размер не подходит всем // Мониторинг общественного мнения: экономические и социальные перемены. 2021. № 1. С. 91—108. <https://doi.org/10.14515/monitoring.2021.1.1895>. (In Eng.)

INTEGRATING EDUCATIONAL TECHNOLOGY IN EAST AFRICA: ONE SIZE DOES NOT FIT ALL

Azeb TADESSE¹—PhD Candidate, Department of Education

E-MAIL: alemma@ucla.edu

<https://orcid.org/0000-0003-1866-9776>

Walter R. ALLEN¹—PhD, Distinguished Professor, Department of Sociology, Department of African American Studies; Allan Murray Carter Chair in Higher Education

E-MAIL: allen@gseis.ucla.edu

<https://orcid.org/0000-0001-5392-4854>

Claudia MITCHELL-KERNAN¹—PhD, Professor Emeritus, Department of Anthropology

E-MAIL: cmkernan@ucla.edu

<https://orcid.org/0000-0002-9168-326X>

¹ University of California, Los Angeles, USA

Abstract. Globalization and the post-industrial economy, combined with a growing number of the youth, has increased the demand for higher education in Africa. Online learning breaks down temporal and geographic barriers creating a digital learning community that offers an opportunity to expand access and meet skills and training demands. Deployment of technology is a principal consideration in implementing an online learning programme, given its centrality in the teaching and learning process. The integration of technology requires a decision-making matrix that provides a phased review of the suitability of solution as well as locally placed contemplation of relevance and user accessibility. Higher education institutions in low connectivity areas

ИНТЕГРАЦИЯ ОБРАЗОВАТЕЛЬНЫХ ТЕХНОЛОГИЙ В ВОСТОЧНОЙ АФРИКЕ: ОДИН РАЗМЕР НЕ ПОДХОДИТ ВСЕМ

ТАДЕСС Азеб—аспирантка, факультет образования, Калифорнийский университет в Лос-Анджелесе, США

E-MAIL: alemma@ucla.edu

<https://orcid.org/0000-0003-1866-9776>

АЛЛЕН Уолтер Р.—PhD, заслуженный профессор, факультет социологии, факультет афроамериканских исследований, Калифорнийский университет в Лос-Анджелесе, США

E-MAIL: allen@gseis.ucla.edu

<https://orcid.org/0000-0001-5392-4854>

МИТЧЕЛЛ-КЕРНАН Клаудиа — PhD, заслуженный профессор, факультет антропологии, Калифорнийский университет в Лос-Анджелесе, США

E-MAIL: cmkernan@ucla.edu

<https://orcid.org/0000-0002-9168-326X>

Аннотация. Глобализация и постиндустриальная экономика в сочетании с ростом численности молодежи увеличили спрос на высшее образование в Африке. Онлайн-обучение устраняет временные и географические барьеры: создавая сообщество цифрового обучения, оно дает возможность расширить доступ и удовлетворить потребности в навыках и обучении. Распространение технологий, с учетом их центральной роли в процессе преподавания и обучения, является основным фактором при реализации программ онлайн-обучения. Для интеграции технологий необходима матрица принятия решений, обеспечивающая поэтапный анализ его пригодности, а также актуальности и доступности для пользователей в локальных условиях. Высшие

need to be deliberative and innovative in selecting which technology best fits their environment and accommodates their students' limitations. This paper discusses the decision-making process of integrating technology for an online learning pilot project analysed through Rogers' Diffusion of Innovation (DOI) theory applied to educational technological adoption. The approach of the project considered local infrastructure, university ICT capacity, and user Internet accessibility in the selection of online learning solutions. The process provides insights and understanding of the decision matrix, not only on online learning solutions but also in the broader issue of integrating technology into brick and mortar institutions.

учебные заведения в районах с низким качеством интернет-связи должны одновременно проявлять осмотрительность и новаторство при выборе технологии, которая лучше всего подходит для их среды и учитывает ограничения студентов. В этой статье обсуждается процесс принятия решений об интеграции технологии для пилотного проекта онлайн-обучения. В основу анализа легла теория диффузии инноваций Эверетта Роджерса в приложении к внедрению технологий в образование. Предложенный аналитический подход учитывает местную инфраструктуру, возможности университетов в области ИКТ и доступность интернета для пользователей при выборе решений для онлайн-обучения. Полученные результаты касаются матрицы решений и аргументов не только в отношении онлайн-обучения, но и в отношении более широкой области интеграции технологий в обычных учебных заведениях.

Keywords: educational technology, low connectivity, Africa, online learning, internet access

Ключевые слова: образовательные технологии, Африка, онлайн-образование, доступ в интернет, ограниченный доступ в интернет

Acknowledgments. The paper is based on online learning program implementation in East Africa as part of the university partnership between the U.S. and East African universities funded by the U.S. government.

Благодарность. Работа основана на опыте реализации программы онлайн-обучения в Восточной Африке в рамках партнерства университетов США и Восточной Африки, финансируемого правительством США.

Introduction

This article examines the decision-making process of adapting Education Technology (EdTech) for online learning within the context of low connectivity environments. Integration of technology is influenced by several factors, from organizational capacity to user readiness. This study focuses primarily on deliberations on infrastructure challenges, notably, capacity, access, and affordability of Internet connection, central to

facilitating online learning. Rogers' [2003] Diffusion of Innovation (DOI) theory provides a framework for analysing the decision-making process in selecting an online learning platform that addresses the challenges in low connectivity environments.

Online learning has the potential to break down barriers such as time and space through the creation of digital learning community facilitated by technology. The ability to access the online community is vital for the success of any online learning initiatives. Technological resources are the scaffolding on which online instruction is constructed and they are essential for delivering online learning. *Afrobarometer*, in *Africa's Digital Divide and the Promise of eLearning*, found 'a substantial digital divide both across and within [African] countries, reflected in uneven access to resources'¹. Access to these digital resources determine the ability of learners to access and participate in online programmes.

A study on connectivity in African higher education concluded that the state of Internet connection in African institutions was 'too little, too expensive, and poorly managed' [Gaiko, 2006]. The survey found that of the institutions planning to launch digital initiatives such as online learning, only 39% had a written ITC strategy, while of institutions who had mounted online learning applications, 50% had no written policy [ibid.]. Of those that have selected and installed an online learning platform, 22% chose Moodle, 12% Blackboard, 15% Kewi, 12% WEBCT, 8% Caroline, while 12% selected other applications and 19% opted for Distance education [ibid.]. It is important to note, though open to modification, of the package solution chosen, 81% were developed externally. Additionally, although these solutions may appear to be convenient, the potential of customised solution to be responsive to the local context, its long-term affordability, and potential scaling to accommodate programme expansion come with a premium.

Technology serves as the foundation for online learning and communication, and unencumbered access to the Learning Management System (LMS) is critical for the success of the online learning programme and positive student learning outcome. Therefore, deployment of technology is a principal consideration in implementing an online learning programme, given its centrality in the teaching and learning process. This is especially true in low connectivity areas where access to the Internet is limited and fee-based. In low connectivity areas, implementing online learning programmes requires extensive planning and assessments to determine the appropriateness and relevance of components, especially the integration of technology [Moore, 2012]. However, higher education institutions in Africa have been adopting systems without long-term strategy, planning, or reflection on infrastructure and connectivity challenges [Gaiko, 2006; Njenga, 2011]².

Background

The public higher education landscape in Africa has attempted to keep abreast of the changes in the labour market and respond to the demands of the new economy focused on service and technology. Although still in the early days, Information and

¹ Krönke M. (2020) Africa's Digital Divide and the Promise of E-Learning. *Afrobarometer Policy Paper No. 66*. P. 1. URL: https://afrobarometer.org/sites/default/files/publications/Policy%20papers/PP66-africas_digital_divide_and_the_promise_of_e-learning-afrobarometer_policy_paper-14june20.pdf (accessed: 20.02.2021).

² Additionally, see Twinomugisha A., Magochi J., Aluoch S. (2004) *Investigation of Bandwidth Consolidation for Partnership Universities*. African Virtual University. URL: http://www.foundation-partnership.org/pubs/pdf/bandwidth_consolidation.pdf (accessed: 20.02.2021).

Communication Technology (ICT) appears to be central in expanding access and meeting new skills and training demands [Altbach et al., 2009]. In many economies across Africa, where increasing infrastructure and teaching staff are prohibitively expensive, technologies offer the key for increasing access to higher education [ibid.]. However, a wholesale adaptation of educational technology may be challenging and even detrimental. Therefore, there is a need for higher education institutions in Africa to be deliberative and innovative in selecting which technology best fits their environment and accommodates their students' limitations.

Educational technology

EdTech is the intersection of technology and education where innovations and advances in technology are integrated to support, supplement, customize, personalize, and diversify instruction³. It is a relatively new area with varying definitions depending on orientation: technological or educational. Within the technical field, EdTech denotes the development and application of tools such as hardware, software, and processes that enhance and promote education⁴. Within education, it is the 'study and ethical practice for facilitating learning and improving performance by creating, using and managing appropriate technological processes and resources'⁵.

There is no limit or strict definition of which technologies are suitable for inclusion in EdTech. In the early days, many of the software and hardware systems widely used in EdTech were initially developed for enterprise-level application in the private sector⁶. Today, there are several Fortune 500 Companies solely devoted to developing technology exclusively for classrooms and learning. However, there are foundational technologies, software, hardware, and interactive media, that are common to the EdTech eco-system.

EdTech software consists of solutions initially developed for business as well as those specifically designed for education. Early EdTech consisted of adapting enterprise-level software to support education. However, growth in the field has led to a sector solely devoted to the development of EdTech technology and training, raising a total of \$ 962 million in 2019, of which \$ 383 million was for software exclusively designed for learning such as learning platform and classroom management tools⁷.

Management Information System (MIS) serves as a central data repository that collects, organizes, stores, processes, and analyses data⁸. A version of MIS for institutions is called Education Management Information System (EMIS). It is specifically

³ See Edtech 2.0 (2017) The Ultimate Guide to Edtech. *Medium*. June 5th. URL: <https://medium.com/@edtech2/the-ultimate-guide-to-edtech-88b9f984ada0> (accessed: 20.02.2021).

⁴ Lazaro H. (2020) What is EdTech and Why Should it Matter to You? *General Assembly Blog*. May 24th. URL: <https://generalassemblyly.com/blog/what-is-edtech/> (accessed: 20.02.2021).

⁵ See post *What is EdTech?* (2013) *EdTechReview*. February 15th. URL: <https://edtechreview.in/dictionary/119-what-is-edtech> (accessed: 20.02.2021).

⁶ Perry Y. (2018) Could Computing in Education with Cloud Volumes ONTAP. *NetApp Blog*. November 8th. URL: <https://cloud.netapp.com/blog/cloud-computing-in-education-trends-and-challenges> (accessed: 20.02.2021).

⁷ Wan T., Millward T. (2019) US Edtech Funding Already Nears \$ 1 Billion in First Half of 2019. *EdSurge Newsletter*. August 7th. URL: <https://www.edsurge.com/news/2019-08-07-us-edtech-funding-already-nears-1-billion-in-first-half-of-2019> (accessed: 20.02.2021).

⁸ Korde B. (2020) Role of Management Information Systems (MIS) in Education Sector. *MasterSoft*. December 25th. URL: <https://www.iitms.co.in/blog/role-of-management-information-system-in-education.html> (accessed: 20.02.2021).

designed for educational institutions and is central to management and instructional operations. EMIS is used to collect and store student data, including personal data, exam records, and can also include housing and library information⁹. At the institutional level, it includes HR and financial management systems, as well as specialised MIS systems such as learning management systems (LMS). Popular programmes in LMS include Blackboard Learning System, Share-Point LMS, and Moodle¹⁰. LMS combines online course management, communication, and collaboration tools which can include a discussion forum, file exchange, real-time chat, interactive whiteboards, group work, registration integration, quizzes/surveys, grade book, student tracking, and content sharing¹¹.

With the advancement in software, hardware requirements for EdTech have expanded from simply computers to include interactive displays, broadband connectivity and support devices such as routers and servers. More powerful computers, as well as personal computers and other devices have led to increased accessibility of instruction and learning. Laptops, desktops, and tablets allow students to experience the full scope of applied technology from fundamental keyboarding to advanced drafting and programming¹². Interactive technology is another driver of EdTech which allows instructors and students to visualize and connect with materials directly. Touch-responsive multimedia displays provide an opportunity to pull limitless resources into the classroom with a tap or swipe of a screen. The goal of each innovation in EdTech is to create a learning environment that diminishes the distance between instructor and student as well as the student and the material.

Online learning

Online learning is a component of Distance Learning (DL), an education process where the learner and instructor are physically separated, and interaction takes place through telecommunication systems [Moore, 2012]. DL has evolved in line with technological innovations over the last three centuries, from postal based to today leveraging ICT [Kentnor, 2015]. Innovations in technology have unleashed DL and given rise to a new component of DL, online learning. Online learning employs computers and the Internet as a delivery mechanism for courses and instruction. An online course can be taught entirely online over the Internet or in a hybrid or blended format, which uses a combination of online and face-to-face interaction [Nguyen, 2015].

The temporal and geographic limitations of classroom instruction mean that a potentially large number of students are left behind [Asunka, 2008]. Not every student seeking an education can travel, relocate or devote full-time to schoolwork, many have family and professional obligations that limit mobility and flexibility. In many cases, those unable to make this accommodation are female students with

⁹ Davis B., Carmean C., Wagner E. D. (2009) The Evolution of the LMS: From Management to Learning. *The Learning Guild*. October 15th. URL: <https://www.learningguild.com/insights/137/the-evolution-of-the-lms-from-management-to-learning/> (accessed: 20.02.2021).

¹⁰ Ibidem.

¹¹ See post by Balram Korde mentioned above.

¹² Lazaro H. (2020) What is EdTech and Why Should it Matter to You? *General Assembly Blog*. May 24th. URL: <https://generalassemb.ly/blog/what-is-edtech/> (accessed: 20.02.2021).

responsibilities and commitments that prevent them from entering a traditional learning classroom¹³. Hence, online instruction creates a classroom without walls and enrollment limits, accessible to previously excluded populations. As long as students have a proper device and Internet connection, they can participate in the digital learning environment.

Since the introduction of online learning, the number of free and fee-based online programmes available has exponentially increased. However, critics of online learning question its appropriateness for instruction at the college level¹⁴. Foundational to this position is the role classroom education plays in socialising students and the intangible benefits of interaction with classmates, professors, and materials that challenge a student's beliefs and perspectives. In campus-based instruction, students are 'taught by expert educators about how to access, analyze, criticize, synthesize, and communicate knowledge from multiple perspectives and disciplines'¹⁵. Although these learning objectives can be accomplished in online learning, there are questions of their suitability for certain disciplines that require hands-on training and students who might not learn and flourish in the isolated, individualised online learning environment.

Literature and studies of US-based online programmes indicate preconditions, such as quality of early learning, digital literacy, technical capacity, etc., strongly determine success in distance learning programmes [Henschke, 2014; Nguyen, 2015; Kauffman, 2015]. Studies have examined 'test scores, student engagement with the class material, improved perception of learning and the online format, stronger sense of community among students, and reduction in withdrawal or failure' [Nguyen, 2015: 310]. The technical elements of online instruction, such as the ability to pause, rewind, and review lectures at will, and flexibility and convivence of attending class, have been attributed to increased learning and engagement of learners in online mode. However, individual learning outcome is primarily determined by learners' motivation and self-directness [Henschke, 2014].

In many developing countries, online learning is controversial and divisive, and this is the case across Africa. Some view it as an 'inferior form of education providing an isolated learning experience <...> [and] a harbinger of global, Western-dominated educational homogenization'¹⁶. Best practices in online learning from the U.S. and the West often 'embed deep political, epistemological, and cultural assumptions that may be incongruent with the cultural knowledge of users in many communities in Africa' [Mawere, Stam, 2019: 421].

The ongoing COVID-19 crisis that led to school closures has highlighted the importance of connectivity for success in online learning. In the U.S., where Internet access

¹³ Trines S. (2018) The Rise of Online Education in Sub-Saharan Africa and South Asia. *WENR: World Education News + Reviews*. August 14th. URL: <https://wenr.wes.org/2018/08/educating-the-masses-the-rise-of-online-education> (accessed: 20.02.2021).

¹⁴ Samuels B. (2013) Being Present. *Inside Higher Education*. January 24th. URL: <https://www.insidehighered.com/views/2013/01/24/essay-flaws-distance-education> (accessed: 20.02.2021).

¹⁵ Ibid.

¹⁶ Trines S. (2018) The Rise of Online Education in Sub-Saharan Africa and South Asia. *WENR: World Education News + Reviews*. August 14th. URL: <https://wenr.wes.org/2018/08/educating-the-masses-the-rise-of-online-education> (accessed: 20.02.2021).

is more readily available than other countries, studies have found that inadequate connectivity has impacted the quality of education¹⁷. A study by Brookings Institute [2020] found that as many as 30% of students had no Internet access at home. The study acknowledges that structural change to deliver Internet access is a broader issue, therefore, areas with low, and in some cases, no connectivity should consider creative approaches to providing instruction¹⁸. The report features examples of solutions that leverage appropriate technology to deliver lessons to students in low connectivity environments. Hence, even in the world's largest economy, one size does not fit all in terms of digital delivery of online learning.

Low connectivity environments

The *Global Connectivity Index (GCI)* assesses the expansion of ICT infrastructure across the world. It lists the top 79 countries according to their level of technology infrastructure development. The GCI categorizes countries that score 65—85% and are leading the way in ICT and AI innovation as *Frontrunners*, those in the range of 40—64% indicating quick expansion of digitization into industry and economy as *Adopters*, and those with scores 23—39%, being in the early stage of ICT development and focusing on expanding connectivity to greater populations as *Starters*¹⁹. The primary metrics for the index are Internet bandwidth and broadband download speed, which are 'Foundation' indicators on which all ICT innovation takes place. The 2019 GCI did not list an African nation in the *Frontrunner* category and only South Africa, with a score of 43%, made it to the *Adopters* list, however, the *Starters* category included Egypt (37%), Morocco (36%), Algeria (31%), Botswana (30%), Ghana (29%), Kenya (29%), Namibia (28%), Nigeria (27%), Tanzania (24%), Uganda (24%), and Ethiopia (23%). The placement of only eleven African countries across 24 available slots is a strong indicator of the emergence of digital technology in Africa. However, the low scores reflect the uneven distribution, where connectivity outside metropolitan areas is still a significant challenge. This finding is consistent with the results of a Pew Research Center study that found, globally, a correlation between national GDP per capita and the proportion of the population online. In the African context, those online are generally young, educated, well-off, and residents of urban centres, who go online for social and entertainment purposes²⁰.

GCI's methodology points to broadband and bandwidth as the primary measures of connectivity at the national level. Thus, low connectivity environment can be characterised as having inadequate availability of bandwidth, defined by the capacity of the Internet connection, as calculated by the rate of data transfer in uploading and

¹⁷ Kim J.-H. (2020) Lack of Internet Access Continues to Impact Students' Online Learning Experience. *The Cavalier Daily*. September 17th. URL: <https://www.cavalierdaily.com/article/2020/09/lack-of-internet-access-continues-to-impact-students-online-learning-experiences> (accessed: 20.02.2021).

¹⁸ Opalka A., Gable A., Nicola T., Ash J. (2020) Rural School Districts Can Be Creative in Solving the Internet Connectivity Gap — but They Need Support. *Brookings*. August 10th. URL: <https://www.brookings.edu/blog/brown-center-chalkboard/2020/08/10/rural-school-districts-can-be-creative-in-solving-the-internet-connectivity-gap-but-they-need-support/> (accessed: 20.02.2021).

¹⁹ Huawei's Global Connectivity Index 2020. URL: <https://www.huawei.com/minisite/gci/en/> (accessed: 20.02.2021).

²⁰ Silver L., Johnson C. (2018) Internet Connectivity Seen as Having Positive Impact on Life in Sub-Saharan Africa. Pew Research Center Report. URL: https://www.pewresearch.org/global/wp-content/uploads/sites/2/2018/10/Pew-Research-Center_Technology-use-in-Sub-Saharan-Africa_2018-10-09.pdf (accessed: 20.02.2021).

downloading²¹. However, an inclusive and individual-level measure of low connectivity includes ‘insufficient bandwidth, inadequate telecommunication infrastructure, irregular power supply, high cost of technology’ [Suhail, 2008: 377] as these factors greatly determine a person’s ability to access and afford Internet connectivity. Afrobarometer found that only 17 % of respondents across 34 African countries are prepared to engage in Internet-facilitated remote learning²². The study conducted a regression analysis model with device ownership and frequency of mobile phone use and Internet access to develop a digital literacy index. The analysis found that 55 % of adult population were completely unprepared, while another 28 % would require additional resources to participate in online learning [ibid.]. The study also found significant urban-rural divide in connectivity metrics, such as household device ownership (64 % and 29 % respectively) and digital literacy (48 % and 18 %), as well as a positive correlation between reliable electricity and digital literacy.

Technological adaptation

The above-mentioned findings have implications for determining technological adaption for online learning in the African context and other low connectivity environments. Prior to adopting online learning solutions, a deeper understanding of the factors impacting adoption, deployment, and user experience is necessary. In terms of connectivity, much of the LM software for online learning delivery and management were developed for environments with technological infrastructure, large device availability, and strong digital literacy. Therefore, the optimal use of these systems requires a powerful personal device, reliable electricity, and broadband Internet connection as well as digital savviness to navigate lessons. Another potential limitation in the African context is the difficulty in meeting essential technical preconditions for online learning across all population groups.

In low connectivity environments, the cost of Internet connection adds to this burden. In Africa, internet users pay the highest rate in the world as a proportion of personal income, and on average the price of 1GB of data is 7.1 % of the average income²³. To further illustrate this point, the bottom nine countries on the Affordability Drivers Index presented in the 2019 *Affordability Report*, which measures the extent of Internet infrastructure and broadband adoption rates, are on the African continent, while top 10 countries are mainly in Latin America and Asia²⁴.

Theoretical framework

Integration of technology for online learning requires a decision-making matrix that provides phased consideration of the suitability of solution as well as locally placed

²¹ Gakio K., Chimwasa G., Nyareza S. (2006) *African Tertiary Institutions Connectivity Survey (ATICS)*. P. 15. URL: <https://idl-bnc-idrc.dspacedirect.org/bitstream/handle/10625/50710/IDL-50710.pdf> (accessed: 20.02.2021).

²² Krönke M. (2020) Africa’s Digital Divide and the Promise of E-Learning. *Afrobarometer policy paper No. 66*. URL: https://afrobarometer.org/sites/default/files/publications/Policy%20papers/pp66-africas_digital_divide_and_the_promise_of_e-learning-afrobarometer_policy_paper-14june20.pdf (accessed: 20.02.2021).

²³ Alliance for Affordable Internet (2019) *The 2019 Affordability Report*. Washington DC: Web Foundation. URL: https://1e8q3q16vyc81g8l3h3md6q5f5e-wpengine.netdna-ssl.com/wp-content/uploads/2019/10/A4AI_2019_AR_Screen_AW.pdf (accessed: 20.02.2021).

²⁴ Ibid.

contemplation of relevance and user accessibility. Rogers' Diffusion of Innovation [2003], a highly influential theory in higher education, is suited as a framework for understanding the process of technological diffusion through a scaffolding approach of the decision-making process. The diffusion of innovation theory (DOI), although developed to explain individual behaviour, applies to group and organizational decision-making, which is ultimately an individual-based participatory process.

Rogers presents four elements in the process of diffusion: *innovation*, the new idea or practice, *communication channels*, the collection and communication of information to form positions and attitudes on the innovation, *time*, the rate of adoption, and *social systems*, defined as 'a set of interrelated units engaged in joint problem solving to accomplish a common goal' [Rogers, 2003: 23]. These components are central in the DOI decision process, characterised as a process motivated by the desire to reduce uncertainty through information gathering and sharing on the benefits and drawbacks of an innovation ultimately leading to a decision to adopt or reject it.

Rogers' theory also articulates a five-step decision-making process consisting of *knowledge*, a cognitive stage creating awareness and in-depth knowledge; *persuasion*, an affective process weighing pros and cons; the *decision* stage of adopting or rejecting the innovation; *implementation*, consisting of technical assistance and reinvention; and the *confirmation* stage, where the decision is reinforced through external support, the absence of which results in replacement or disenchantment [Rogers, 2003]. This process is central to the adoption of online learning technology in Africa, where it illustrates an institutional consultation process. However, as noted in Njenga [2011], the incomplete nature of the knowledge phase is a significant risk for institutions, leading to unsuited technological adoption followed by replacement or disenchantment.

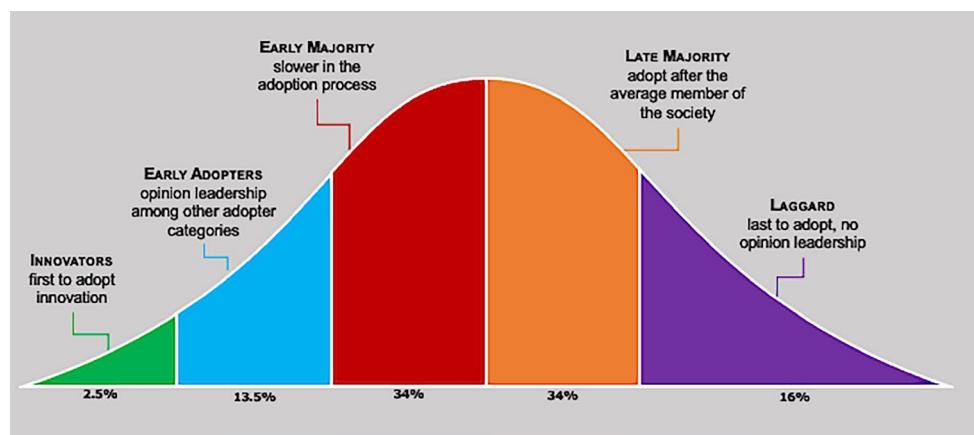


Fig. 1. Rogers' Diffusion of Innovations

An essential component of Rogers' theory is the time-based categorization of adopters charted on a normal distribution curve where those one standard deviation away from the mean are either Early Majority (left) or Late Majority (right), and two standard deviations are Early Adopters (left) or Laggards (right), and finally, those who are in the

left corner beyond two standard deviations are the Innovators, those who joined the system first (see Fig. 1). The relative risk and reward of adoption are analogous to the timeframe of adoption where those that join earlier are able to withstand greater risk and failure of innovation. However, those that join later, as Late Majority, are more risk averse and rely on network information for adoption.

African institutions engaged in technological adoption at this point in the timeframe of digital innovation fall within the Late Majority category, where networks provide the basis for building knowledge and understanding of systems. The high risk of failure and loss attendant with being innovators present a challenge at the individual and organizational level. The harsh funding climate, faced by many institutions, and the significant peril, associated with failure, render innovation a frontier with considerable consequences.

However, Rogers did not provide a quantifiable definition or standard for diffusion and, lacking consistency and cohesion, the theory rendered stagnant [Greenhalgh et al., 2005]. Criticism of DOI includes the focus on the individual in conceptualising the decision-making process, even though it draws from organizational and rational theories from a wide range of fields and disciplines [Ciborra, 1997]. Some authors have questioned the applicability of DOI theory to the diffusion of IT innovations, in light of the complexity and dynamic nature of IT diffusion on individuals and organizations [Hanseth et al., 1996]. Additionally, others have pointed to the lack of in-depth analysis of social factors in analysing the success or failure of adoption of innovations [Grudin, 1988].

Integration of technology

Background and context

Between 2012—2015, a partnership project, funded by the U.S. government, between a US-based research university and a national university in East Africa implemented a pilot online learning programme²⁵. The online learning programme was part of a project to increase equity in terms of providing equal access to the university's programmes across gender groups. The online learning pilot was part of a component aimed at increasing the rate of participation of women in higher education. The process consisted of transforming an existing post-graduate distance learning programme into a hybrid online learning programme to provide a flexible learning environment, thereby increasing access to advanced training for women and others unable to participate in traditional learning programmes.

The post-graduate programme targeted educators, employed in the field, who had previously earned a diploma/degree in education and were pursuing a post-graduate degree to move up in their position. The piloted online programme was a hybrid modality, with lecturers using the online platform to distribute assignments, lecture notes and to conduct online discussions while still maintaining the weekly face-to-face session in the learning centres. However, the length and frequency of face-to-face engagement steadily decreased through the course of the semester. Students use

²⁵ Here and further in the text, when describing the programme and its outcomes, the authors rely on the materials of the report: Lancaster I., Mirembe J. (2015) Evaluation of HED/USAID Women's Leadership Program — Rwanda (2012—2015). USAID. URL: https://pdf.usaid.gov/pdf_docs/PA00KPDP.pdf (accessed: 20.02.2021).

the online platform to complete and submit assignments, conduct research, access readings (both mandatory and suggested supplementary), send questions to instructors, and engage in online discussion forums. Developing a mixed instructional model provided the needed flexibility for increasing the capacity of the institution to provide access to the programme and for participants to advance their training.

The pilot online programme enrolment was mainly mature adults employed full-time who, upon completion of the programme, intended to use their new degree to move forward in their careers. The selection of the post-graduate programme was based mainly on the higher rate of enrolment of mature students in the distance learning version of programme and the potential of this student population to adapt to new learning environments. Learning theory suggests that the learning characteristics of adults are best suited for the independent learning environment of online learning as advanced by Knowles [1970] andragogy. It posits that adult learning, *andragogy* (adult-led), is an approach for learners who have experience and some knowledge and can actively participate in the learning process. This learning process entails different underlying factors than child learning, *pedagogy* (child-led), which is well suited for children who lack experience and knowledge and require teacher directed structured instruction [Knowles, 1970].

The project also provided an opportunity to compare learning outcomes between the previously utilised distance education method and the new system using technology. The enrolment for this programme in correspondence distance mode was 50, but for the online programme, it increased to 70 and finally settled at 74. The use of technology did not seem to deter registration but increased participation. The dropout rate for this programme in the past was low with only 2 to 4 students, however, the online pilot did not experience any dropout.

Technological adaptation decision framework

Knowledge and persuasion

The initial steps to mounting the online learning project consisted of expansion of knowledge of the innovation and persuasion of users at all levels. In designing the pilot, the project had to determine the process of gradually introducing technology, to both students and lecturers. To this end, a readiness assessment was conducted to determine the skills and knowledge of participants and design the onboarding process for the online learning programme. The assessment featured a section on Language Readiness, Technology Readiness exploring the level of computer literacy, previous experience with online learning, and familiarity with various digital resources, as well as learning characteristic profile assessment. Result of the assessment was used to design training and roll out of the pilot online learning programme by addressing identified gaps in skills required to participate in the programme successfully. An additional lecturer assessment, consisting of questions to determine their readiness to teach in the online modality, was conducted among instructors.

All participants, students and lecturers, did not have previous experience with the online learning system and a very small number considered themselves computer literate at the start of the programme implementation. The average participant did not own a computer and did not have regular Internet access at home. This aligns with

findings that indicate that only 28% of the population in Africa is online and only 20% of adults have access to a smartphone *and* a computer, however, 50% can access either a smartphone or a laptop²⁶. A polling of the 74 pilot programme participants found that students planned to download lessons at work, at the learning centre, and in cyber cafes and then working on their assignments offline. Due to the pervasiveness of low digital literacy among participants, it was necessary to build competency and confidence of users gradually.

The lecturers were also relatively new to online learning and received basic training in uploading their syllabi, tracking student progress, and monitoring assignments. During the pilot, lecturers were trained on the platform and online course delivery, thus gaining familiarity with the e-learning and its functionalities. To fully utilize the system, lecturers were trained in converting their courses into interactive digital lessons and integrate video and other digital media. Once lecturers can deliver their lectures through video (live and pre-recorded), the pilot would go entirely online, and in class lectures would eventually be eliminated.

In selecting the devices for the pilot, functionality and accessibility were two primary considerations. Online learning devices for students were assessed based on usability and price. In line with the *Afrobarometer* study²⁷, participants owned mobile phones, however, many did not own laptops or tablets. Compatibility and availability of the appropriate device is central to participant success in online learning. Initially, Windows mobile phones were considered, however, the relatively small screen proved challenging during the testing phase for extended reading and writing assignments. The project settled on laptops at the price range of \$ 300, which fulfilled accessibility requirement while providing users, of all levels, easy functionality and usability.

The computers selected for the pilot proved to be easy to operate for the students and lecturers. Once they received the necessary training, they were able to utilize the devices not only for the programme but also in other areas of their lives. A number of the participants reported using the computer in their classrooms, which increased their students' interest in the lessons. Others commented on how having a laptop motivated them to conduct online research and explore the other functionalities of the device. Overall, participants reported that they considered themselves computer literate and able to use the computers in both their professional and private lives after the training.

The devices were loaned to the students for the pilot period; however, ownership of the laptop was the goal. Upon gaining experience and familiarity with the computer, participants realised the expanded impact on their lives. Participants were interested in a mechanism that would allow a purchase of a laptop, and lecturers were keen to integrate it into their teaching. They also felt that the online learning programme was

²⁶ Krönke M. (2020) Africa's Digital Divide and the Promise of E-Learning. *Afrobarometer Policy Paper No. 66*. URL: https://afrobarometer.org/sites/default/files/publications/Policy%20papers/pp66-africas_digital_divide_and_the_promise_of_e-learning-africanbarometer_policy_paper-14june20.pdf (accessed: 20.02.2021); Alliance for Affordable Internet (2019) *The 2019 Affordability Report*. Washington DC: Web Foundation. URL: https://1e8q3q16vyc81g8l3h3md6q5f5e-wpengine.netdna-ssl.com/wp-content/uploads/2019/10/A4AI_2019_AR_Screen_AW.pdf (accessed: 20.02.2021).

²⁷ Krönke M. (2020) Africa's Digital Divide and the Promise of E-Learning. *Afrobarometer Policy Paper No. 66*. URL: https://afrobarometer.org/sites/default/files/publications/Policy%20papers/pp66-africas_digital_divide_and_the_promise_of_e-learning-africanbarometer_policy_paper-14june20.pdf (accessed: 20.02.2021).

more cost effective, in the long run, eliminating expenses such as making copies, transportation to learning centres, and purchasing supplementary readings.

Decision, implementation and confirmation

The DOI theory framed the process of selecting an online learning system where a wide range of factors are considered in a decision-making process to address uncertainty in adopting the innovation. The analysis of the project within the DOI framework considered local infrastructure, university ICT capacity, and user Internet accessibility in the selection of online learning solution. The pilot project faced several decision points in meeting the primary requirements, including:

- Connectivity—the e-learning platform must have an offline capability to address issues of connectivity and availability of the Internet.
- Ease of usability—the system must be easy to learn and navigate for users, lecturers, and administrators. The assessment results indicated that students and lecturers had limited digital literacy, making an easy-to-use system necessary. The system and interface needed to be intuitive, and as much as possible, free from developer assumptions.
- Alteration—the ability to customize, as needed, at any point in the future without additional costs was essential. Flexibility to add features, change visuals, and the interface was important.
- Scalability—ability to scale up or down the platform as needed. In the early days of the programme, it was anticipated that few features would be utilised; however, as users gained proficiency, additional features would be used widely.

Table 1. Comparison of solutions

	Package solution	Institutionally developed solution
OWNERSHIP	No institutional ownership	Full ownership by the institution
CUSTOMIZATION	<ul style="list-style-type: none">— Requires a programmer with specific knowledge unavailable locally— Expensive, additional fee for customization required	<ul style="list-style-type: none">— Fully customizable locally— Additional manipulation and customization at the discretion of the institution without additional costs
SYSTEM SUPPORT	<ul style="list-style-type: none">— No dedicated support— Local talent difficult to secure	<ul style="list-style-type: none">— Developer support available locally— Institutional capacity developed to serve as the first line of support
OFFLINE CAPABILITY	No offline capacity	Offline option available
LICENSE	Annual license fee	No license fee

As the project was contemplating solutions, there was a national rollout of a top-selling online LMS in various institutions. An informal observation found that the national rollout faced challenges of affordability, capacity to administer the system, and user

connectivity. Therefore, the project conducted a comparative analysis of using the online platform that was part of the national rollout or developing a system that would be more responsive to the requirements discussed above. The result indicated that the cost of adapting and launching the top-selling LMS solution that was part of the national rollout was cost-prohibitive (see *table 1*). Moreover, the institution undertaking the online programme lacked the ICT infrastructure and personnel capacity to administer the system.

The decision-making process resulted in the project opting to build an online learning platform developed by a local firm, with the leasing of a regionally developed platform as an alternative. The information gathering and persuasion stage consisted of consultation with technology partners and the review and testing of a regionally developed platform customised for the African context. However, closer examination revealed the alternative challenges including:

- the solution turned out to be irreconcilable with the limited infrastructure capacity of the institution and presented a number of financial and organizational concerns;
- although the solution offered a certain level of customization, user adoption to the functionality proved to be a barrier;
- expanding use of the platform for use in other online programmes would be cost-prohibitive which raised issues of sustainability. A study conducted by the institution estimated that the use of a leased online learning platform would result in a 46% increase in programme fees.

Challenges during the trial phase of the regionally developed solution proved that an initial investment in building an e-learning platform could respond to the requirements and need for sustainability and could have a lasting impact. There was a sustained search that resulted in securing a national technology firm, which was contracted to develop an online learning solution to project specifications and provide training for the project online platform administrators, who will take over the management and maintenance, as well as lecturers and students.

In addition to information and experience gained from a review of the regionally development solution, there were a number of decision points which led to the development of learning platform. An important consideration was user connectivity given the inadequacy of Internet connection, nationwide and in participants' households. Integrating an App for the platform offered offline functionality and access, a vital feature for low connectivity areas. Moreover, the online learning platform, housed in the internal system at each learning centre site, allowed students to access it through an internal network, neutralising issues with connectivity and cost.

Additional advantages that were a consideration for selection of a locally developed online learning solution included:

- building internal capacity, both human and technological, as well as national capacity;
- direct consultation and input from users embedding an understanding of the local context;

- ability to load the solution locally in satellite learning centres for users to access through the intranet, which allowed to download and upload lessons faster and at no cost;
- increased student enrolment in online programmes at will without additional costs;
- increased number of courses offered at will without additional charges.

Maintenance for the online learning platform was designed at two levels. The institution entered a service contract with the platform developer for secondary level support that would provide institution-based administrators, the first line of support for the platform, with maintenance services. The system was directly housed on the institution's servers which provided IT personnel and the platform administrator with direct access to the solution, where they can assign users and control permissions for levels of access. The fact that the firm was located within a few miles of the institution, facilitated the development of close working relationships, targeted training and support and eliminated the need for additional travel costs to address technical issues with the system.

The final solution for the online learning programme was a customised, scalable online learning platform owned, and not licensed, by the institution, which removed the high cost of leasing with an App functionality for offline use. Eliminating leasing fee for the online learning solution offered the opportunity to re-direct funds allocated for leasing to assist students with payments for computers. A contemplated formula includes allocating a portion of the registration fee towards funding approximately half of the cost of the computer while students would enter a payment schedule for the remaining half of the purchase price.

Conclusion

This pilot study of the integration of EdTech in East Africa provides insights and understanding on the decision matrix and consideration, not only on online learning solutions but also on the broader issue of integrating technology into brick and mortar institutions. It provided an opportunity to explore areas that might require attention and adjustment. The experience of the lecturers and students was critical to understanding the functionality of online learning platforms as a learning environment.

Reducing uncertainty was the driver for the DOI process for the project in light of the limited resources and shortage of institutional and infrastructure capacity. Although the project arrived late in the general context of international development of online learning, nationally, it was innovative and thus, preceded through the DOI steps of building awareness, building consensus, two-step decision, one resulting in discontinuance and one in implementation and confirmation. The complexity of the package system and regionally developed solution previewed were rendered impractical, delivering the relative advantage to the final option of internally developing a customised solution. The ability to create a solution that aligned with requirements and functionality suited to local users offered compatibility, which in DOI is the 'degree to which an innovation is perceived as consistent with the existing values, past experiences, and needs of potential adopters' [Rogers, 2003: 15]. Ultimately, the final decision rested on the solution's relative advantage, compatibility, complexity, and observability.

References (Список литературы)

- Altbach P.G., Reisberg L., Rumbley L. E. (2009) Trends in Global Higher Education: Tracking an Academic Revolution. A Report Prepared for the UNESCO 2009 World Conference on Higher Education. URL: http://www.cep.edu.rs/public/Altbach,_Reisberg,_Rumbley_Tracking_an_Academic_Revolution,_UNESCO_2009.pdf (accessed: 20.02.2021).
- Asunka S. (2008) Online Learning in Higher Education in Sub-Saharan Africa: Ghanaian University Students' Experiences and Perceptions. *The International Review of Research in Open and Distributed Learning*. Vol. 9. No. 3. <https://doi.org/10.19173/irrodl.v9i3.586>.
- Ciborra C. U. (ed.) (1997) Groupware and Teamwork: Invisible Aid Or Technical Hindrance? Chichester: John Wiley & Sons, Inc.
- Greenhalgh T., Robert G., Macfarlane F., Bate P., Kyriakidou O., Peacock R. (2005) Storylines of Research in Diffusion of Innovation: A Meta-Narrative Approach to Systematic Review. *Social Science & Medicine*. Vol. 61. No. 2. P. 417—430.
- Grudin J. (1988) Why CSCW Application Fail: Problems in the Design and Evaluation of Organizational Interfaces. In: *CSCW '88: Proceedings of the 1988 ACM Conference on Computer-Supported Cooperative Work*. New York, NY: Association for Computing Machinery. P. 85—93. <https://doi.org/10.1145/62266.62273>.
- Hanseth O., Monteiro E., Halting M. (1996) Developing Information Infrastructure: The Tension Between Standardization and Flexibility. *Science, Technology, & Human Values*. Vol. 21. No. 4. P. 407—426. <https://doi.org/10.1177/01622439602100402>.
- Henschke J. A. (2014) Andragogical Curriculum for Equipping Successful Facilitators of Andragogy in Numerous Contexts. In: Wang, V. X., Bryan, V. C. (eds.) *Andragogical and Pedagogical Methods for Curriculum and Program Development*. IGI Global. P. 142—168. <https://doi.org/10.4018/978-1-4666-5872-1.ch008>.
- Kauffman H. (2015) A Review of Predictive Factors of Student Success in and Satisfaction with Online Learning. *Research in Learning Technology*. Vol. 23. <https://doi.org/10.3402/rlt.v23.26507>.
- Kentnor H. (2015) Distance Education and the Evolution of Online Learning in the United States. *Curriculum and Teaching Dialogue*. Vol. 17. No. 1. P. 21—34.
- Knowles M. S. (1970) *The Adult Learner: A Neglected Species*. Houston, TX: Gulf Publishing Company.
- Mawere M., van Stam, G. (2019) eLearning in an African Place: how 'alien' eLearning Models are Failing Many in Africa. In: Nielsen P., Kimaro H. (eds.) *Information and Communication Technologies for Development. Strengthening Southern-Driven Cooperation as a Catalyst for ICT4D*. ICT4D 2019. IFIP Advances in Information and Communication Technology. Vol 552. Cham: Springer. P. 422—433. https://doi.org/10.1007/978-3-030-19115-3_35.

Moore M. G. (ed.) (2012) *Handbook of Distance Education*. New York: Routledge.
<https://doi.org/10.4324/9780203803738>.

Nguyen T. (2015) The Effectiveness of Online Learning: Beyond No Significant Difference and Future Horizons. *MERLOT Journal of Online Learning and Teaching*. Vol. 11. No. 2. P. 309—319.

Njenga J. K. (2011) eLearning Adoption in Eastern and Southern African Higher Education Institutions. *A Thesis Submitted in Fulfillment of the Requirements for the Degree of Doctor of Philosophy*. Cape Town: University of the Western Cape, The Department of Information Systems. URL: <https://core.ac.uk/download/pdf/58913753.pdf> (accessed: 20.02.2021).

Rogers E. M. (2003) *Diffusion of Innovation*. New York: Free Press.

Suhail N.A. (2008) Assessing Implementation of Blended Learning in Constrained Low Bandwidth Environments. In: *Special Topics in Computing and Ict Research: Strengthening the Role of Ict in Developing Countries*. Kampala: Fountain Publishers. P. 377—390.

МОНИТОРИНГ МНЕНИЙ

Правильная ссылка на статью:

Мониторинг мнений (ВЦИОМ): январь — февраль 2021 // Мониторинг общественного мнения: экономические и социальные перемены. 2021. № 1. С. 109—128.

For citation:

Public Opinion Poll (VCIOM): January — February 2021. (2021) *Monitoring of Public Opinion: Economic and Social Changes*. No. 1. P. 109—128.

МОНИТОРИНГ МНЕНИЙ: ЯНВАРЬ — ФЕВРАЛЬ 2021

Результаты ежедневных опросов «ВЦИОМ—Спутник». Метод опроса — телефонное интервью по стратифицированной двухосновной случайной выборке стационарных и мобильных номеров объемом 1600 респондентов в возрасте от 18 лет. Выборка построена на основе полного списка телефонных номеров, задействованных на территории РФ. Данные взвешены на вероятность отбора и по социально-демографическим параметрам. Максимальный размер ошибки с вероятностью 95% не превышает 2,5%. Помимо ошибки выборки, смещение в данные опросов могут вносить формулировки вопросов и различные обстоятельства, возникающие в ходе полевых работ.

СОДЕРЖАНИЕ ДАЙДЖЕСТА

ПОЛИТИКА

«СПРАВЕДЛИВАЯ РОССИЯ», «ЗА ПРАВДУ» И «ПАТРИОТЫ РОССИИ»: ПЕРСПЕКТИВЫ ОБЪЕДИНЕНИЯ.....	110
ГОСУДАРСТВО И ОБЩЕСТВО: ЦЕЛИ, ПРИОРИТЕТЫ, ИМПЕРАТИВЫ.....	112

СОЦИАЛЬНЫЕ И ЭКОНОМИЧЕСКИЕ ПРОБЛЕМЫ

ПОТРЕБИТЕЛЬСКИЕ ПЛАНЫ — 2021.....	115
ОБРАЗОВАНИЕ В РОССИИ: ВОСТРЕБОВАННОСТЬ, ДОСТУПНОСТЬ, КАЧЕСТВО.....	117
ГОД РОССИЙСКОЙ НАУКИ И ТЕХНОЛОГИЙ.....	120

ОБРАЗ ЖИЗНИ

СЕМЬЯ, ДРУЗЬЯ И БЕЗОПАСНОСТЬ	122
УДАЛЕННАЯ РАБОТА ПО-РУССКИ: ПЛЮСЫ И МИНУСЫ.....	124
ТРУД — 2021: ЗАПРОС НА БАЛАНС ЖИЗНЬ/РАБОТА	127

ПОЛИТИКА

«СПРАВЕДЛИВАЯ РОССИЯ», «ЗА ПРАВДУ» И «ПАТРИОТЫ РОССИИ»: ПЕРСПЕКТИВЫ ОБЪЕДИНЕНИЯ.....	110
ГОСУДАРСТВО И ОБЩЕСТВО: ЦЕЛИ, ПРИОРИТЕТЫ, ИМПЕРАТИВЫ	112

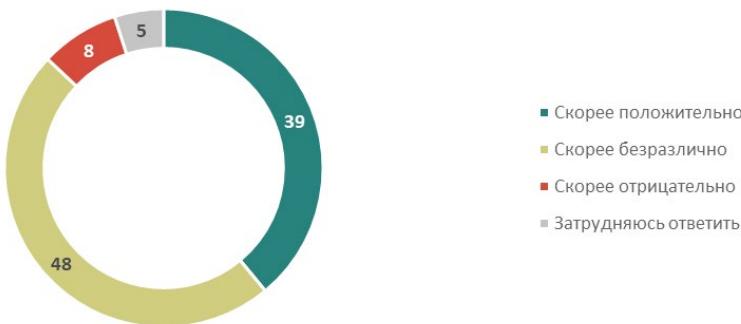
«СПРАВЕДЛИВАЯ РОССИЯ», «ЗА ПРАВДУ» И «ПАТРИОТЫ РОССИИ»: ПЕРСПЕКТИВЫ ОБЪЕДИНЕНИЯ

27 января 2021 г.

Информированность об объединении этих трех партий среди россиян низкая — в той или иной степени о нем знают лишь 28% опрошенных. При этом 46% россиян допускают для себя возможность проголосовать за объединенную партию. Вновь образованная партия «Справедливая Россия — Патриоты — За правду» ориентирована скорее на развитие России по своему особому пути, считают 67% респондентов. С тем, что партия будет выступать за усиление социальной справедливости, согласны 55% наших соотечественников, и 22% предполагают, что ее деятельность будет нацелена на развитие рыночной экономики. Если бы выборы в Государственную Думу России состоялись в ближайшее воскресенье, за объединенную партию «Справедливая Россия — Патриоты — За правду» проголосовало бы 11% россиян.



Рис. 1. В январе три партии, «Справедливая Россия», «За правду» и «Патриоты России», приняли решение объединиться в новую партию, которая будет называться «Справедливая Россия — Патриоты — За правду». Вы знаете об этом объединении или сейчас впервые слышите? (закрытый вопрос, один ответ, % от опрошенных)



*Рис. 2. Как Вы относитесь к решению об объединении трех партий:
«Справедливая Россия», «За правду» и «Патриоты России»?
(закрытый вопрос, один ответ, в % от информированных об объединении партий)*



*Рис. 3. Какое из суждений, на Ваш взгляд,
в большей степени характерно для новой объединенной партии?
(закрытый вопрос, в % от опрошенных, представлены доли согласных с утверждением)*

ГОСУДАРСТВО И ОБЩЕСТВО: ЦЕЛИ, ПРИОРИТЕТЫ, ИМПЕРАТИВЫ

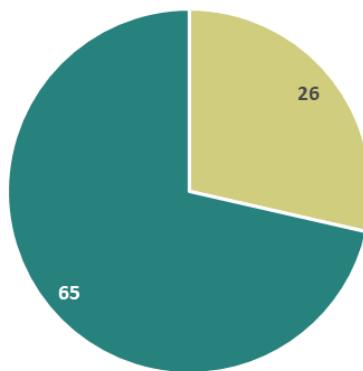
январь и февраль 2021 г.

С утверждением, что главная обязанность государства — всегда заботиться обо всех и оказывать социальную поддержку, согласились 96 % опрошенных россиян. Несколько меньше согласия вызвало утверждение, что политика государства в первую очередь должна быть направлена на сохранение и укрепление традиционных ценностей (83 %). Еще меньше респондентов солидаризовались с тезисом, что «благополучие человека в большей мере зависит от него самого, чем от устройства общества и государственной политики» (64 %). Чуть больше половины наших соотечественников (65 %) полагают, что роль государства в жизни общества должна вырасти, однако четвертая часть (26 %) высказываются за минимизацию госвмешательства.

Топ-5 приоритетов в развитии России, на взгляд респондентов ВЦИОМ, выглядит так: власть должна уделять внимание качеству и доступности медицины (38 %), повышению доступности образования (25 %), открытию новых производств (23 %), решению жилищной проблемы (19 %) и повышению социальных гарантий различным группам населения (15 %). Проблемы обороноспособности (12 %) и другие в первую пятерку приоритетов не вошли. Вопрос о великоледжавности поделил россиян на две практически равновеликие группы. Каждый второй (51 %) согласился с утверждением, что «Россия должна быть великой державой, с мощными вооруженными силами и влиять на политические процессы в мире», однако 41 % опрошенных придерживаются другого мнения, а именно, что «первоочередным приоритетом страны должна быть забота о благосостоянии собственных граждан».

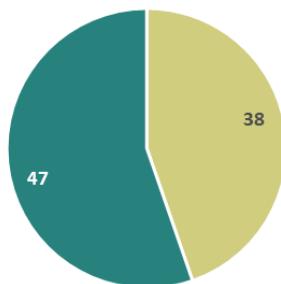


*Рис. 4. Вы согласны или не согласны со следующим утверждением: ...
(закрытый вопрос, один ответ, % от всех опрошенных,
представлены сумма ответов «определенко согласен» и «скорее согласен»)*



- Государство должно минимально вмешиваться в экономику, бизнес и другие сферы жизни
- Необходимо усилить роль государства в разных сферах, в том числе экономике и бизнесе

*Рис. 5. С каким из следующих мнений Вы в большей степени согласны?
(закрытый вопрос, один ответ, % от всех опрошенных)*



- Личные интересы - это главное для человека, их нельзя ограничивать даже ради блага общества
- Людям следует ограничивать свои личные интересы во имя интересов государства и общества

*Рис. 6. С каким из следующих мнений Вы в большей степени согласны?
(закрытый вопрос, один ответ, % от всех опрошенных)*



Рис. 7. Если говорить о развитии России, то, по Вашему мнению, развитию каких направлений власть должна уделить приоритетное внимание?
(закрытый вопрос, не более 2 ответов, % от всех опрошенных)



Рис. 8. С каким из следующих мнений Вы в большей степени согласны?
(закрытый вопрос, один ответ, % от всех опрошенных, представлены сумма ответов «определенко согласен» и «скорее согласен»)

СОЦИАЛЬНЫЕ И ЭКОНОМИЧЕСКИЕ ПРОБЛЕМЫ

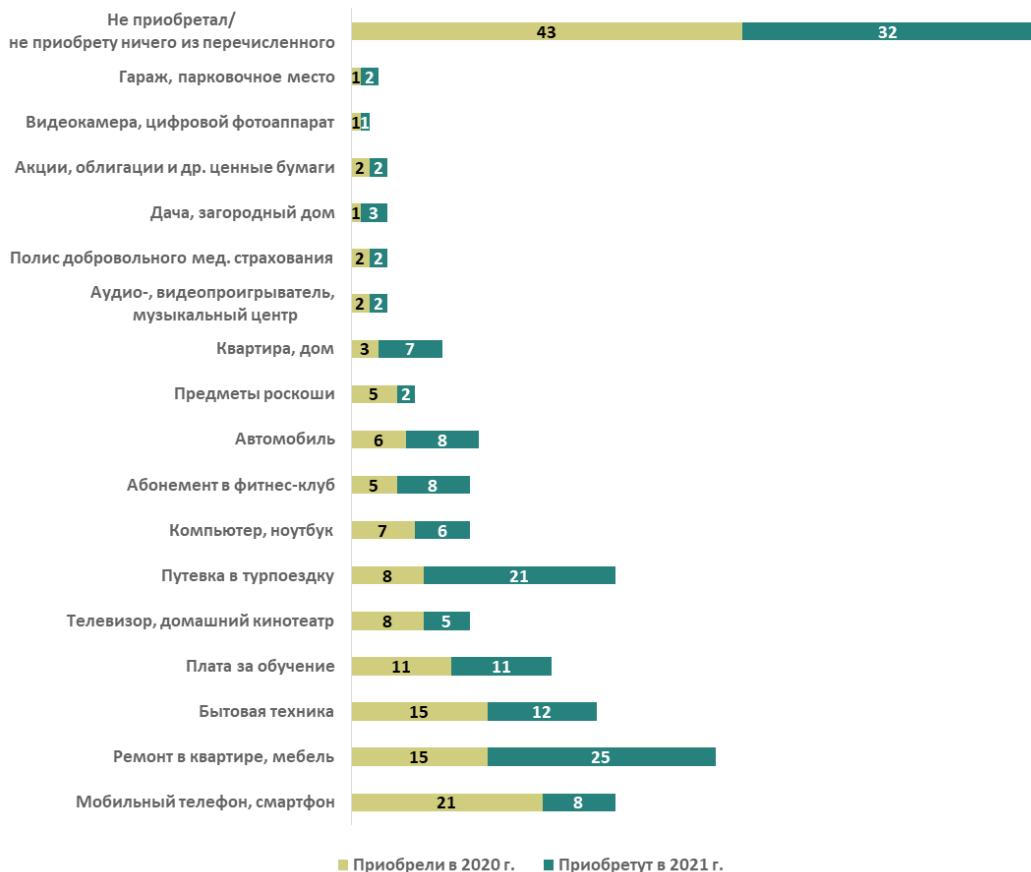
ПОТРЕБИТЕЛЬСКИЕ ПЛАНЫ — 2021.....	115
ОБРАЗОВАНИЕ В РОССИИ: ВОСТРЕБОВАННОСТЬ, ДОСТУПНОСТЬ, КАЧЕСТВО	117
ГОД РОССИЙСКОЙ НАУКИ И ТЕХНОЛОГИЙ.....	120

ПОТРЕБИТЕЛЬСКИЕ ПЛАНЫ — 2021

24 января 2021 г.

Самым частым приобретением года россиян уже на протяжении нескольких лет оказывается мобильный телефон или смартфон: в 2020 г. его купили 27 % наших соотечественников. На втором месте самых частых приобретений прошедшего года стал ремонт в квартире — 21 %. В список популярных приобретений попала также бытовая техника (19 %), по 11 % россиян оплатили образование и купили телевизор или домашний кинотеатр, еще по 8 % — ноутбук или компьютер, приобрели путевку в турпоездку. Владельцами нового автомобиля стали в 2020 г. 7 % респондентов, столько же купили абонемент в фитнес-клуб. Приобретали предметы роскоши и квартиры 6 % наших соотечественников.

В 2021 г. каждый четвертый россиянин планирует сделать ремонт в квартире или купить мебель (26 %). Второй по популярности статьей расходов, запланированной респондентами на текущий год, является туристическая поездка — ее хотят совершить 20 % опрошенных. Бытовую технику планируют приобрести 15 % россиян, еще 12 % оплатят обучение. Купить мобильный телефон или автомобиль в 2021 г. хотят 11 % наших соотечественников.

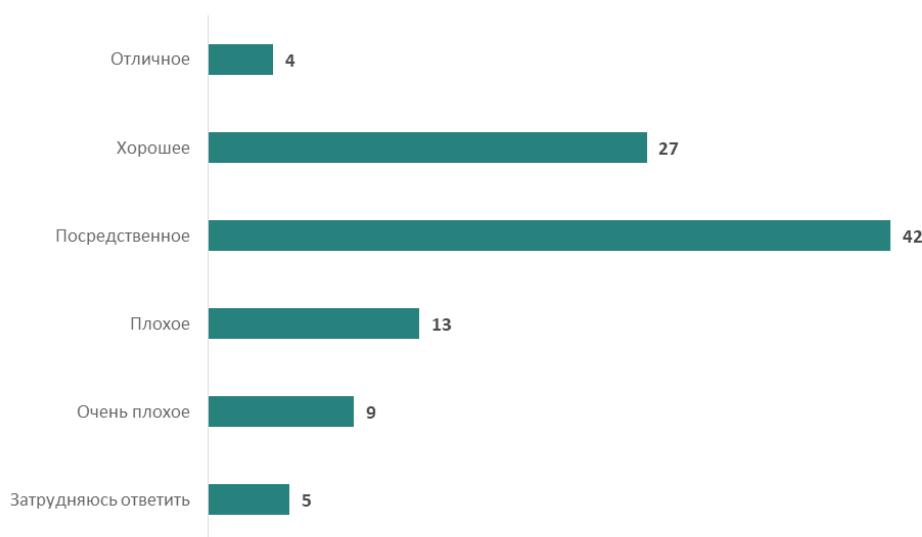


*Рис. 1. Перечислите, пожалуйста, что из следующего списка товаров/услуг Вы приобрели в 2020 г. и планируете приобрести в 2021 г. для себя, своей семьи?
(закрытый вопрос, любое число ответов, % от всех опрошенных)*

ОБРАЗОВАНИЕ В РОССИИ: ВОСТРЕБОВАННОСТЬ, ДОСТУПНОСТЬ, КАЧЕСТВО

14 января 2021 г.

Треть россиян (31 %) считают, что образование в нашей стране «хорошее» или «отличное». Тем не менее 42 % опрошенных оценивают его как посредственное, а каждый пятый (21 %) считает его «плохим» или «очень плохим». Наличие образования воспринимается респондентами как необходимость. Только 5 % вошедших в выборку жителей страны сообщили, что им безразлично, какой уровень образования получит их ребенок. Вузовское образование остается самым востребованным на протяжении многих лет (81 %). Прямую зависимость между уровнем образования и материальным благополучием человека видят 32 % наших соотечественников, большинство же считают, что прямой зависимости не существует (62 %). Факторами, влияющими на образование, получаемое ребенком, по мнению россиян, оказываются материальное благополучие семьи (53 %), желание ребенка учиться (48 %), желание семьи дать ему достойное образование (34 %) и, наконец, индивидуальные способности ребенка (29 %). В целом высшее образование сейчас менее доступно, чем в советское время, полагают 59 % опрошенных.



*Рис. 2. Как бы Вы оценили состояние нашей системы образования?
(закрытый вопрос, один ответ, %)*



Рис. 3. Какое образование Вы хотели бы, чтобы получили Ваши дети, внуки?
(закрытый вопрос, один ответ, %)



Рис. 4. Если говорить о семье, в которой растет ребенок, от чего прежде всего зависит, какое образование он получит?
(закрытый вопрос, не более трех ответов, %)

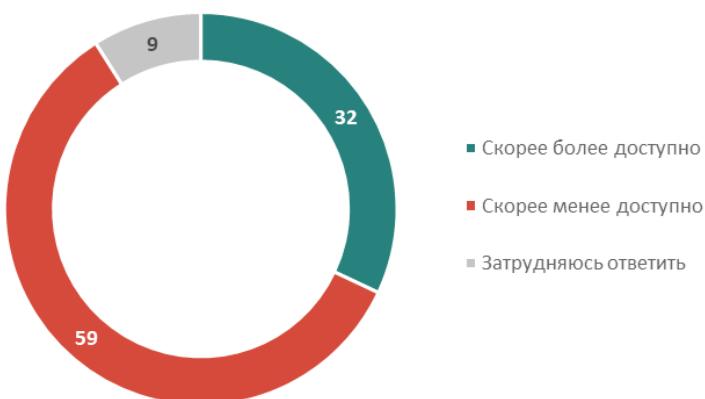


Рис. 5. По Вашему мнению, сейчас высшее образование более доступно или менее доступно для россиян, чем в советские годы?
(закрытый вопрос, один ответ, % от всех опрошенных)

ГОД РОССИЙСКОЙ НАУКИ И ТЕХНОЛОГИЙ

30 января 2021 г.

Самым важным достижением российской и мировой науки последних десятилетий россияне считают борьбу с коронавирусом и разработку вакцины (16%). Второй по частоте ответ — военные разработки/ВПК/новая техника (10%). Немногим менее половины респондентов (44%) уверены, что мировая наука сейчас на подъеме, и треть опрошенных (31%) полагают, что она, наоборот, в застое. Российская наука с разной степенью опережает развитие мировой науки, по мнению 31% россиян, однако более половины наших соотечественников считают, что она скорее находится в позиции отстающего: 34% полагают, что она немного отстает от мировой, а 21% — что значительно. Чтобы ученые смогли сделать прорывные открытия, нужно увеличить финансирование науки, заявили 56% респондентов. Еще третья (34%) придерживается мнения, что необходимо также повысить компетентность ученых, поскольку иначе финансовые вливания не приведут к результату. По мнению каждого второго россиянина (51%), решать судьбу отечественной науки должны сами ученые и научное сообщество, и только каждый пятый (23%) считает, что решения должны принимать государство и органы власти.

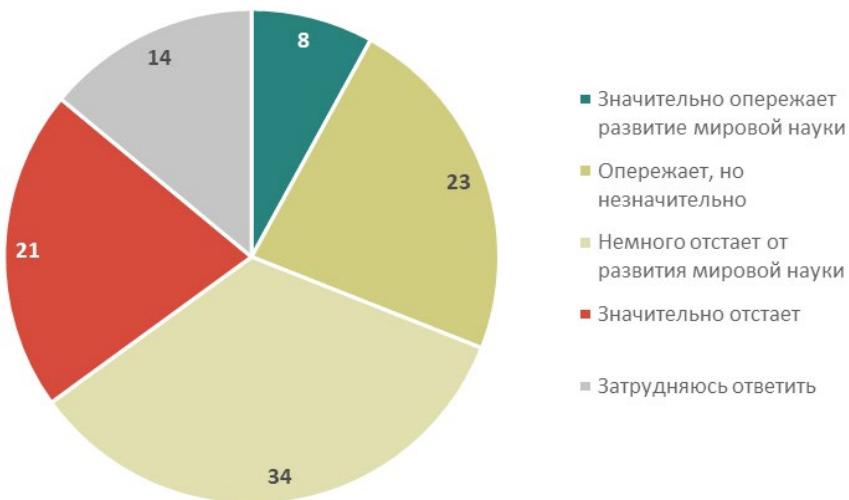


Рис. 6. Восьмого февраля в нашей стране отмечается День российской науки. По Вашему мнению, какими темпами развивается современная российская наука в сравнении с мировой наукой?
(закрытый вопрос, один ответ, % от всех опрошенных)

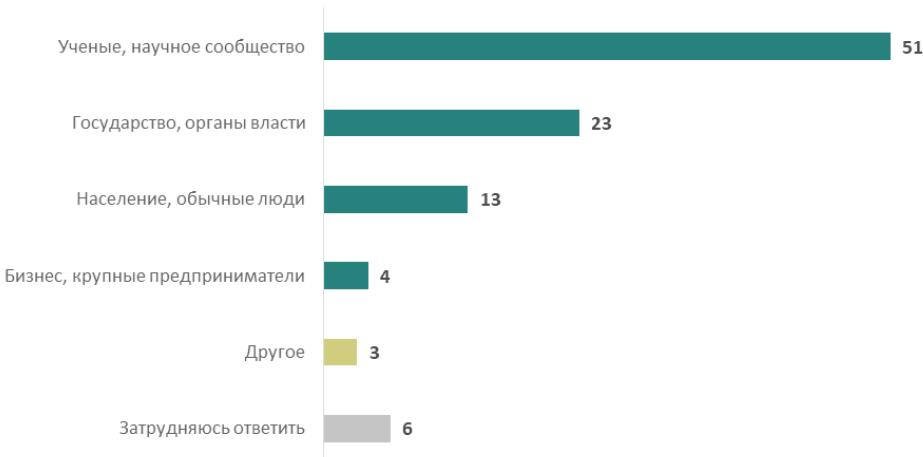


Рис. 7. Кто, на Ваш взгляд, должен принимать решение о том, какие задачи должна сегодня решать российская наука, в каком направлении развиваться? (закрытый вопрос, один ответ, % от всех опрошенных)



Рис. 8. Назовите, пожалуйста, достижения российской и мировой науки последних десятилетий (открытый вопрос, до трех ответов, % от всех опрошенных)

ОБРАЗ ЖИЗНИ

СЕМЬЯ, ДРУЗЬЯ И БЕЗОПАСНОСТЬ	122
УДАЛЕННАЯ РАБОТА ПО-РУССКИ: ПЛЮСЫ И МИНИСЫ.....	124
ТРУД—2021: ЗАПРОС НА БАЛАНС ЖИЗНЬ/РАБОТА.....	127

СЕМЬЯ, ДРУЗЬЯ И БЕЗОПАСНОСТЬ

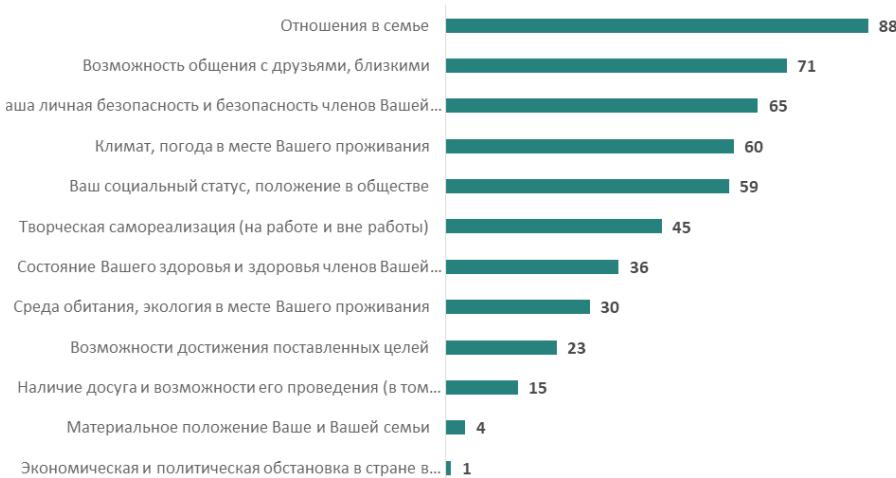
14 января 2021 г.

Отношения в семье — сфера, в которой россияне испытывают наивысшую удовлетворенность¹ (88 пунктов). На втором и третьем месте — общение с друзьями и близкими (71 пункт) и личная безопасность (65 пунктов). Высокая степень удовлетворенности наших соотечественников климатом, погодой в месте их проживания (60 пунктов). Всего на 1 пункт отстает социальный статус в обществе: удовлетворенность им оценивают на 59 пунктов. Творческая самореализация (в работе и вне ее) удовлетворяет россиян на 45 пунктов, а сфера здоровья — личного и близких — на 36 пунктов. Экология оценивается респондентами в 30 пунктах, а наличие возможностей достижения своих целей — на 23 пункта. Досугом и способами его проведения россияне удовлетворены на 15 пунктов. Завершают список сфер материальное положение (4 пункта) и экономическая, политическая обстановка в стране (1 пункт).



*Рис. 1. Насколько Вы удовлетворены следующими сторонами своей жизни?
 (закрытый вопрос, один ответ по строке, % от всех опрошенных)*

¹ Индекс удовлетворенности сферами жизни рассчитывается по ответу на вопрос «Насколько Вы удовлетворены следующими сторонами своей жизни?» как разница между суммами положительных и отрицательных ответов и может принимать значение от –100 до 100 пунктов.



*Рис. 2. Насколько Вы удовлетворены следующими сторонами своей жизни?
(индекс в пунктах)*

УДАЛЕННАЯ РАБОТА ПО-РУССКИ: ПЛЮСЫ И МИНУСЫ

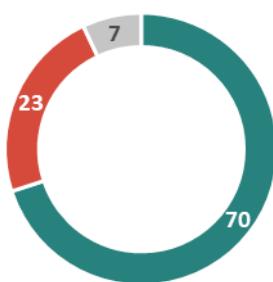
28 января 2021 г.

До пандемии большинство россиян (84 %) работали в основном из офиса / на предприятии. Только 4 % имели опыт удаленной работы и еще 6 % комбинированной — частично удаленно, частично из офиса. В настоящее время три четверти работающих россиян (75 %) трудятся преимущественно в офисе / на предприятии. Каждый десятый респондент (11 %) работает в основном удаленno и еще 12 % — комбинированно. Чаще всего о работе из офиса / на предприятии заявляли жители городов-миллионников (74 %) и до 100 тыс. жителей (83 %). Более чем половине работающих россиян удобно трудиться удаленno (64 %). Треть наших соотечественников сообщили об обратном (29 %). Удобно работать в удаленном режиме чаще женщинам (70 % vs 60 % мужчин), жителям Москвы и Санкт-Петербурга (75 %). Каждый второй россиянин (52 %), работающий удаленno или комбинированно, считает, что плюсов и минусов в таком формате работы примерно поровну. Каждый четвертый (25 %) нашел больше плюсов в удаленном режиме работы, а каждый пятый — больше минусов (20 %). Женщины чаще мужчин видят плюсы в удаленном формате работы (30 % vs 22 %). Большинство россиян (84 %), кто раньше работал преимущественно в комбинированном режиме или ездил на работу, при переходе на удаленный формат не сменили место проживания.

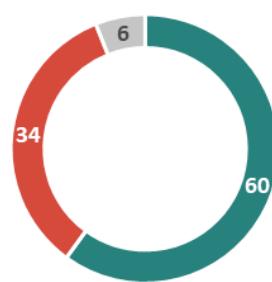


Рис. 3. В этом году из-за ситуации с коронавирусом сотрудники некоторых компаний были переведены на удаленный режим работы. Скажите, а как Вы работаете в настоящее время? (закрытый вопрос, один ответ, % от работающих россиян)

Женщины

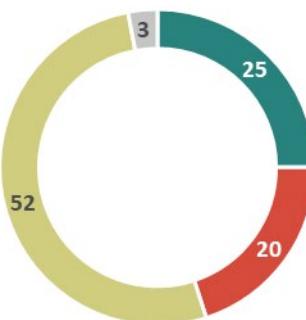


Мужчины



- Скорее удобно ■ Скорее неудобно
- Затрудняюсь ответить ■ Затрудняюсь ответить

*Рис. 4. В целом Вам удобно или неудобно работать в удаленном формате?
(закрытый вопрос, один ответ, % от тех, кто работает удаленно или комбинированно)*



- Скорее больше плюсов ■ Скорее больше минусов
- Плюсов и минусов примерно поровну ■ Затрудняюсь ответить

*Рис. 5. На Ваш взгляд, в удаленной работе больше плюсов или больше минусов,
или плюсов и минусов примерно поровну?
(закрытый вопрос, один ответ, % от тех, кто работает удаленно или комбинированно)*



*Рис. 6. Сейчас я считаю Вам несколько утверждений об удаленной работе, а Вы скажите, согласны ли Вы с ними или нет
(закрытый вопрос, один ответ, % от тех, кто работает удаленно или комбинированно)*

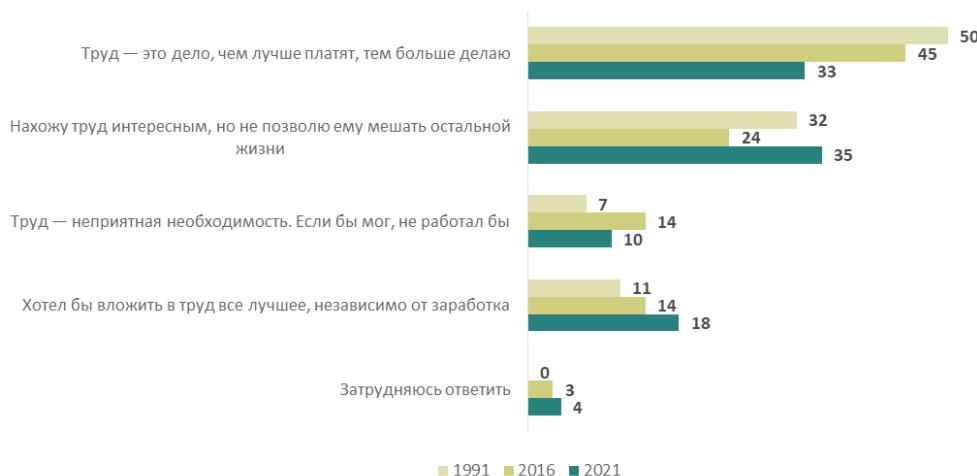


*Рис. 7. Какой формат работы Вы предпочли бы в дальнейшем?
(закрытый вопрос, один ответ, % от тех, кто работает)*

ТРУД—2021: ЗАПРОС НА БАЛАНС ЖИЗНЬ/РАБОТА

16 февраля 2021 г.

Треть работающих россиян находят труд интересным, но не позволяют ему мешать остальной своей жизни (35 %). Почти такая же доля респондентов относятся к труду как к делу: чем больше платят, тем больше и будет сделано (33 %). Вложить в свой труд лучшее, независимо от оплаты труда, хотел бы почти каждый пятый опрошенный (18 %). Только 10 % наших соотечественников считают труд неприятной необходимостью и предпочли бы не работать вовсе. Основная мотивация работать лучше для россиян — достойная оплата труда (44 %). Уважение трудовых успехов сотрудников (19 %), внедрение новых инструментов (18 %), нормальные условия труда (16 %) — все это тоже способствовало бы улучшению продуктивности респондентов на их рабочем месте. Свои трудовые показатели улучшили бы 15 % опрошенных, если бы им предоставляли возможность обучаться профессии в процессе работы, особенно актуально это для молодежи. По мнению россиян, организация труда (13 %) и компетентность руководства (13 %) также сыграли бы роль в улучшении их производительности труда.



*Рис. 8. Скажите, чем для Вас является труд?
(закрытый вопрос, один ответ, % от работающих)*



*Рис. 9. При каких условиях Вы стали бы работать на своем рабочем месте лучше, чем сейчас?
(закрытый вопрос, не более трех ответов, % от работающих)*

DOI: [10.14515/monitoring.2021.1.1790](https://doi.org/10.14515/monitoring.2021.1.1790)



А. В. Смирнов

ЦИФРОВОЕ ОБЩЕСТВО: ТЕОРЕТИЧЕСКАЯ МОДЕЛЬ И РОССИЙСКАЯ ДЕЙСТВИТЕЛЬНОСТЬ

Правильная ссылка на статью:

Смирнов А. В. Цифровое общество: теоретическая модель и российская действительность // Мониторинг общественного мнения: экономические и социальные перемены. 2021. № 1. С. 129–153. <https://doi.org/10.14515/monitoring.2021.1.1790>.

For citation:

Smirnov A. V. (2021) Digital Society: Theoretical Model and Russian Reality. *Monitoring of Public Opinion: Economic and Social Changes*. No. 1. P. 129–153. <https://doi.org/10.14515/monitoring.2021.1.1790>. (In Russ.)

ЦИФРОВОЕ ОБЩЕСТВО: ТЕОРЕТИЧЕСКАЯ МОДЕЛЬ И РОССИЙСКАЯ ДЕЙСТВИТЕЛЬНОСТЬ

СМИРНОВ Андрей Владимирович — кандидат экономических наук, старший научный сотрудник лаборатории демографии и социального управления, Институт социально-экономических и энергетических проблем Севера ФИЦ Коми НЦ УрО РАН, Сыктывкар, Россия
E-MAIL: av.smirnov.ru@gmail.com
<https://orcid.org/0000-0001-6952-6834>

Аннотация. В статье рассмотрена теоретическая модель цифрового общества, основанная на четырех концептах: сверхсвязность, платформизация, датификация и алгоритмическое управление. С помощью модели описан механизм углубления цифровизации общества — от переноса в цифровую среду отдельных практик и социальных взаимодействий до создания социального порядка на основе больших данных. Анализ панельных данных лонгитюдного обследования РМЭЗ НИУ ВШЭ за 2003—2018 гг. и географически привязанных данных проекта «Виртуальное население России» показал, что, несмотря на впечатляющий рост некоторых показателей, цифровизация российского общества носит фрагментарный характер как в разрезе социальных групп и территорий, так и по сферам общественной жизни. Препятствует устранению этого неравенства в первую очередь низкая устойчивость интернет-практик индивидов. Изучение содержания паспортов национальных проектов, утвержденных в 2018 г., позволило оценить перспективы формирования в России цифрового общества и выявить барьеры адаптации населения к изменяющимся условиям.

DIGITAL SOCIETY: THEORETICAL MODEL AND RUSSIAN REALITY

Andrey V. SMIRNOV¹ — Cand. Sci. (Econ.), Senior Researcher at the Laboratory of Demographic and Social Management
E-MAIL: av.smirnov.ru@gmail.com
<https://orcid.org/0000-0001-6952-6834>

¹ Institute for Socio-Economic and Energy Problems of the North, Federal Research Centre “Komi Science Centre of the Ural Branch of the Russian Academy of Sciences”, Syktyvkar, Russia

Abstract. The article considers a theoretical model of digital society based on four concepts: super-connectivity, platformisation, datafication, and algorithmic governance. The model describes how the digitalisation of society deepens: from the transfer of individual practices and social interactions to a new social order based on big data. Analysis of panel data from the 2003–2018 longitudinal survey (RLMS HSE) and geospatial data from the ‘Virtual population of Russia’ project shows that, despite considerable growth in specific indicators, the digitalisation of the Russian society across various social groups, territories and areas of public life appears fragmented. Elimination of this disparity is primarily hindered by the low stability of an individual’s online practices. Analysis of the national projects approved in 2018 helped assess the prospects of the digital society in Russia and identify barriers to public adaptation to the changes.

Ключевые слова: цифровое общество, цифровизация, платформизация, большие данные, цифровое неравенство, Россия

Благодарность. Исследование выполнено при финансовой поддержке РФФИ и БРФФИ в рамках научного проекта № 20-510-00007.

Keywords: digital society, digitalization, platformisation, big data, digital divide, Russia

Acknowledgments. The study was funded by the Russian Foundation for Basic Research (RFBR) and the Belarusian Republican Foundation for Fundamental Research (BRFFR), project No. 20-510-00007.

Введение

Во многих регионах мира формируется цифровое общество, для которого характерны новые формы «техно-социальной жизни» [Redshaw, 2019], ставшие результатами внедрения информационно-коммуникационных технологий. В цифровом обществе упрощаются социальные взаимодействия, повышается информационная открытость, снижаются издержки периферийности. Повсеместное применение цифровых технологий способствует экономическому росту и повышению качества жизни населения¹. В то же время исследователи отмечают и риски, которые несет цифровизация: формирование новых структур социального неравенства [van Dijk, 2013; Grimshaw, 2017], усиление зависимости развивающихся стран от технологических лидеров, рост влияния транснациональных корпораций [Ганичев, Кошовец, 2019], нарушение приватности и использование цифровых технологий для ограничения прав и свобод человека [Matzner, Ochs, 2019].

Хотя количество публикаций, посвященных этой проблематике, в общественных науках за последние годы увеличилось многократно (см. рис. 1), теоретические основы, понятийный аппарат и методы изучения цифрового общества и цифровой социологии только формируются [Lupton, 2015; Selwyn, 2019]. Исследования охватывают как самые общие вопросы, вроде соотношения «реального» и «виртуального» [Колозари迪, Макушева, 2018], сущности «цифровизации» [Dufva, Dufva, 2019], так и конкретные ее проявления в различных сферах жизни. Разработаны теоретические концепции, объясняющие преобразования, происходящие при формировании цифрового общества [Katzenbach, Bächle, 2019]. Тем не менее процессы углубления цифровизации, последовательно охватывающей все более фундаментальные социальные структуры, требуют дальнейшего осмысления.

В данной статье предпринята попытка объяснить углубление и расширение процессов цифровизации общества с помощью объединения четырех теоретических концептов: сверхсвязность, платформизация, датификация и алгоритмическое управление. Чтобы проследить за тем, как цифровая среда проникает в разные

¹ Согласно исследованию Европейской комиссии, 75% опрошенных жителей Европейского союза считают, что новейшие цифровые технологии положительно влияют на экономику, 67% — на качество их жизни. См.: Attitudes Towards the Impact of Digitisation and Automation on Daily Life. Special Eurobarometer 460. Report. European Union, 2017. <https://www.doi.org/10.2759/83566>.

сферах человеческой жизни, изучались практики взаимодействия населения с интернет-средой. Цель исследования — улучшить понимание процессов цифровизации общества, рассматривая интернет-практики в рамках предложенной модели. В фокусе исследования — цифровизация российского общества в 2003—2019 гг. Отдельные аспекты проиллюстрированы примерами из стран Запада и Азии, лидирующих в цифровой трансформации общества. Для более комплексного рассмотрения цифровизации применялись как традиционные данные выборочных обследований населения, так и данные, полученные в цифровой среде.

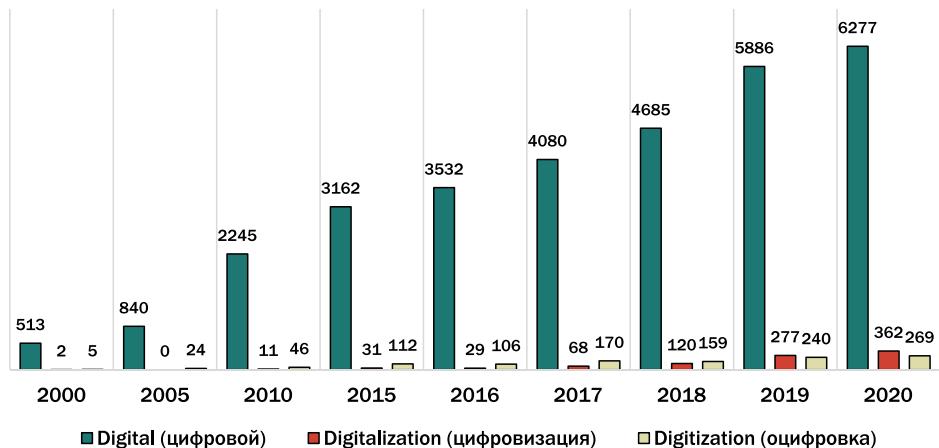


Рис. 1. Количество статей (социальные науки) по ключевым словам в базе данных Scopus, 2000—2020 гг.²

В начале статьи раскрыты основные теоретические концепты цифрового общества, соотношения между ними и их роль в процессах цифровизации. Затем описана методика количественного исследования интернет-практик россиян. Представлены основные результаты эмпирического исследования, показывающие особенности цифровизации российского общества. Рассмотрены перспективы дальнейшей цифровизации в России и ее ожидаемые последствия. В заключении обобщены результаты исследования и обозначены некоторые нерешенные научные проблемы.

Теоретическая модель цифрового общества

В научной литературе еще нет устоявшегося определения цифрового общества. В данном исследовании мы будем понимать под цифровым обществом современную стадию развития информационного общества, в которой важнейшее значение имеет не информация в целом, а прежде всего ее цифровой формат, методы оцифровки, кодирования и передачи информации [Khazieva, Khaziev, Klyushina, 2018: 349]. Ключевая роль в цифровом обществе отводится интернет-среде, развитие которой служит источником множества общественных изменений.

² Scopus. URL: <https://www.scopus.com/search/form.uri> (дата обращения: 10.01.2021).

Концепции цифрового общества находят поддержку как в исследовательских программах, так и в практике управления наиболее развитых стран. Так, в 2016 г. Ассоциация университетов Нидерландов объединила ученых из 14 университетов для проработки вопросов, возникших в связи с появлением цифрового общества в рамках исследовательской инициативы «Цифровое общество»³. Похожие инициативы реализуются в Цюрихском университете⁴, Университете Брауна⁵, Университете штата Аризона⁶ и в других организациях. Исследовательская программа «Развивающееся цифровое общество» германского Института интернета и общества Александра фон Гумбольдта⁷ внесла значительный вклад в концептуализацию цифрового общества как области исследования общественных наук. Правительство Японии в 2016 г. приняло стратегию «Общество 5.0», в рамках которой основанное на цифровых технологиях общество рассматривается как следующий этап развития информационно-ориентированного общества [Deguchi, 2020]. Основная цель стратегии состоит в улучшении качества жизни человека путем максимальной интеграции цифровой среды с физическим пространством.

Для описания ключевых характеристик цифрового общества и процесса его формирования (цифровизации) используем теоретическую модель, включающую четыре концепта, уже разработанных в научной литературе.

Первый концепт — *сверхсвязность* (от англ. superconnected) — характеризует высочайшую степень проникновения цифровых технологий в повседневную жизнь человека [Chayko, 2018]. Технологии увеличивают количество социальных взаимодействий до невозможного ранее уровня. Возникают новые формы взаимодействий с физическими предметами (например, интернет вещей) и искусственным интеллектом [Резаев, Трегубова, 2019; Coombs et al., 2020]. С помощью интернет-ресурсов люди участвуют во многих повседневных делах, например, следят за новостями; спорят о политике, спорте, музыке и кино; поддерживают социальные связи с друзьями и семьей; совершают покупки; знакомятся и ищут работу [Golder, Macy, 2014: 144].

Сверхсвязность может снижать негативное влияние расстояний и национальных границ на жизнь человека, экономить время, повышать культурный и образовательный уровень. Однако кроме очевидных преимуществ сверхсвязность порождает и проблемы, такие как новые формы зависимостей, распространение ложной информации, возможности слежки и др. Для достижения сверхсвязности требуется не только материальная база, обеспечивающая первичный доступ к информационным технологиям, но и широкое распространение навыков их использования [Dobrinskaya, Martynenko, 2019]. Поэтому в цифровом обществе возрастает значимость развития человеческого капитала, переосмысливаются механизмы социальной мобильности под влиянием цифровых технологий [Erokhina, Mukhametov, Sheremetiev 2019].

³ Digital Society. URL: <https://www.thedigitalsociety.info> (дата обращения: 25.01.2021).

⁴ Digital Society Initiative. URL: <https://www.dsi.uzh.ch/en.html> (дата обращения: 25.01.2021).

⁵ The Digital Society. URL: <https://www.brown.edu/research/projects/digital-society> (дата обращения: 25.01.2021).

⁶ Digital Society Initiative. URL: <https://research.wpcarey.asu.edu/digital-society> (дата обращения: 25.01.2021).

⁷ The Evolving Digital Society: Concepts, Discourses, Materialities. URL: <https://www.hiig.de/en/research/evolving-digital-society> (дата обращения: 25.01.2021).

Второй концепт — **платформизация** (*platformisation*) — отражает процесс проникновения цифровых платформ в различные сферы жизни. Под цифровыми платформами понимаются (пере)программируемые цифровые инфраструктуры, которые облегчают и формируют взаимодействия между пользователями и поставщиками услуг посредством систематического сбора, алгоритмической обработки, монетизации и распространения данных [Poell, Nieborg, van Dijck, 2019: 3]. Платформы позволяют «переносить» в цифровую среду самые разные формы социальных взаимодействий: общение (социальные сети и мессенджеры), покупки (виртуальные доски объявлений и электронные торговые площадки), поиск работы (онлайн-рекрутмент) и др. «Онлайн-мир не идентичен офлайновому, но он вполне реален. Пользователи, которые стремятся к статусу, восхищению окружающими, социальному одобрению и вниманию в своих онлайновых отношениях, приносят эти желания с собой в онлайн-сети» [Golder, Macy, 2014: 143].

Цифровые технологии дают больше преимуществ крупным транснациональным корпорациям [Ганичев, Кошовец, 2019], способным создавать платформы, функционирующие по всему миру, и объединять их в цифровые экосистемы путем многочисленных расширений и поглощений смежных платформ. Благодаря высокой степени взаимной интеграции платформ экосистемы становятся более привлекательными для потребителей. Так, почти во всех странах Запада доминируют многочисленные цифровые платформы «Большой пятерки» крупнейших американских компаний: Google-Alphabet, Amazon, Facebook, Apple и Microsoft [van Dijck, 2019: 2].

Третий концепт — **датификация** (*datafication*) — процесс количественной оценки и монетизации человеческой жизни с помощью цифровой информации [Mejias, Couldry, 2019: 2]. При использовании цифровых технологий генерируется цифровой след социальных взаимодействий [Billari, Zagheni, 2017], что «представляет беспрецедентные возможности для сбора как экспериментальных, так и наблюдательных данных в масштабах, одновременно огромных и микроскопических» [Golder, Macy, 2014: 131]. В результате формируются «большие данные», к свойствам которых Р. Китчин относит: огромный объем, высокую скорость производства и детализацию, внутреннее многообразие, исчерпывающую полноту, взаимосвязь с другими массивами данных, расширяемость и масштабируемость [Kitchin, 2014: 1—2]. Д. Лейзер и Дж. Рэдфорд выделяют три типа больших данных по источникам и характеру их получения: цифровая жизнь (фиксация социального поведения, опосредованного цифровыми технологиями), цифровые следы (записи о предпринятых действиях, а не сами действия) и оцифрованная жизнь (переход «аналогового» поведения в цифровую форму) [Lazer, Radford, 2017: 21—22].

Большие данные включают в себя не только новые виды цифровых данных, но и инструменты их получения, обработки и анализа. Они открывают столь обширные возможности производства знаний, что была разработана концепция четвертой парадигмы науки, связанной с интенсивным анализом большого объема данных и появлением принципиально новых форм научного метода [Kitchin, 2014]. Суммарное количество записей открытых микроданных о населении, доступных ученым, уже измеряется миллиардами [Ruggles, 2014: 287]. Еще больше данных доступно государству и крупным корпорациям. Как и в случае с платформизацией,

крупнейшие корпорации являются основными бенефициарами датификации, в результате чего возникают проблемы неравного доступа к информации (информационная асимметрия) и невозможности индивида контролировать процесс производства данных о себе.

Четвертый концепт — алгоритмическое управление (*algorithmic governance*) — выдвигает на первый план идею о том, что цифровые технологии особым образом конструируют социальную среду. Алгоритмическое управление — форма социального упорядочения, которая опирается на координацию между участниками, основана на правилах и включает в себя особенно сложные компьютерные эпистемологические процедуры [Katzenbach, Ulbricht, 2019: 2]. Алгоритмическое управление происходит не всегда преднамеренно и целенаправленно, но ему свойственны децентрализация и стабильность, позволяющие людям формировать свои ожидания о результатах работы алгоритмов управления и скоординированно действовать в соответствии с этими ожиданиями.

Системы алгоритмического управления существенно различаются по степени автоматизации и прозрачности. Они имеют разнообразные проявления: от модерации контента в социальных сетях и предоставления государственных услуг в цифровой форме до комплексных систем управления обществом, таких как система социального кредита в Китае [Liang et al., 2018], которая с помощью больших данных, собираемых из множества информационных систем, оценивает социальный рейтинг гражданина и может применять к нему поощрения или санкции: запрет на работу в госучреждениях, отказ в авиабилетах, запрет на обучение детей в частных школах и др. Эта система подвергается критике как «цифровая диктатура» и «цифровой концлагерь». Похожим образом функционируют системы, разработанные для контроля распространения коронавируса COVID-19 в некоторых странах и регионах. Более глубокое внедрение алгоритмического управления неизбежно вызовет множество этических проблем⁸ и недоверие со стороны гражданского общества [Falk, Römmel, Silverman, 2017].

Явления, описываемые представленными концептами, происходят одновременно, поддерживают и усиливают друг друга. Сверхсвязность создает основу для существования цифровых платформ и экосистем, которые генерируют большие данные, позволяющие осуществлять алгоритмическое управление социальными системами. Однако в разных странах и регионах явления происходят с различной интенсивностью. Цифровизация практик индивидов на микроуровне — базовое условие для осуществления других процессов формирования цифрового общества. Платформизация нацелена на перенос в цифровую среду социальных и экономических взаимодействий индивидов, организаций, органов власти; на их систематизацию и осмысление. Датификация открывает новые возможности для накопления и систематизации информации об обществе, которая может быть использована в разных целях. Наконец, алгоритмическое управление способно конструировать социальный порядок, устанавливая или контролируя институциональные основы социальной жизни.

⁸ Так, в футурологическом труде «Сумма технологии» С. Лем рассматривал проблему делегирования управления социальными системами алгоритмам, правильность решений которых ввиду их масштаба и сложности человек не сможет оценить и, следовательно, внести корректизы.

Процессы цифровизации так или иначе проявляются в практиках взаимодействия людей с интернет-средой (в научной литературе применяются такие термины, как «интернет-практики», «цифровые практики» и др.). Поэтому изучение этих практик и вопросов цифрового неравенства позволит улучшить понимание процессов цифровизации общества и адаптации населения к изменяющимся социальным условиям.

Социологическое осмысление цифрового неравенства началось с изучения цифрового разрыва в физической или материальной возможности доступа к новым цифровым технологиям. Современные исследования цифрового неравенства включают дополнительные уровни анализа: неравенство в цифровых навыках и мотивации, различие в практиках использования цифровых ресурсов и их влияние на жизненные шансы [Dimaggio, Hargittai, 2001; Dimaggio et al., 2004; Van Dijk, 2012]. Исследования выявили влияние возраста и уровня образования на неравенство в цифровых навыках [van Deursen, van Dijk, 2011] и продемонстрировали механизмы его формирования в процессе обучения [Rafalow, 2014]. Поскольку цифровое неравенство может усиливать степень социального неравенства и ограничивать участие части населения в жизни общества, разрабатывается концепция «цифрового исключения» [Van Deursen et al., 2017].

В России уже проведен ряд исследований интернет-практик и цифрового неравенства. Выявлена структура медиапредпочтений населения [Реутов, Тришина, 2015] и определены основные факторы цифрового неравенства [Волченко, 2016]. Оценено влияние поселенческого аспекта на цифровой разрыв [Шабунова, Груздева, Калачикова, 2020]. Показана динамика различных цифровых практик россиян, построена типология пользователей по практикам использования интернета и проанализировано проникновение интернета в разных социально-экономических и демографических группах [Богданов, Лебедев, 2017]. Рассмотрены три уровня цифрового разрыва: материальная база, навыки использования цифровых технологий и жизненные шансы [Dobrinskaya, Martynenko, 2019]. Проанализировано поведение социальных групп с позиций адаптации к цифровым технологиям, описаны стимулы повышения информационно-цифровой активности представителей разных типов социальной адаптации [Шиняева, Полетаева, Слепова, 2019].

В данной работе благодаря анализу лонгитюдных данных акцент сделан на выявлении условий, благоприятствующих возникновению практик и на оценке устойчивости уже возникших практик в контексте предложенной теоретической модели.

Методика исследования

В связи со сложным техно-социальным характером изучаемых явлений принято решение использовать источники двух видов. Во-первых, это данные, характеризующие реальное население России, а именно результаты выборочных обследований населения. Во-вторых, данные, отражающие характеристики «виртуального населения» цифрового пространства: метрики популярности интернет-ресурсов «Яндекс.Радар»⁹ и данные проекта «Виртуальное население

⁹ Яндекс.Радар. URL: <https://radar.yandex.ru/> (дата обращения: 25.01.2021).

России»¹⁰. Проект содержит географически привязанные демографические данные, полученные путем обработки анкет пользователей социальной сети «ВКонтакте» (около 200 млн анкет, собранных в январе — марте 2015 г., в 88 млн из которых указано место жительства или последнее место учебы на территории России). Данные проекта позволили оценить степень пространственной неравномерности проникновения цифровых технологий не только на региональном, но и на муниципальном уровне.

Использование социальных сетей и других цифровых данных дополняет, но не заменяет традиционные методы опроса. Виртуальное население не соответствует реальному, а представляет собой его проекцию, искаженную цифровой средой. Новые данные интернет-платформ ставят исследователей перед множеством серьезных препятствий, среди которых фрагментарность демографической информации, недостаточная презентативность и обобщаемость, уязвимость к изменениям в процессе генерации данных и атакам, ошибки в работе платформ, распространение фальшивых и спам-аккаунтов, возможность создания нескольких учетных записей и ввода недостоверных данных, манипуляции пользователей данными для достижения тайных целей, ограниченный или платный доступ к данным, контроль и воздействие на платформы со стороны государств и корпораций, этические ограничения [Golder, Macy, 2014; Lazer, Radford, 2017].

Если в начале века Россия по уровню интернет-проникновения соответствовала развивающимся странам (2,9% населения использовали интернет в 2001 г. при среднемировом значении 8,1%), то к 2019 г. она почти достигла уровня развитых стран (82,6%, при среднемировом значении 51,4%)¹¹. Причем более 60% населения России пользуются интернетом каждый день или почти каждый день. Чаще всего россияне используют интернет для посещения социальных сетей, среди которых первое место по популярности занимает «ВКонтакте»¹², что делает эту соцсеть одним из самых обширных и презентативных цифровых источников информации о российском населении. Тем не менее следует учитывать и ограничения данных, полученных таким образом. Например, жители пригородов склонны указывать в профилях в качестве места жительства крупные города, центры систем расселения [Замятиной, Яшунский, 2018], что в некоторой степени искажает результаты пространственного анализа.

Особое внимание в исследовании уделялось выявлению условий возникновения и исчезновения у индивидов интернет-практик путем анализа панельных данных лонгитюдного обследования РМЭЗ НИУ ВШЭ¹³. Гипотеза состоит в том, что темпы формирования в России цифрового общества существенно ограничивает низкая устойчивость практик (частота сохранения уже возникших практик). Для

¹⁰ Проект «Виртуальное население России» реализован при поддержке Русского географического общества под руководством Н. Ю. Замятиной и А. Д. Яшунского. URL: <http://webcensus.ru/> (дата обращения: 01.06.2020).

¹¹ По данным Международного союза электросвязи. URL: <https://www.itu.int/en/ITU-D/Statistics/Pages/stat/>; <https://data.worldbank.org/indicator/IT.NET.USER.ZS?locations=RU-1W> (дата обращения: 25.01.2021).

¹² По данным сервиса «Яндекс.Радар» на декабрь 2019 г. URL: <https://radar.yandex.ru/search> (дата обращения: 25.01.2021).

¹³ Российский мониторинг экономического положения и здоровья населения НИУ ВШЭ (RLMS HSE), проводимый НИУ ВШЭ и ООО «Демоскоп» при участии Центра народонаселения Университета Северной Каролины в Чапел Хилле и Института социологии ФНИСЦ РАН. URL: <http://hse.ru/rims> (дата обращения: 25.01.2021).

лонгитюдного анализа интернет-практик применялась следующая последовательность действий. Сначала репрезентативный массив обследования за год t (с 2004 по 2018 г.) дополнялся данными из полной выборки за предыдущий год ($t-1$) для респондентов из исходного массива. Затем полученные массивы были объединены и из них исключены респонденты моложе 14 лет, поскольку детский вопросник не содержит вопросов об использовании интернета.

В соответствии с вопросником обследования анализировалось наличие или отсутствие у респондентов практик использования интернета в следующих девяти целях (функциях): (1) для получения новостей, (2) общения, (3) получения справочных материалов, (4) повышения культурного уровня, (5) развлечений, (6) посещения социальных сетей, (7) покупок товаров и услуг, (8) работы, (9) учебы. На основе полученного массива панельных данных выявлялись пять видов событий, связанных с возникновением и исчезновением интернет-практик у респондентов (см. табл. 1). Всего в результате анализа 228 879 анкет 39 496 уникальных респондентов выявлено более 1,16 млн событий. С помощью полученного массива данных определены частоты возникновения и исчезновения практик для разных социальных групп и временных периодов с помощью формул 1 и 2:

$$\text{Частота возникновения практик} = \frac{e_2 + e_3}{e_1 + e_2 + e_3}; \quad (1)$$

$$\text{Частота исчезновения практик} = \frac{e_5}{e_4 + e_5}, \quad (2)$$

где e_i — количество выявленных событий вида i (обозначения групп событий см. в табл. 1).

Таблица 1. Группы и количество выявленных событий

Вид события	Группа событий	Условия			Выявлено событий, единиц	
		Наличие практик		Респондент пользовался интернетом в году $t-1$		
		В году $t-1$	В году t			
Практика не возникла	e_1	Нет	Нет	—	810 690	
Практика возникла (первичная)	e_2	Нет	Да	Нет	31 409	
Практика возникла (вторичная)	e_3	Нет	Да	Да	44 713	
Практика сохранилась	e_4	Да	Да	—	219 754	
Практика исчезла	e_5	Да	Нет	—	49 304	

Расчеты производились с помощью алгоритма, реализованного автором на языке программирования Julia 1.5, с использованием пакета обработки табличных данных `DataFrames.jl`. Фоновые картограммы для визуализации пространственных данных созданы с применением графического пакета `VegaLite.jl`.

Для выявления основных тенденций цифровизации российского общества анализировалось содержание паспортов национальных и федеральных проектов, утвержденных в России в 2018 г.¹⁴ Изучались следующие разделы паспортов: цели, индикаторы, ожидаемые результаты, мероприятия. Учитывая большие объемы финансирования, можно ожидать по крайней мере частичной реализации мероприятий национальных проектов. Поэтому векторы развития цифрового общества, сформулированные в этих документах, заслуживают детального рассмотрения.

Цифровизация российского общества: фрагментарная сверхсвязность

Межстрановое сравнение показывает (см. рис. 2), что, хотя Россия почти не уступает развитым странам по интегральному показателю использования интернета населением, она отстает от них по большинству частных показателей. Исключение составляют социальные сети и звонки, широко распространенные в России. Особенно силен разрыв в использовании интернета для дистанционного обучения, поиска работы, чтения газет, журналов или книг и коммуникации посредством электронной почты. Последнее, вероятно, объясняется более поздней цифровизацией в России — электронная почта уже перестала быть наиболее популярной формой коммуникации в сети.

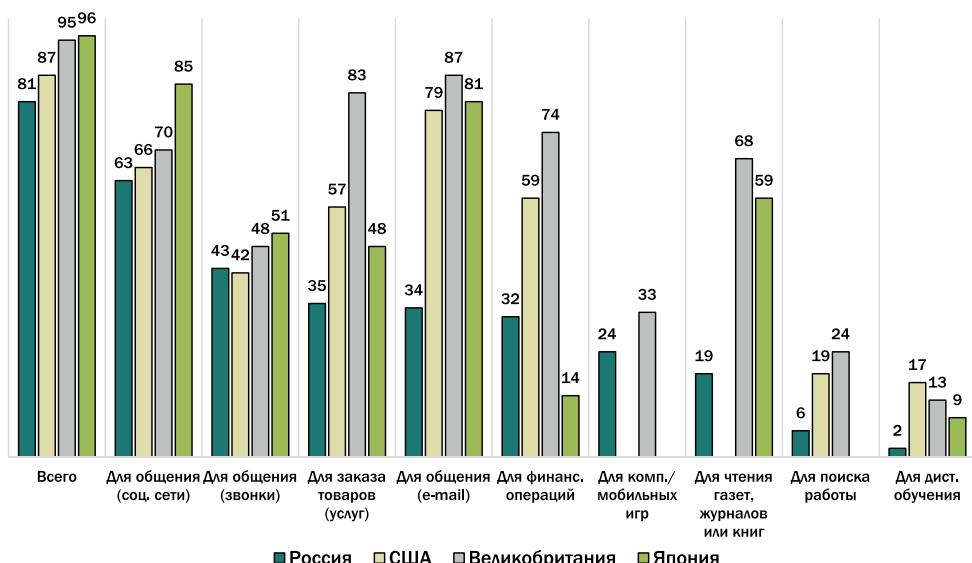


Рис. 2. Использование интернета населением по странам и целям, 2018 г., %¹⁵

¹⁴ Цифровая экономика 2024. Материалы. URL: <https://digital.ac.gov.ru/materials/passport/> (дата обращения: 25.01.2021).

¹⁵ Или за ближайшие годы, по которым имеются данные. В процентах от общей численности населения в возрасте от 15 до 74 лет. Россия — расчеты ИСИЭЗ НИУ ВШЭ по данным Росстата; Великобритания — Евростат; другие страны — ОЭСР, МСЭ. Составлено по данным краткого статистического сборника «Цифровая экономика: 2020», НИУ ВШЭ. URL: <https://issek.hse.ru/digec2020> (дата обращения: 25.01.2021).

Проследим динамику использования интернета по целям. За 15 лет доля использующих интернет выросла в 7 раз, а для общения — в 14 раз (см. рис. 3). Стремительно растут показатели использования интернета для покупок (в 27 раз), повышения культурного уровня, получения новостей (по 11) и развлечений (10). Однако удельный вес населения, использующего интернет для работы, и в особенности для обучения, растет довольно медленно (в 6 и 3 раза соответственно). Можно сделать вывод, что не во всех сферах общественной жизни достаточно быстро осваиваются практики использования интернет-технологий. Несмотря на развитие в последние годы онлайн-образования, эта сфера по-прежнему привлекательна для относительно небольшой доли населения. Однако по итогам 2020 г. следует ожидать значительного повышения этих показателей вследствие реализации карантинных мероприятий.

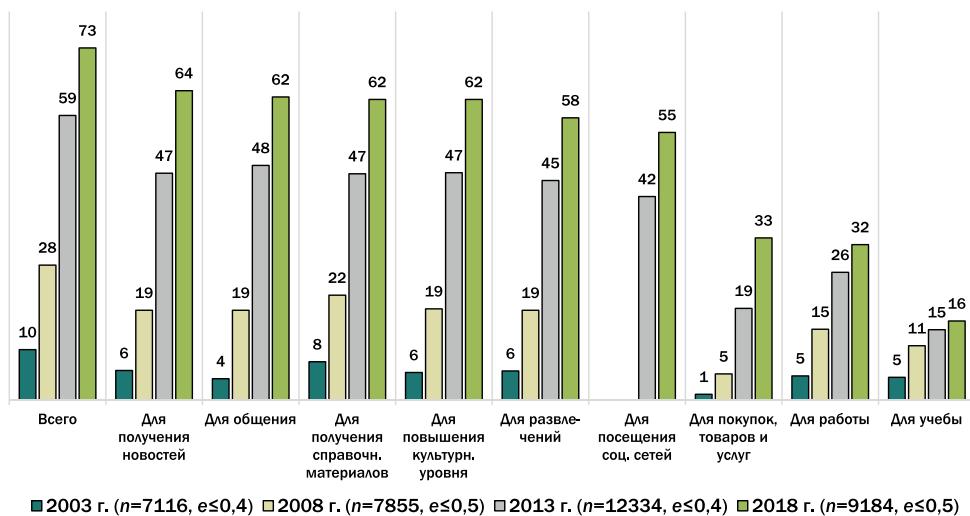


Рис. 3. Использование интернета населением России по целям, 2003—2018 гг., %¹⁶

График изменения частот возникновения и исчезновения интернет-практик в России (см. рис. 4), рассчитанный в соответствии с приведенной выше методикой, отражает сразу несколько важных закономерностей. Во-первых, на ранних стадиях цифровизации российского общества частота исчезновения практик многократно превосходила частоту возникновения. Практики были крайне неустойчивыми и быстро исчезали. Тем не менее показатели цифровизации росли, поскольку имелся высокий потенциал для усвоения практик населением. Во-вторых, частота возникновения практик на протяжении всего анализируемого периода увеличивалась, а исчезновения — уменьшалась. По-видимому, эти тенденции

¹⁶ В процентах от численности респондентов в возрасте от 15 до 74 лет. В скобках указаны объем выборки (n) и стандартные ошибки долей (e) в процентных пунктах. В 2003 и 2008 гг. вопросы об использовании интернета задавали только респондентам, пользующимся персональным компьютером; затем, в связи с распространением мобильного интернета,— всем респондентам. Составлено по данным обследования РМЭЗ НИУ ВШЭ.

сохраняются и в дальнейшем. В-третьих, поскольку к 2018 г. оба показателя почти сравнялись, можно ожидать дальнейшего распространения интернет-практик не только в краткосрочной, но и в долгосрочной перспективе.

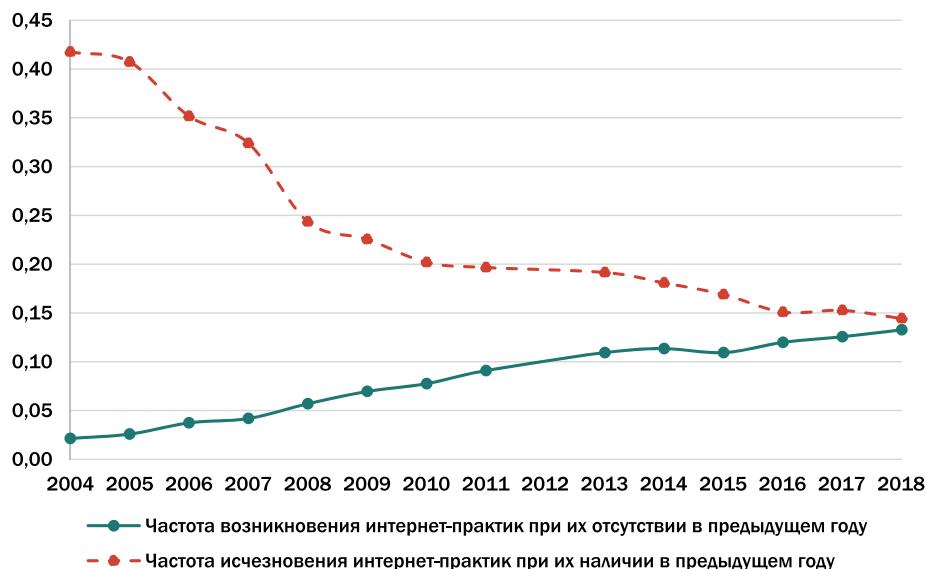


Рис. 4. Частоты возникновения и исчезновения интернет-практик жителей России, 2004—2018 гг.¹⁷

Об углублении цифровизации общества можно судить по тому факту, что среднее число целей, в которых взрослый житель России использует интернет, увеличилось с 0,4 в 2003 г. до 4,4 в 2018 г., то есть выросло более чем на порядок. Пик прироста пришелся на 2006—2012 гг. Куда менее оптимистичная картина складывается при рассмотрении показателей в разрезе социальных групп.

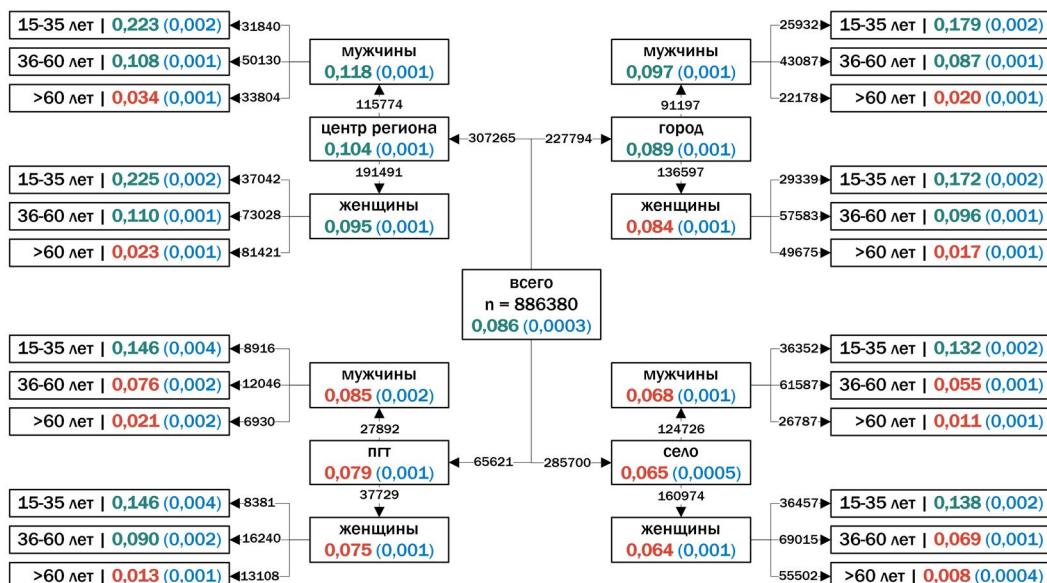
Полученные данные согласуются с другими эмпирическими исследованиями о проникновении интернета в России по различным социально-демографическим группам [Волченко, 2016; Богданов, Лебедев, 2017] и типам поселений [Шабунова и др., 2020]. Сильна зависимость от возраста респондентов. Пик освоения практик приходится на возраст от 19 до 22 лет, затем частота снижается, к 45 годам — вдвое, к 60 годам — вчетверо. Существенной разницы в показателях между полами нет, но мужчины несколько активнее осваивают практики в младших и старших возрастах, а женщины — в средних (30—60 лет). Частота освоения цифровых практик в поселках городского типа (пгт) на 21 % выше, чем в сельской местности, в городах — на 36 %, в столицах субъектов Российской Федерации — на 59 %, в Москве — на 85 %. К тому же в столичных городах практики более устойчивы. В сельской местности и в пгт ежегодно исчезает почти каждая четвертая практика. Работающие осваивают практики в 2,4 раза активнее, чем безработные, руководители и специалисты — вдвое активнее, чем рабочие. Существенное пре-

¹⁷ Составлено по данным обследования РМЭЗ НИУ ВШЭ. Данные на 2012 г. не приводятся в связи с изменением в этом году вопросника в части использования интернета населением.

имущество в освоении практик также дает наличие высшего образования (в два раза) и знание иностранного языка.

Чтобы продемонстрировать, что разные структуры цифрового неравенства существуют независимо и усиливают воздействие друг друга, рассмотрим древовидную схему на рисунке 5. Она классифицирует события возникновения или невозникновения практик по демографическим характеристикам людей: тип поселения, пол, возрастная группа. Подписи к стрелкам показывают количество рассмотренных событий в той или иной группе, числа в прямоугольниках — вероятность возникновения практики и стандартную ошибку доли (в скобках). Частоты ниже среднего по выборке уровня отмечены красным цветом. Такой подход позволяет независимо оценить влияние нескольких факторов на цифровое неравенство.

Различия между административными центрами регионов, городами, пгт и сельскими поселениями проявляются во всех шести половозрастных группах. Аналогичным образом и возраст дифференцирует показатели освоения практик во всех типах поселений и у респондентов обоих полов. Хотя по всей выборке мужчины осваивают практики лучше женщин, этот разрыв связан прежде всего с более низкой продолжительностью жизни мужчин, которая приводит к малому удельному весу мужчин старших возрастов. Внутри одних и тех же возрастных групп различия в освоении практик невелики. Разница между полярными значениями (молодые жители региональных центров и пожилые в сельской местности) более чем двадцатикратная. Если рассматривать частоты исчезновения/сохранения практик, то будут наблюдаться похожие структуры неравенства.



Обозначения: Группа событий | Частота возникновения интернет-практики (стандартная ошибка)

Рис. 5. Частоты возникновения интернет-практик по демографическим характеристикам¹⁸

¹⁸ Составлено по данным РМЭЗ НИУ ВШЭ за 2003—2018 гг.

Для более глубокого анализа далее ограничимся 10-летним возрастным промежутком (от 36 до 45 лет) и рассмотрим, как освоение интернет-практик в этой группе связано с социально-трудовыми характеристиками населения (см. рис. 6). Влияние уровня образования на частоту возникновения практик прослеживается во всех типах поселений. Среднее профессиональное образование дает преимущество в 1,3—1,5 раза над средним общим, а высшее образование — в 2,1—2,6 раза. У людей, владеющих иностранным языком, показатель еще в 1,2—2,0 раза выше. Влияние занятости и профессиональной группы на частоту возникновения практик также не зависит от возраста или типа поселения, в котором проживает респондент (исключение — разница между специалистами высшего уровня квалификации и руководителями в региональных столицах). На цифровое неравенство также могут влиять семейное положение, наличие детей, состояние здоровья, уровень дохода, миграционная история и многие другие факторы, но анализ этих структур неравенства выходит за пределы статьи.

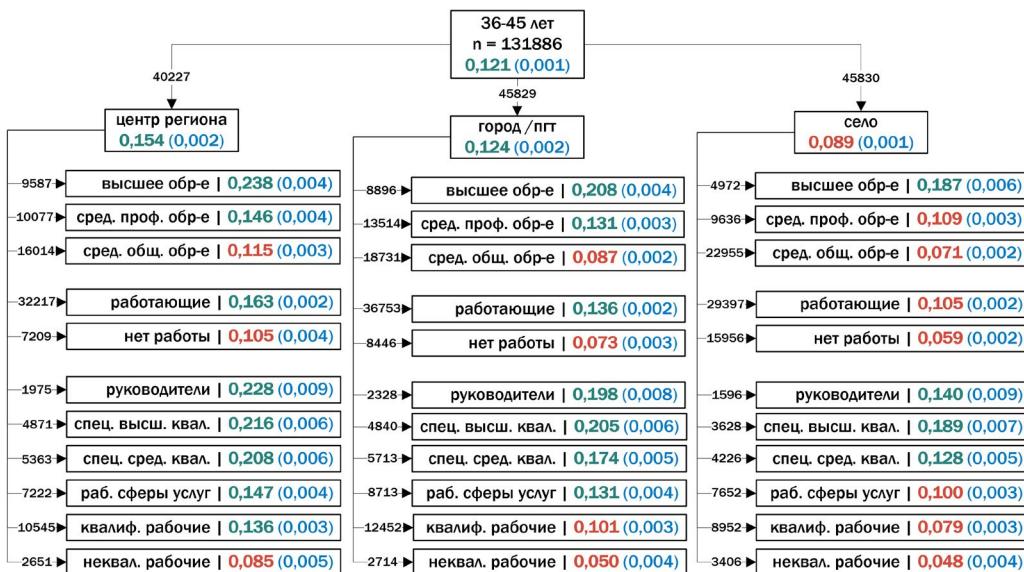


Рис. 6. Частоты возникновения интернет-практик у населения в возрасте 36—45 лет по социально-трудовым характеристикам. Составлено по данным РМЭЗ НИУ ВШЭ за 2003—2018 гг.

Начиная пользоваться интернетом, респонденты чаще обращаются к нему для общения, получения справочных материалов и развлечений. Для работы, учебы и покупки товаров интернет начинают использовать позже, с опытом взаимодействия с интернет-средой. Поскольку такие практики сильно влияют на жизненные шансы, их низкая устойчивость сдерживает преодоление цифрового неравенства. Исчезают чаще всего практики покупок товаров и услуг, реже всего — посещения социальных сетей и общения. Во всех рассмотренных случаях сильна дифференциация между социально-демографическими группами: особенно часто перестают использовать интернет респонденты старше 60 лет, а также живущие в пгт и сельской местности.

Показатель доли активных пользователей интернета по субъектам РФ в 2015 и 2019 гг. (см. рис. 7) демонстрирует, что наибольшая степень интернет-проникновения характерна для регионов Севера и Арктики, где население моложе и высока степень урбанизации. Отстающие регионы расположены в Центральном федеральном округе и на Юге Сибири, где выше доля населения старше трудоспособного возраста. Разница между регионами-лидерами и отстающими почти двукратная.

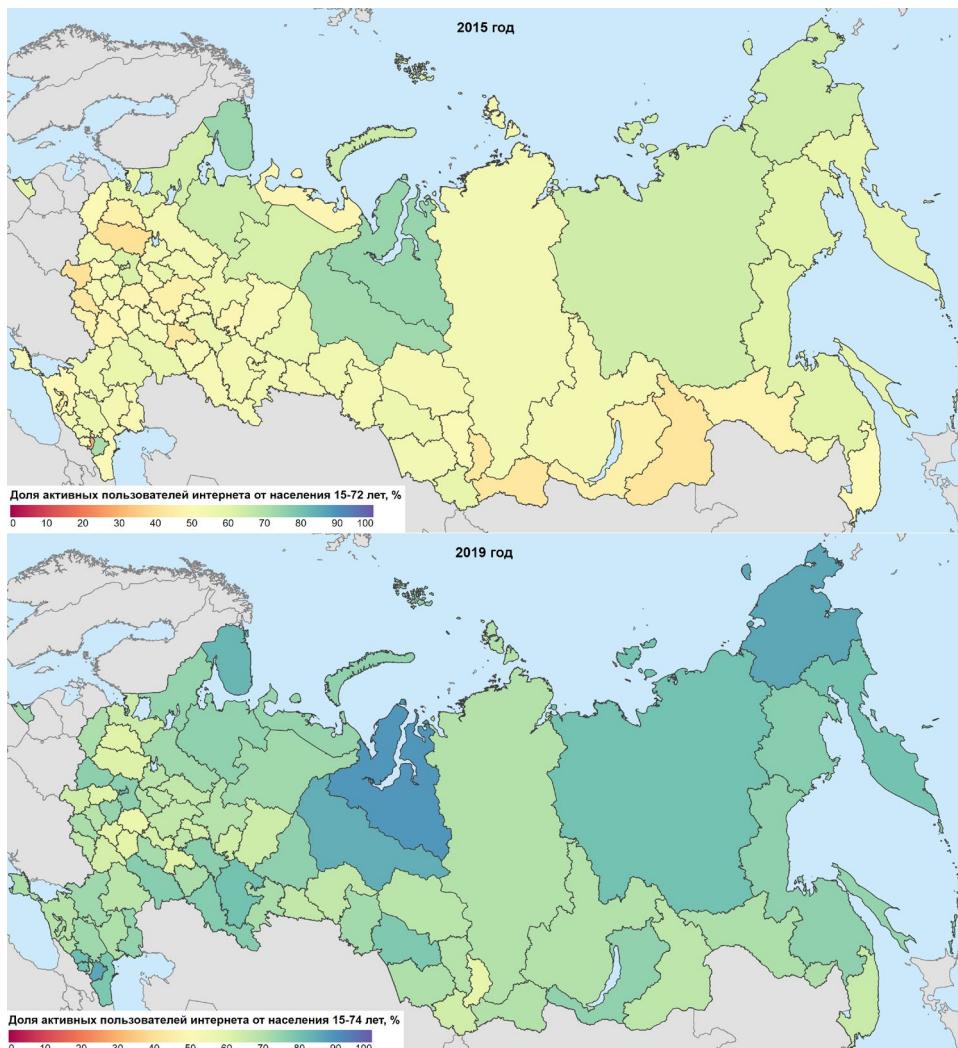


Рис. 7. Доля населения, использующего интернет каждый день или почти каждый день, по субъектам РФ в 2015 и 2019 гг.¹⁹

¹⁹ Составлено по данным выборочного обследования населения по вопросам использования информационно-коммуникационных технологий Росстата. URL: https://gks.ru/free_doc/new_site/business/it/fed_nabl-croc/index.html (дата обращения: 01.10.2020).

Среднее арифметическое значение показателя использования интернета по 85 регионам России увеличилось за четыре года с 53,8 % до 71,3 %, а дисперсия сократилась с 58,4 до 45,8. Иными словами, пространственное неравенство в использовании интернета в региональном разрезе сокращается. Для более полного понимания пространственных закономерностей цифровизации интерес представляют муниципальные данные, но официальная статистика на настоящий момент не дает подходящих показателей на муниципальном уровне. Поэтому воспользуемся данными проекта «Виртуальное население России», генерируемыми самими пользователями в цифровой среде (см. рис. 8).

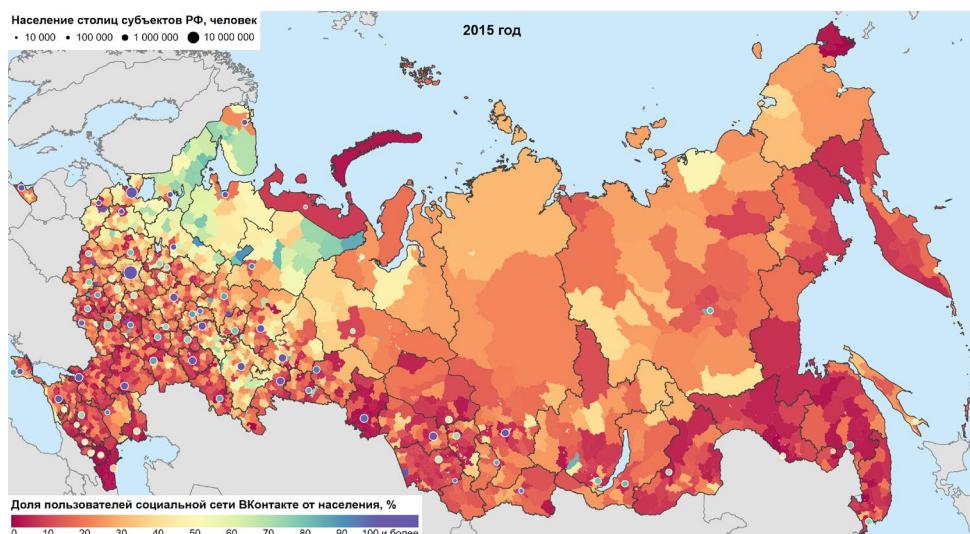


Рис. 8. Отношение числа пользователей социальной сети «ВКонтакте» к численности населения по муниципальным образованиям России на начало 2015 г., %

Карта демонстрирует, что в большинстве крупных городов и административных центров субъектов РФ доля пользователей социальной сети близка к общей численности населения. Однако муниципальные районы, занимающие большую часть территории страны, отстают от городских округов многократно. Вероятно, после 2015 г. пространственное неравенство на муниципальном уровне несколько снизилось, но по-прежнему велико.

Оценить степень неравномерности цифровизации на муниципальном уровне можно с помощью кривой Лоренца (см. рис. 9) и коэффициента Джини. Для построения графика 2361 муниципальное образование России отсортировано в порядке неубывания доли пользователей социальной сети «ВКонтакте» от населения. На рисунок последовательно наносились точки, отражающие долю показателей нарастающим итогом (кумулятивно). Расчеты показывают, что на 25 % населения, проживающего в городах и районах с наименьшей степенью проникновения социальной сети приходится лишь 5 % пользователей, а на 50 % населения России — 20 % пользователей. Коэффициент Джини, рассчитанный по графику, со-

ставил 0,409, что почти соответствует показателю неравенства доходов в России за тот же год (0,413).

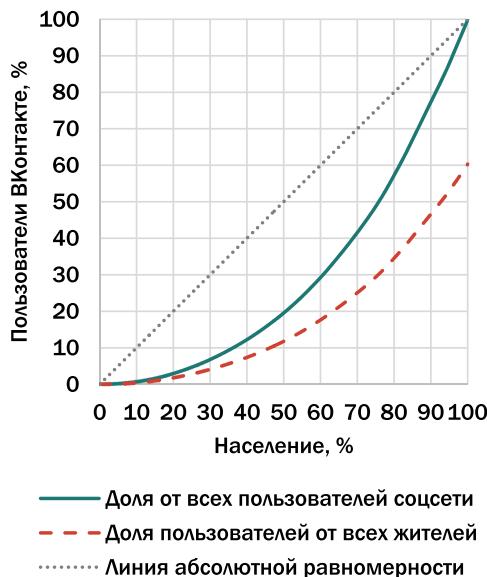


Рис. 9. Кривая неравномерности распределения по муниципальным образованиям России пользователей социальной сети «ВКонтакте» в 2015 г.

Столь низкое распространение интернет-практик в сельских и периферийных территориях во многом схоже с инфраструктурным неравенством [Лыткина, Смирнов, 2019]. Цифровое неравенство может привести к распространению среди россиян технопессимизма. Эмпирические исследования показывают, что часть жителей, в особенности старших возрастов, избирает пассивный тип адаптации к цифровизации [Шиняева, Полетаева, Слепова 2019]. Поэтому в российских условиях можно говорить лишь о фрагментарной сверхсвязности, характерной для молодого населения крупных городов. В рамках реализации национального проекта «Цифровая экономика» долю домохозяйств, имеющих возможность широкополосного доступа к интернету, планируется увеличить с 75 % в 2018 г. до 97 % в 2024 г., а долю населения, обладающего цифровой грамотностью и ключевыми компетенциями цифровой экономики, — с 26 % до 40 %. Акцент национальных проектов на создании инфраструктуры, а не на формировании навыков у всего населения, не позволяет ожидать быстрого перелома в сокращении цифрового неравенства.

Перспективы экспансии цифровых экосистем и барьера адаптации населения

Для изучения платформизации в России обратимся к данным сервиса «Яндекс. Радар» (см. табл. 2). Первые строчки во всех представленных рейтингах занимают интернет-ресурсы, являющиеся цифровыми платформами, то есть предо-

ставляющие инфраструктуру для размещения контента самими пользователями. Специфика России состоит в том, что, в отличие от большинства стран Европы, на многих ключевых рынках лидируют отечественные разработки (например, поисковая система «Яндекс» и социальная сеть «ВКонтакте»). Это говорит о высокой конкурентоспособности российской IT-сферы, хотя может быть барьером для расширения взаимодействий россиян с иностранцами в интернет-среде.

Таблица 2. Наиболее популярные в России интернет-ресурсы по категориям на конец 2019 г.*

Место	Категории			
	Поисковые системы	Социальные сети	Видео	Интернет-магазины
1	1115 — Яндекс	192 — ВКонтакте	129 — YouTube	33 — AliExpress
2	774 — Google	112 — Instagram	124 — Яндекс.Видео	31 — WildBerries
3	31 — Mail.ru	108 — Одноклассники	25 — ivi.ru	28 — OZON.ru
4	3 — Rambler	67 — Facebook	11 — Rutube	25 — М. Видео
5	22 — остальные	20 — Twitter	8 — THT-Online	23 — DNS

* Поисковые системы — млн запросов за неделю, остальные категории — млн посетителей за месяц. Составлено по данным «Яндекс.Радар» на декабрь 2019 г.

В разных категориях рейтинга встречаются платформы одних и тех же компаний, что свидетельствует о формировании технологическими гигантами цифровых экосистем путем поглощений и экспансии на новые рынки²⁰. Горизонтальное объединение платформ в рамках единой инфраструктуры позволяет компаниям получать больше данных о пользователях, привлекать их на другие свои платформы. Хотя у россиян и имеется некоторый простор для выбора той или иной экосистемы, существуют барьеры мобильности. Переход на другую экосистему сопряжен по крайней мере с потерей времени, сложностью переноса данных между платформами, возможными трудностями в решении повседневных задач без привычных инструментов.

Помимо частных цифровых экосистем, в России активно расширяется и государственная — на базе Единого портала государственных услуг Российской Федерации. Одно из проявлений экспансии этой экосистемы состоит в предоставлении скидок на государственные пошлины или при уплате штрафов гражданам, получающим услуги через портал. Такая практика направлена на стимулирование цифровизации, но в краткосрочной перспективе она может усиливать социальное неравенство, поскольку привилегии получают наиболее приспособленные группы населения, владеющие цифровыми навыками.

²⁰ Яркий пример — компания «Яндекс», поглотившая многие крупные платформы в смежных сферах («Яндекс.Карты», «КиноПоиск», Auto.ru), вышедшая на рынки такси («Яндекс.Такси»), доставки еды («Яндекс.Еда»), финансовых и др. Аналогично в Mail.ru Group входят социальные сети «ВКонтакте» и «Одноклассники», платформа доставки еды Delivery Club (совместно со Сбербанком), картографический сервис MAPS.ME и многие другие крупные интернет-платформы. Экосистема «Сбер» (Сбербанка) помимо финансовых услуг включает платформы «СберМаркет», «СберАвто», «СберКлауд», «СберФуд» и др.

Датификация — один из приоритетов российской государственной политики. Это зафиксировано в целевых показателях национальных и федеральных проектов. Уже в 2022 г. планируется подключить к Единой государственной информационной системе в области здравоохранения все медицинские организации, а в 2024 г. — внедрить во всех регионах целевую модель цифровой образовательной среды. Доля взаимодействий граждан и организаций с государственными органами, осуществляемых в цифровом виде, планируется увеличить с 15 % в 2017 г. до 70 % в 2024 г. При этом 100 % приоритетных государственных услуг будут предоставлять без необходимости личного посещения. В единую облачную платформу планируется перенести 70 % информационных систем и ресурсов федеральных органов власти. Будущие переписи населения планируется проводить с широким применением цифровых технологий и больших данных²¹.

Очевидно, что даже при частичной реализации намеченных целей или при переносе мероприятий на более поздний срок объем накапливаемых государством данных о российском обществе возрастет многократно. Эти данные могут найти применение в научных исследованиях и совершенствовании государственного управления, но одновременно с этим возрастают риски утечек персональных данных и других злоупотреблений накопленной информацией. Дополнительным фактором ускорения цифровизации уже в глобальном масштабе становится пандемия коронавирусной инфекции COVID-19, форсировавшая переход в интернет-среду образования, работы, покупок и других сфер жизни. В условиях подобных кризисов цифровое неравенство представляет особенно большую опасность [Beaunoyer, Dupré, Guitton, 2020], поскольку пропадает возможность получения услуг в традиционных формах.

Алгоритмическое управление в России сегодня применяется ограниченно. Самая массовая реализация — Портал государственных услуг, в котором на конец 2019 г. было зарегистрировано свыше 100 млн человек. Он автоматизирует регламенты предоставления государственных услуг, обычно не требующие принятия сложных решений. Наиболее комплексный и сложный проект в сфере алгоритмического управления — «Умный город» — включен в нацпроект «Жилье и городская среда». В него входят мероприятия по внедрению цифровых технологий в управление городской средой, энергетическими системами, водными ресурсами, транспортом, общественной безопасностью [Попов, Семячков, 2020]. Планируется применение больших данных и систем искусственного интеллекта, создание «цифровых двойников» городов, широкое вовлечение граждан в решение вопросов городского развития посредством цифровой платформы. Мероприятия проекта будут реализованы преимущественно в крупных городах. Многие из наработок «Умного города» были использованы для контроля за соблюдением гражданами самоизоляции во время пандемии COVID-19.

Дальнейшая экспансия цифровых технологий во все сферы жизни неизбежна. Она будет осуществляться в рамках крупных частных и государственных цифровых экосистем. Цифровое неравенство выступает главным барьером ее реализации, поскольку миллионы людей, не владеющих материальной базой, навыками или

²¹ В России в 2020 г. пройдет первая цифровая перепись населения // ТАСС. 2019. 29 сентября. URL: <https://tass.ru/obschestvo/6942094> (дата обращения: 25.01.2021).

мотивацией для взаимодействия с интернет-средой, могут быть лишены возможности равного доступа к общественным благам. Другие проблемы — возможности утечек или нецелевого использования персональных данных — особенно опасны в цифровую эпоху.

Заключение

Концепты цифрового общества показали свою применимость как аналитический инструмент для изучения социальных процессов в современной России. Теоретическая модель продемонстрировала степень углубления и расширения процессов цифровизации, позволила оценить перспективы их дальнейшего развития. Вместе с тем она показала и фрагментарность формируемого сегодня в России цифрового общества, а также выявила риски для людей, которые не смогут адаптироваться к общественным изменениям.

Практическая значимость исследования состоит в возможности использования выявленных закономерностей при разработке стратегий и мероприятий, затрагивающих вопросы цифровизации экономики и общественной жизни. Для углубления понимания цифрового общества в будущих исследованиях следует обратить внимание на такие вопросы, как влияние цифровизации на жизненные стратегии, барьеры мобильности между цифровыми экосистемами, допустимые границы применения алгоритмического управления обществом, разработка методического инструментария для изучения цифрового общества.

Список литературы (References)

- Богданов М. Б., Лебедев Д. В. Пользование сетью интернет в России в 2003—2015 гг. // Вестник Российского мониторинга экономического положения и здоровья населения НИУ ВШЭ (RLMS-HSE) / отв. ред.: Козырева П. М.. Вып. 7. М.: НИУ ВШЭ, 2017. С. 129—145.
- Bogdanov M. B., Lebedev D. V. (2017). The Use of the Internet in Russia from 2003 to 2015. In: *The Bulletin of the RLMS-HSE. Issue 7*. Moscow: HSE. P. 129—145. (In Russ.)
- Волченко О. В. Динамика цифрового неравенства в России // Мониторинг общественного мнения: экономические и социальные перемены. 2016. № 5. С. 163—182. <https://doi.org/10.14515/monitoring.2016.5.10>.
- Volchenko O. V. (2016) Dynamics of Digital Inequality in Russia. *Monitoring of Public Opinion: Economic and Social Changes*. No. 5. P. 163—182. <https://doi.org/10.14515/monitoring.2016.5.10>. (In Russ.)
- Ганичев Н. А., Кошовец О. Б. Технологический прорыв на базе развития цифровой экономики: возможности, проблемы, риски // Проблемы прогнозирования. 2019. № 6. С. 48—59. URL: <https://ecfor.ru/publication/tehnologicheskij-proryv-na-osnove-razvitiya-tsifrovoj-ekonomiki/> (дата обращения: 25.01.2021).
- Ganichev N. A., Koshevets O. B. (2019) Integrating Russia into the Global Project of Digital Transformation: Opportunities, Problems and Risks. *Studies on Russian Economic Development*. No. 6. P. 627—636. <https://doi.org/10.1134/S1075700719060030>.

Замятин Н. Ю., Яшунский А. Д. Виртуальная география виртуального населения // Мониторинг общественного мнения: экономические и социальные перемены. 2018. № 1. С. 117—137. <https://doi.org/10.14515/monitoring.2018.1.07>.

Zamyatina N. Yu., Yashunsky A. D. (2018) Virtual Geography of Virtual Population. *Monitoring of Public Opinion: Economic and Social Changes*. No. 1. P. 117—137. <https://doi.org/10.14515/monitoring.2018.1.07>. (In Russ.)

Колозариди П. В., Макушева М. О. Интернет как проблемное поле социальных наук // Мониторинг общественного мнения: экономические и социальные перемены. 2018. № 1. С. 1—11. <https://doi.org/10.14515/monitoring.2018.1.01>.

Kolozaridi P.V., Makusheva M. O. (2018) The Internet as a Problematic Field of Study in Social Sciences. *Monitoring of Public Opinion: Economic and Social Changes*. No. 1. P. 1—11. <https://doi.org/10.14515/monitoring.2018.1.01>. (In Russ.)

Лыткина Т. С., Смирнов А. В. Российский Север в условиях глобальной неолиберальной политики: преодоление пространственного неравенства или вытеснение? // Мир России. 2019. Т. 28. № 3. С. 27—47. <https://doi.org/10.17323/1811-038X-2019-28-3-27-47>.

Lytkina T.S., Smirnov A.V. (2019) The Russian North in the Context of Global Neoliberal Politics: Overcoming Spatial Inequality or Expulsion. *Universe of Russia*. Vol. 28. No. 3. P. 27—47. <https://doi.org/10.17323/1811-038X-2019-28-3-27-47>. (In Russ.)

Попов Е. В., Семячков К. А. Умные города. М.: Юрайт, 2020.

Popov E. V., Semyachkov K. A. (2020). Smart Cities. Moscow: URAIT. (In Russ.)

Резаев А. В., Трегубова Н. Д. «Искусственный интеллект», «онлайн-культура», «искусственная социальность»: определение понятий // Мониторинг общественного мнения: экономические и социальные перемены. 2019. № 6. С. 35—47. <https://doi.org/10.14515/monitoring.2019.6.03>.

Rezaev A. V., Tregubova N. D. (2019) Artificial Intelligence, On-line Culture, Artificial Sociality: Definition of the Terms. *Monitoring of Public Opinion: Economic and Social Changes*. No. 6. P. 35—47. <https://doi.org/10.14515/monitoring.2019.6.03>. (In Russ.)

Реутов Е. В., Тришина Т. В. Интернет-практики и информационные предпочтения населения // Социологические исследования. 2015. № 4. С. 140—148. URL: <http://socis.isras.ru/article/5201> (дата обращения: 25.01.2021).

Reutov E. V., Trishina T. V. (2015) People's Internet-Practices and Information Preferences. *Sociological Studies*. No. 4. P. 140—148. URL: <http://socis.isras.ru/article/5201> (accessed: 25.01.2021). (In Russ.)

Шабунова А. А., Груздева М. А., Калачикова О. Н. Поселенческий аспект цифрового неравенства в современной России // Проблемы развития территории. 2020. № 4. <https://doi.org/10.15838/ptd.2020.4.108.1>.

Shabunova A. A., Gruzdeva M. A., Kalachikova O. N. (2020) Settlement Aspect of Digital Inequality in Modern Russia. *Problems of Territory's Development*. No. 4. P. 7—19. <https://doi.org/10.15838/ptd.2020.4.108.1>. (In Russ.)

Шиняева О.В., Полетаева О.В., Слепова О.М. Информационно-цифровое неравенство: поиски эффективных практик адаптации населения // Мониторинг общественного мнения: экономические и социальные перемены. 2019. № 4. С. 68—85. <https://doi.org/10.14515/monitoring.2019.4.04>.

Shinyaeva O.V., Poletaeva O.V., Slepova O.M. (2019) Information and Digital Inequality: Searching for Effective Population Adaptation Practices. *Monitoring of Public Opinion: Economic and Social Changes*. No. 4. P. 68—85. <https://doi.org/10.14515/monitoring.2019.4.04>. (In Russ.)

Beaunoyer E., Dupéré S., Guitton M.J. (2020) COVID-19 and Digital Inequalities: Reciprocal Impacts and Mitigation Strategies. *Computers in Human Behavior*. Vol. 111. Article 106424. <https://doi.org/10.1016/j.chb.2020.106424>.

Billari F., Zagheni E. (2017) Big Data and Population Processes: A Revolution? In: *SIS 2017. Statistics and Data Science: New Challenges, New Generations*. Florence: Firenze University Press. P. 167—178. <https://doi.org/10.36253/978-88-6453-521-0>.

Chayko M. (2018) Superconnected. London: Sage.

Coombs C., Hislop D., Taneva S.K., Barnard S. (2020) The Strategic Impacts of Intelligent Automation for Knowledge and Service Work: An Interdisciplinary Review. *The Journal of Strategic Information Systems*. Vol. 29. No. 4. Article 101600. <https://doi.org/10.1016/j.jsis.2020.101600>.

Deguchi A. et al. (2020) Society 5.0. A People-Centric Super-Smart Society. Singapore: Springer. <https://doi.org/10.1007/978-981-15-2989-4>.

Dimaggio P., Hargittai E. (2001) From the ‘Digital Divide’ to ‘Digital Inequality’: Studying Internet Use as Penetration Increases. *Working Paper #15*. Princeton University, Center for Arts and Cultural Policy Studies.

Dimaggio P., Hargittai E., Celeste C., Shafer S. (2004) Digital Inequality: From Unequal Access to Differentiated Use. In: Social Inequality. New York, NY: Sage. P. 355—400.

Dobrinskaya D. E., Martynenko T. S. (2019) Defining the Digital Divide in Russia: Key Features and Trends. *Monitoring of Public Opinion: Economic and Social Changes*. No. 5. P. 100—119. <https://doi.org/10.14515/monitoring.2019.5.06>.

Dufva T., Dufva M. (2019) Grasping the Future of the Digital Society. *Futures*. Vol. 107. P. 17—28. <https://doi.org/10.1016/j.futures.2018.11.001>.

Erokhina O. V., Mukhametov D. R., Sheremetiev A. V. (2019) New Social Reality: Digital Society and Smart City. In: *Wave Electronics and Its Application in Information and Telecommunication Systems (WECONF)*. P. 1—6. <https://doi.org/10.1109/weconf.2019.8840644>.

Falk S., Römmele A., Silverman M. (2017) The Promise of Digital Government. In: *Digital Government*. Switzerland: Springer. P. 3—23. <https://doi.org/10.1007/978-3-319-38795-6>.

- Grimshaw M. (2017) Digital Society and Capitalism. *Palgrave Communications*. Vol. 3. Article 28. <https://doi.org/10.1057/s41599-017-0020-5>.
- Golder S. A., Macy M. W. (2014) Digital Footprints: Opportunities and Challenges for Online Social Research. *Annual Review of Sociology*. Vol. 40. No. 1. P. 129—152. <https://doi.org/10.1146/annurev-soc-071913-043145>.
- Katzenbach C., Bächle T. C. (2019) Defining Concepts of the Digital Society. *Internet Policy Review*. Vol. 8. No. 4. <https://doi.org/10.14763/2019.4.1430>.
- Katzenbach C., Ulbricht L. (2019) Algorithmic Governance. *Internet Policy Review*. Vol. 8. No. 4. <https://doi.org/10.14763/2019.4.1424>.
- Khazieva N., Khaziev A., Klyushina E. (2018) Digital Society: The Experience of the Philosophical Understanding of a Problem. *Journal of History Culture and Art Research*. Vol. 7. No. 4. P. 347—353. <https://doi.org/10.7596/taksad.v7i4.1856>.
- Kitchin R. (2014) Big Data, New Epistemologies and Paradigm Shifts. *Big Data & Society*. P. 1—12. <https://doi.org/10.1177/2053951714528481>.
- Lazer D., Radford J. (2017) Data ex Machina: Introduction to Big Data. *Annual Review of Sociology*. Vol. 43. No. 1. P. 19—39. <https://doi.org/10.1146/annurev-soc-060116-053457>.
- Liang F., Das V., Kostyuk N., Hussain M. M. (2018) Constructing a Data-Driven Society: China's Social Credit System as a State Surveillance Infrastructure. *Policy & Internet*. Vol. 10. No. 4. P. 415—453. <https://doi.org/10.1002/poi3.183>.
- Lupton D. (2015) Digital Sociology. London, New York: Routledge.
- Matzner T., Ochs C. (2019) Privacy. *Internet Policy Review*. Vol. 8. No. 4. <https://doi.org/10.14763/2019.4.1427>.
- Meijas U. A., Couldry N. (2019) Datafication. *Internet Policy Review*. Vol. 8. No. 4. <https://doi.org/10.14763/2019.4.1428>.
- Poell T., Nieborg D., van Dijck J. (2019) Platformisation. *Internet Policy Review*. Vol. 8. No. 4. <https://doi.org/10.14763/2019.4.1425>.
- Rafalow M. H. (2014) The Digital Divide in Classroom Technology Use: A Comparison of Three School. *International Journal of Sociology of Education*. Vol. 3. No. 1. P. 67—100. <https://doi.org/10.4471/rise.2014.04>.
- Redshaw T. (2019) What is Digital Society? Reflections on the Aims and Purpose of Digital Sociology. *Sociology*. Vol. 54. No. 2. P. 425—531. <https://doi.org/10.1177/0038038519880114>.
- Ruggles S. (2014) Big Microdata for Population Research. *Demography*. Vol. 51. P. 287—297. <https://doi.org/10.1007/s13524-013-0240-2>.
- Selwyn N. (2019) What is Digital Sociology? Cambridge, UK: Polity Press.

van Deursen E. J., Helsper E., Eynon R., van Dijk J. A. G. M. (2017) The Compoundness and Sequentiality of Digital Inequality. *International Journal of Communication*. Vol. 11. P. 452—473.

van Deursen E. J., van Dijk J. (2011) Internet Skills and the Digital Divide. *New Media & Society*. Vol. 13. No. 6. P. 893—911. <https://doi.org/10.1177/1461444810386774>.

van Dijk J. A. G. M. (2012) The Evolution of the Digital Divide: The Digital Divide Turns to Inequality of Skills and Usage. In: Bus J., Crompton M., Hildebrandt M., Metakides G. (eds.) *Digital Enlightenment Yearbook 2012*. Amsterdam: IOS Press. P. 57—75. <https://doi.org/10.3233/978-1-61499-057-4-57>.

van Dijk J. A. G. M. (2013) Inequalities in the Network Society. In: Orton-Johnson K., Prior N. (eds.) *Digital Sociology*. London: Palgrave Macmillan. P. 105—124. <https://doi.org/10.1057/9781137297792>.

van Dijck J. (2019) Governing Digital Societies: Private Platforms, Public Values. *Computer Law & Security Review*. Vol. 36. No. 105377. <https://doi.org/10.1016/j.clsr.2019.105377>.

DOI: [10.14515/monitoring.2021.1.1784](https://doi.org/10.14515/monitoring.2021.1.1784)



А. П. Климович

РИСКИ КОНЦЕНТРАЦИИ ВЛАСТИ В ЭПОХУ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА: МЕЖДУ СЛЕДЯЩИМ КАПИТАЛИЗМОМ И ЦИФРОВЫМ ЛЕВИАФАНОМ

Правильная ссылка на статью:

Климович А. П. Риски концентрации власти в эпоху искусственного интеллекта: между следящим капитализмом и цифровым Левиафаном // Мониторинг общественного мнения: экономические и социальные перемены. 2021. № 1. С. 154—170. <https://doi.org/10.14515/monitoring.2021.1.1784>.

For citation:

Klimovich A. P. (2021) Risky Concentration of Power in the Age of AI: Between Surveillance Capitalism and Digital Leviathan. *Monitoring of Public Opinion: Economic and Social Changes*. No. 1. P. 154–170. <https://doi.org/10.14515/monitoring.2021.1.1784>. (In Russ.)

РИСКИ КОНЦЕНТРАЦИИ ВЛАСТИ В ЭПОХУ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА: МЕЖДУ СЛЕДЯЩИМ КАПИТАЛИЗМОМ И ЦИФРОВЫМ ЛЕВИАФАНОМ

КЛИМОВИЧ Александр Павлович — кандидат философских наук, научный руководитель Лаборатории по изучению больших данных, Финансовый университет при Правительстве Российской Федерации (Липецкий филиал), Липецк, Россия

E-MAIL: allexgut@gmail.com
<https://orcid.org/0000-0002-8912-7635>

Аннотация. В статье рассмотрен феномен роста влияния лидеров IT-индустрии, предпринята попытка анализа этого явления. Описаны технологические особенности платформенной архитектуры, способствующие концентрации власти в руках крупнейших концернов информационного рынка. Обрисована бизнес-модель большой пятерки GAFAM, истолкованная в качестве ключевого источника власти цифровых гигантов. Представлено описание качественного отличия цифрового капитализма начала XXI века от индустриального капитализма прошедшего столетия. Изложены подходы критики цифрового капитализма, опирающиеся на опыт критических теорий классиков социологии. Рассмотрены принципы выстраивания отношений государственных структур с IT-индустрией с точки зрения баланса и концентрации власти, освещены авторитарные и демократические варианты взаимодействия.

Ключевые слова: GAFAM, цифровой капитализм, платформенный капитализм, следящий капитализм, новые информационные технологии, большие

RISKY CONCENTRATION OF POWER IN THE AGE OF AI: BETWEEN SURVEILLANCE CAPITALISM AND DIGITAL LEVIATHAN

Alexander P. KLIMOVICH¹ — Cand. Sci. (Soc.), Scientific Adviser at the Big Data Lab

E-MAIL: allexgut@gmail.com
<https://orcid.org/0000-0002-8912-7635>

¹ Financial University under the Government of the Russian Federation (Lipetsk Branch), Lipetsk, Russia

Abstract. The paper is focused on a remarkable growth of the IT industry giants' influence. The study describes specific features of platform architecture allowing the largest information market companies to increase their market power concentration. The Big Five (GAFAM) business model is interpreted in terms of the key source of the giants' power. The author describes qualitative differences between the capitalism in the early 21st century and the 20th-century industrial capitalism. The author outlines approaches to digital capitalism criticism resting on the critical theories of the sociological classics. The author considers rules of the government - IT-industry relationship in terms of balance and concentration of power and touches upon authoritarian and democratic interaction patterns.

Keywords: GAFAM, digital capitalism, platform capitalism, surveillance capitalism, information technology, artificial intelligence, Big data, critical theories,

данные, критические теории, Франкфуртская школа, демократия, авторитаризм, социальный кредит

the Frankfurt school, democracy authoritarianism, social credit

Цифровые технологии стремительно набирают вес в экономике, государственном управлении, образовании, медицине, медийном пространстве, социальной и приватной жизни каждого человека. Одновременно растет их влияние на современное общество. Особенно четко этот рост прослеживается в сфере экономики, там, где количественные показатели капитализации непосредственно отражают долю влияния компаний на рынке. В статье приведены некоторые статистические данные, отражающие этот процесс. Ниже будет показано, что на сегодняшний день следует говорить не только о количественных изменениях, проходящих под влиянием IT-технологий, но и о возникновении глубоких, качественных перемен в обществе. Будет объяснена целесообразность введения в практику термина «эпоха цифрового капитализма» и описано отличие наступающей эры от эпохи индустриального капитализма. Основная проблема, с которой работал автор, заключается в поиске подходов к анализу рисков концентрации власти при формировании цифрового капитализма. В статье предпринята попытка привлечь к этому анализу критику капиталистических отношений, проделанную рядом социологов на протяжении XX века. Вторая проблема, рассмотренная автором, состоит в анализе взаимодействия государственного управления с лидерами информационного монополизированного сектора.

Принцип генерирования добавленной стоимости при цифровом капитализме

Для демонстрации степени воздействия цифровых технологий на повседневную жизнь современного человека приведем свежую статистику. Наиболее показательными выглядят данные в области интернет-поиска, социальных сетей и электронной коммерции. Обратим внимание, насколько распространился в мире интернет. Количество активных пользователей в 2020 г. составило 4,54 млрд человек, что приблизительно составляет 56 % жителей планеты. Таким образом, на сегодняшний день каждые шесть из десяти человек активно пользуются всемирной паутиной. На долю поисковой системы Google приходится 76 % всех поисковых запросов в мире. Каждую секунду самая популярная поисковая система отвечает на более чем 40 тыс. поисковых запросов. Продукты и услуги Google насчитывают более 1 млрд пользователей по всему миру, и их число продолжает расти. У сайтов Google самая большая клиентская база с более чем четверть-миллиардной пользовательской аудиторией. В качестве ведущего домена Google насчитывает 74 млрд посещений в месяц, за ним следует приобретенная в 2006 г. кампанией Google платформа YouTube с 27 млрд посещений¹.

Список активных пользователей социальных сетей в мире достиг 3,8 млрд человек. В среднем люди проводят в социальных сетях более двух часов в день.

¹ Основная статистика и факты поиска Google (2020 г.) // Techblog. 2020. 15 октября. URL: https://techblog.sdstudio.top/osnovnaja-statistika-i-fakty-poiska-google-2020-g/#1_Kazduu_sekundi_v_Google_vypolnaetsa_40_000_zaprosov (дата обращения: 17.02.2021).

Facebook с 2,5 млрд активных пользователей в месяц доминирует среди социальных сетей, каждую минуту в нем регистрируются 400 новых пользователей². Приобретенная в 2010 г. компанией Facebook социальная сеть Instagram имеет более миллиарда активных пользователей в месяц. У Instagram в 58 раз больше активности на одного подписчика, чем у Facebook. В совокупности «инстаграмеры» «лайкают» в день более четырех миллиардов сообщений³.

Вместе с распространением интернета стремительно растет рынок электронной коммерции. Общий объем продаж на этом рынке в 2019 г. составил 4,2 трлн долл. В 2022 г. это число превысит 6,54 трлн долл. По данным Statista, доля розничных продаж в электронной коммерции составляет около 15 %, и с каждым годом растет⁴. По результатам 2020 г. более половины онлайн-продаж в мире приходится на Amazon. Только лишь в Америке компания ежедневно осуществляет более 3 млн заказов⁵.

Взглянув на список публичных финансовых предприятий с наибольшей рыночной капитализацией, мы обнаружим, что на второй квартал 2020 г. его возглавляли пять крупнейших IT-компаний: Alphabet (Google), Apple, Facebook, Amazon, Microsoft. Так называемая большая пятерка GAFAM с 2016 г. регулярно занимает первые позиции в рейтинге самых ценных компаний мира⁶. За последние два года четырем из пяти гигантов удалось преодолеть порог капитализации в 1 трлн долларов⁷. Совокупная стоимость этих пяти компаний превысила 5,3 трлн долл.,⁸ что составило более 40 % суммарной стоимости акций всех компаний, торгующихся на американской бирже Nasdaq⁹. Анализируя динамику развития технологигантов, трудно не обратить внимание на беспрецедентный рост объемов и скорости их рыночной стоимости. Пятерка лидеров IT-индустрии за последние несколько лет продемонстрировала невиданный в истории экономический рывок, побив все рекорды концентрации капитала.

Глобальные трансформации, вызванные цифровыми технологиями, порождают дебаты о необходимости разработки новых институтов и правил регулирования экономики и общества [Lobel, 2016]. Феномен роста крупнейших игроков цифровой индустрии воспринимается рядом аналитиков как индикатор качественного изменения мировой политэкономической системы и в академической среде отражается введением в дискуссию таких терминов, как «цифровой капитализм» [Staab, 2019], «платформенный капитализм» [Srnicek, 2016; Langley, Leyshon,

² Общая статистика Facebook. URL: <https://www.websitehostingrating.com/ru/facebook-statistics/#chapter-1> (дата обращения: 17.02.2021).

³ 40 + Instagram Статистика и факты для 2020. URL: <https://www.websitehostingrating.com/ru/instagram-statistics/> (дата обращения: 17.02.2021).

⁴ Статистика Интернета 2020: сайты, домены, хостинг, трафик. URL: <https://sdvv.ru/articles/elektronnaya-kommertsiya/statistika-interneta-2020-sayty-domeny-khosting-trafi/> (дата обращения: 17.02.2021).

⁵ Статистика Amazon. URL: <https://amzmln.com/stati/amazon/statistika-amazon/> (дата обращения: 17.02.2021).

⁶ Ovide Sh., Molla R. Technology Conquers Stock Market. 2016. August 2. URL: <https://www.bloomberg.com/opinion/articles/2016-08-02/tech-giants-form-fab-five-to-dominate-stock-valuation-chart> (дата обращения: 17.02.2021).

⁷ List of Public Corporations by Market Capitalization // Wikipedia. URL: https://en.wikipedia.org/wiki/List_of_public_corporations_by_market_capitalization#2000 (дата обращения: 17.02.2021).

⁸ Größte Unternehmen der Welt nach ihrem Marktwert im Jahr 2020 // Statista. URL: <https://de.statista.com/statistik/daten/studie/12108/umfrage/top-unternehmen-der-welt-nach-marktwert/> (дата обращения: 17.02.2021).

⁹ NASDAQ 100 // Insider. URL: https://markets.businessinsider.com/index/nasdaq_100 (дата обращения: 17.02.2021).

2016] или «следящий капитализм» [Zuboff, 2018]. Каждый термин отражает специфику исследовательского подхода. Цифровой капитализм рассматривает экономическую логику бездефицитных рынков, платформенный капитализм делает ставку на рассмотрение технологической основы, следящий капитализм концентрируется на аспекте сбора поведенческих данных.

С технологической точки зрения ИТ-гиганты — это цифровые платформы, предлагающие услуги на информационном рынке. Платформа представляет собой «киберпространство», формируемое при помощи алгоритмов, в которых пользователи совершают операции, взаимодействуют и заключают сделки [Zysman, Kenney, 2016: 11]. Операции могут быть классифицированы как экономические, социальные или технические действия участников. С другой стороны, под платформами понимаются управляемые компьютером, программируемые и алгоритмически структурированные технологические архитектуры, на которых может быть основано множество конкретных приложений и которые сегодня образуют центральную техническую инфраструктуру сети [Gillespie 2010: 350].

Платформы условно можно разделить на несколько основных типов. Прежде всего, это поисковые платформы. Другой тип — сети и платформы для обмена сообщениями, например Facebook, WhatsApp, Instagram, Twitter или Snapchat. Кроме того, важную роль играют медиаплатформы, такие как YouTube, Netflix или Spotify. Особое место занимают торговые платформы, среди которых Amazon, Alibaba, eBay или Zalando. Стоит также обозначить посреднические платформы, например, в сфере услуг вождения (Uber, Lyft), бронирования путешествий и гостиничных номеров (Airbnb Expedia, booking.com) или знакомств (Match, Parship) [Dolata, 2019].

Существует множество платформ, и не все из них техногиганты. Некоторые еще находятся в положении независимых стартапов, другие уже поглощены более крупными и стали частью холдинга одного из монополистов (как YouTube и WhatsApp). Приведенные выше статистические данные иллюстрируют тенденцию современной цифровой индустрии к монополизации. Этому факту есть вполне конкретные объяснения. Ключевая особенность развития современных технологий заключается в принципе «победитель получает все», опирающемся на пять специфических для платформенной архитектуры факторов [Moore, Tambini, 2018: 33].

1. Программное обеспечение, формирующее платформы, имеет высокие фиксированные затраты на разработку, но минимальные расходы на копирование и распространение цифрового контента. То есть удельные затраты на создание платформ обратно пропорциональны объему продаж. Это дает значительное конкурентное преимущество политике максимально быстрого и широкого распространения продукта на рынке (фактор масштабирования).

2. Платформы вбирают в себя свойства сетевых структур, полезность которых прямо пропорциональна количеству подключенных к сети участников [Jeffrey, 1974: 16]. В более выигрышном положении оказывается сеть, успевшая включить в себя наибольшее количество абонентов (сетевой фактор).

3. На платформах, осуществляющих связь продавцов с покупателями, возникает симметрия между предложением и спросом. Количество участников одной части рынка (например, отелей) зависит от количества участников другой части рынка (путешественников, бронирующих номера) (фактор симметрии).

4. Эффект, связанный со сложностью переключения с одной платформы на другую, где требуется дополнительное усилие пользователя по освоению нового интерфейса. Чем больше объем такого обучения, тем выше стоимость переключения.

5. Платформы, обладая практически неограниченными возможностями транспортировки данных, могут осуществлять их повсеместный сбор и использовать для обучения искусственного интеллекта. Получая доступ к большим данным, лидеры рынка становятся обладателями информации для тактического и стратегического улучшения качества продукта, что позволяет им выигрывать конкуренцию у игроков, не имеющих этого доступа (фактор больших данных).

Все эти эффекты действуют в одном направлении, повышая шансы победителя и закрепляя его положение лидера, способствуют образованию доминантных игроков и монополизации рынков. Однако последний (пятый) эффект имеет особое значение. Он основан на применении технологии машинного обучения искусственного интеллекта. Широкомасштабное применение этой технологии проводит черту между периодом «оцифрованного капитализма», элементы которого проявились в 1960-е годы (после первых попыток автоматизации процессов производства), и эрой современного «цифрового капитализма», в которой поведенческие данные используются в качестве основы генерирования добавленной стоимости, а сбор этих данныхложен в основу бизнес-моделей техногигантов, занявших ключевые позиции влияния на рынке цифровых технологий [Staab, 2019: 38].

Бизнес-модели крупнейших акторов цифровой индустрии различны. Не все они основаны на продаже онлайн-рекламы, однако у всех есть общее свойство — они собирают данные пользователей и используют машинное обучение искусственного интеллекта для их обработки в целях дальнейшей оптимизации своего продукта. Американская экономистка Шошана Зубофф провела подробный критический разбор бизнес-моделей техногигантов GAFAM, в котором показала, что в основу создания добавленной стоимости всех этих корпораций положен сбор пользовательских поведенческих данных [Zuboff, 2018]. Анализ действий ключевых игроков цифрового рынка показывает, что сбор и обработка больших данных — первостепенная задача всех крупнейших акторов IT-индустрии. Постепенно в круг собирателей данных вовлекаются все больше отраслей и участников. Интернет вещей, индивидуальная медицина, автономное вождение — это лишь некоторые примеры способов массового сбора поведенческих данных. В современном мире следящего капитализма большие данные стало принято называть «новой нефтью» по аналогии с тем, что именно нефть была основой строительства крупного капитала прошлого столетия.

Появление фордизма и тейлоризма стало началом применения новой экономической логики, сформировавшей капитализм массового производства XX столетия. Подобным образом на рубеже XXI века в Кремниевой долине была открыта экономическая логика, потеснившая классический принцип массового производства с вершины олимпа прибыльности. Построенная на производстве продукта, основанного на сборе данных, эта логика стала фундаментом бизнес-модели капитализма XXI века [Климович, 2020а].

Логично, что вопрос концентрации капитала в конечном счете должен сводиться к вопросу о концентрации власти. В современных условиях, когда технологии

усиливают тенденции к монополизации, проблема распределения и баланса власти приобретает новый масштаб. В такой ситуации перед социологами стоит задача проведения всестороннего анализа цифрового капитализма, а вместе с тем и пересмотра роли и значения технологий в обществе.

Подходы к критическому анализу цифрового капитализма

Несмотря на то что цифровой капитализм использует качественно новую логику создания добавленной стоимости, отличающую его от классического капитализма, все же он не перестает быть капитализмом — формацией, основанной на частной собственности и свободном рыночном предпринимательстве, направленном на повышение прибыли и увеличение капитала. Поэтому, делая попытку осмыслиения социальных эффектов цифрового капитализма, логично обратиться к опыту критики индустриального общества, проделанной классиками социологии, авторами критических теорий XX столетия.

Ключевым свойством платформенной архитектуры оказывается сбор и обработка информации. Именно сбор больших данных обеспечивает IT-лидерам конкурентное преимущество, позволяет удерживать доминантное положение и распространять экономическое политическое и общественное влияние. Возможность сбора и анализа поведенческих данных приводит к открытию нового источника знаний — новой силы, способной стать локомотивом развития. По мере того, как поведение людей превращается в информацию, распределение прогностического знания в обществе становится доминирующим принципом социального порядка. Знание, извлекаемое из поведенческих данных, генерируемых в экспоненциально возрастающем объеме, распределяется неравномерно. Это явление усугубляется по мере накопления данных на серверах IT-гигантов, которые отнюдь не заинтересованы делиться ими с остальным миром.

Проблема дифференциации доступа к экономическим, социальным, политическим и культурным ресурсам была рассмотрена основателем французской социологии Эмилем Дюркгеймом более ста лет назад. По мнению французского социолога, разделение труда легло в основу установления разного рода взаимозависимостей, которые, объединив членов социума в более широком контексте солидарности, способствовали установлению социального и морального порядка современного индустриального общества [Дюркгейм, 1991]. Однако ученый отметил, что чрезмерная социальная дифференциация, вызванная неравноправным распределением ресурсов, становится источником патологий, в результате которых возникает так называемое аномическое разделение, вызывающее социальную дистанцию, несправедливость и конфликты вместо установления взаимодействия и взаимозависимости, формирующих общественную солидарность.

В цифровом обществе, когда право распределения прогностическим знанием все отчетливее проявляет себя как форма власти, неравномерность распределения этого знания угрожает последствиями, о которых предостерегал Дюркгейм. Наблюдение, осуществляемое платформами посредством сбора и обработки персонализированных данных, обеспечивает IT-корпорациям конкурентное преимущество и создает прогрессирующую асимметрию знаний, повышает риски возникновения эффектов, описанных французским социологом.

В контексте анализа последствий неравномерности распределения доступа к новому типу знания имеет смысл вспомнить о классическом «эффекте Матфея». Этот термин, предложил американский социолог Роберт Мертон для описания феномена прогрессирующего неравенства, состоящего в том, что сторона, уже обладающая преимуществами, продолжает их накапливать, в то время как другая, изначально ограниченная, оказывается обделена еще сильнее [Мертон, 1993]. Как было показано выше, этот эффект отчетливо проявляется в платформенной архитектуре.

Отдельно отметим, что проблема цифрового неравенства, обсуждаемая в контексте неравномерного доступа к интернету [Быков, Халл, 2011], скорее всего, не является специфической проблемой цифрового капитализма. Более того, логика цифрового капитализма как раз состоит в том, чтобы обеспечить максимально широкийхват сбора поведенческих данных, и для реализации этой цели инвесторы, зарабатывающие на продаже предиктивной информации, по всей видимости, не будут жалеть средств на развитие инфраструктуры сбора транспортировки и обработки данных, что в принципе подтверждается статистикой о динамике распространения интернета по всему миру¹⁰.

Продолжая сопоставление критики индустриального общества с контекстом цифрового капитализма, нельзя не обратиться к опыту критических теорий, предметом которых является анализ буржуазно-капиталистических отношений, направленный на раскрытие механизмов господства и подавления. Основатели Франкфуртской школы Теодор Адорно и Макс Хоркхаймер выяснили, что благодаря чрезмерной фокусировке капиталистического производства на извлечении прибыли и приумножении капитала оптимизация и эффективность в индустриальном обществе были возведены в ранг высшего приоритета.

Рассмотрев эффективность применительно к понятию «самоцель», они выделили тип человеческой деятельности, направленный на повышение эффективности, но игнорирующий вопрос о цели этого повышения. По мнению немецких социологов, этот принципложен в основу так называемого инструментального разума, который начал доминировать в западной культуре со времен эпохи Просвещения. Примат господства инструментального разума стал питающей почвой, на которой впоследствии пустил корни современный западный капитализм. Развитый модерн характеризуется тем, что общество неизбежно превращается в систему, в которой все элементы соподчинены функциональной зависимостью. Каждая часть воспринимает другую часть в качестве ресурса. То, что не является ресурсом, игнорируется системой и, исключаясь из внимания, низводится до статуса неинтересного, безразличного, исключенного. Напротив, то, что является ресурсом, используется с максимальной эффективностью, вовлекаясь в структуру социальных связей. Логика капиталистического общества полностью отражает стремление человека инструментализировать окружающую среду. В конечном счете этой цели подчиняется вся деятельность общества, включая современную

¹⁰ Вся статистика интернета на 2020 год — цифры и тренды в мире и в России. URL: <https://www.web-canape.ru/business/internet-2020-globalnaya-statistika-i-trendy/> (дата обращения: 17.02.2021); Список стран по числу пользователей интернета. URL: https://ru.wikipedia.org/wiki/Список_стран_по_числу_пользователей_Интернета (дата обращения: 17.02.2021).

науку. Говоря словами Т. Адорно и М. Хоркхаймера, единственное, чему люди хотят научиться у природы, это как ее использовать, чтобы полностью поработить [Хоркхаймер, Адорно, 1997: 17].

Перенося критику немецких классиков на цифровой капитализм, можно провести некоторые параллели. В частности, любой IT-предприниматель, как и всякий субъект капиталистических отношений, преследует цель повышения оптимизации производимого продукта. Однако в случае с глобальными платформами, чья бизнес-модель опирается на извлечение прибыли за счет сбора персонализированных данных, происходит нечто особенное. Оптимизация их продукта основана на анализе поведенческой информации и создании нового типа приватизированного прогностического знания, что в итоге делает возможным доминирование IT-концернов во всех областях деятельности — экономике, политике, науке, образовании, медицине. Если в индустриальном обществе в погоне за прибылью инструментализации подвергалась окружающая природа, то в цифровом обществе инструментализации подвергается сам человек. Поставленный на службу производства удобного и «бесплатного» продукта, субъект, предоставляемый бесконтрольный доступ к персональным данным, жертвует своим правом свободного, независимого выбора. Механизм этого обмена обществу еще предстоит изучить, но для данной статьи важно подчеркнуть, что у IT-гигантов есть все основания скрывать стоящую за сбором персональных данных суть инструментализации человека и вуалировать реальное положение дел подобно тому, как это делали крупные концерны эпохи индустриализации, лоббируя интересы производств и отраслей, наносящих непоправимый вред окружающей среде.

В контексте оценки рисков влияния цифровых технологий на общество следовало бы обратить внимание на критический анализ классического капитализма, проделанный еще одним представителем Франкфуртской школы Гербертом Маркузе. В работе «Одномерный человек» ученый анализирует саму возможность противостояния негативным эффектам, которые возникают в результате распространения технического прогресса и модернизации общества. Изучая происходившие в XX веке изменения, Маркузе приходит к выводу, что в связи с успехами применения научно-рационалистической методологии как в научном, так и в публичном пространстве устанавливается так называемое одномерное мышление, исключающее альтернативные подходы к развитию познания [Marcuse, 1970]. Немецкий классик констатирует, что разнообразие перспектив развития общества схлопывается в одномерность отсутствия альтернативы не путем наложения запрета, а с помощью отключения у индивида самого осознания возможности альтернативы. Любой выход «за рамки» маркируется как нецелесообразный, бесполезный и даже нереалистичный, таким образом превращаясь, по выражению Роберта Мертона, в самоисполняющееся пророчество [Мертон, 2006]. То есть с момента того, как человек принимает безальтернативность, она реализуется как действительный факт. Объясняя причины сформировавшейся в результате технического прогресса безальтернативности, Маркузе различает истинные и ложные потребности человека. Истинные потребности возникают на основе собственных представлений и интересов личности, а ложные имеют природу внешнего воздействия, навязываются ему извне. Именно внешнее воздействие,

по мнению представителя Франкфуртской школы, ответственно за нормализацию одномерного мышления.

Применяя доводы немецкого социолога к контексту анализа платформенного капитализма, следует обратить внимание на особое значение злоупотребления технологиями сбора поведенческих данных. Простая логика подсказывает, что чем шире возможности манипулирования человеком, тем менее заметен этот процесс для него самого, ровно как верно и обратное. В мире неконтролируемого сбора поведенческих данных с новой силой возрастают риски реализации сценария Маркузе, когда одномерное мышление угрожает отстранением человека от самостоятельного формирования собственного будущего. Платформенная архитектура, обеспечивающая сбор персональных данных, создает благоприятную среду для распространения технологий подталкивания [Талер, Санстейн, 2017]. В сочетании с технологией микротаргетинга, опирающегося на анализ больших данных и действующего индивидуально на каждого субъекта в автоматизированном режиме, возрастают риски манипулирования массовым сознанием. Создается угроза отключения альтернативных вариантов развития целого общества путем персонального убеждения каждого пользователя в нецелесообразности или невозможности их реализации. По словам Паула Шварца, чем больше информации известно о человеке, тем легче им управлять. Поэтому, чтобы защитить лежащую в основе демократии свободу, необходимо тщательно продумать, как общество должно обращаться с информацией, каким образом оно будет гарантировать равноправный доступ к ней и как должна обеспечиваться ее надежная защита [Schwartz, 1989].

Продолжая сопоставлять анализ рисков цифрового капитализма с классической критикой индустриального общества, нельзя оставить без внимания, пожалуй, самого известного представителя Франкфуртской школы второго поколения Юргена Хабермаса. По мнению Хабермаса, модернизация предоставляет человеку определенный шанс построения свободного демократического общества, реализовать который можно посредством механизма ограничения власти через публичность [Хабермас, 2016]. Такой принцип контроля подразумевает выяснение истины не в режиме отсылки к авторитету, а через всеобщее и добровольное принуждение лучшего аргумента.

Хабермас — представитель теории дискурса, в его концепции закон и право определяются истиной, а не авторитетом. Публичное пространство состоит из участников, выполняющих роль гражданских репрезентантов, активно защищающих свою позицию.

Таким образом, по мере разворачивания модернизационных процессов именно публичность становится медиумом построения демократии. Дискурс, выступая в качестве процедурного момента установления истины, становится эссенциальным свойством демократического строя. Демократия представляет собой процедурное явление, реализованное посредством широкомасштабного применения дискурсивных практик, в рамках реализации которых Хабермас различает два типа действий. Коммуникативное действие происходит в контексте планов участников коммуникации и направлено на достижение взаимного понимания между ними. Такое действие построено на совместном координировании

своих намерений в зависимости от прояснения и понимания противоположных позиций. Стратегическое действие направлено на достижение индивидуального успеха. Реализуя стратегическое действие, актор поступает, невзирая на интересы и потребности другого, принимая во внимание лишь возможности, обеспечиваемые доступом к ресурсам. В стратегическом действии акторы стремятся взять свои цели за ориентир и пытаются оказать влияние на решения других акторов [Habermas, 1988: 131].

Такие элементы человеческой реальности, как знания, социальный порядок, общественное пространство, относятся к структурам «жизненного мира» и могут быть репродуцированы исключительно посредством механизма взаимопонимания, на основе коммуникативного действия. В процессе модернизации общества от «жизненного мира», в котором доминирует коммуникативное действие, постепенно отделяется «система», включающая в себя область экономики и административного управления. Системные процессы, репродуцируемые на основе стратегического действия, достигая высокой степени эффективности, имеют тенденцию распространения, и в ряде случаев просматривается тренд вытеснения этими процессами менее эффективных, не приносящих сиюминутный результат коммуникативных действий. Критика Хабермаса показывает, что в ряде случаев этот феномен следует диагностировать как социальную патологию, при этом сам процесс ученый называет колонизацией жизненного мира системой.

Применяя терминологию Хабермаса к анализу цифрового капитализма, следует отметить, что коммуникативное действие, составляющее основу жизненного мира, опирается на право свободного выбора, который участники коммуникации предположительно должны совершить в процессе поиска взаимного понимания. То есть акторы должны находиться в равных категориях доступа к информации о мотивах действия друг друга. В случае применения технологий микротаргетинга и машинного обучения собираемые персональные данные человека используются для формирования предсказаний его поведения. При взаимодействии с платформой пользователь, передавая свои персональные данные, делится информацией о собственных мотивах, однако мотивы действия платформы остаются для него закрытыми и неясными. Чем больше пользовательских данных собирает платформа, тем больше у нее информации о том, как будет действовать пользователь. Напротив, чем более сложной становится платформа, тем меньше пользователь имеет представление о принципах ее функционирования. Взаимодействуя с платформой в режиме ассиметричного доступа к прогностической информации, человек с каждым актом взаимодействия теряет контроль над собственным действием, передавая его инстанции — обладателю этой информации, то есть самой платформе. Такое взаимодействие не может быть отнесено к классу коммуникативного действия, так как участники находятся в разных категориях доступа к информации о мотивах действия друг друга. Однако оно имеет все признаки стратегического действия ввиду того, что основано на доступе к ресурсу — прогностической информации. Повсеместное распространение действия платформ в современном мире, скорее всего, следует интерпретировать как процесс колонизации жизненного мира системой, при котором логика эффективности вытесняет логику понимания и при необходимости может оправдывать применение механизмов манипулиро-

вания [Климович, 2019]. Этот вывод вызывает особое беспокойство в контексте анализа рисков концентрации власти цифрового капитализма.

IT и государство

Анализируя проблему монополизации власти, необходимо рассмотреть тему взаимодействия IT-индустрии с государственными структурами. Характер их взаимодействия зависит от того, с каким типом государства мы имеем дело — демократическим или авторитарным. Несмотря на то что наблюдаются некоторые общие процессы, характерные для обеих ситуаций, все же различия достаточно велики, так что проанализируем оба варианта по отдельности. Вариант демократического государства хорошо иллюстрирует пример США, а случай взаимодействия IT-технологий с авторитарным типом государства — пример Китая.

Рассматривая ситуацию в Соединенных Штатах, начнем с того, что на практике часто наблюдается тесное взаимодействие техногигантов с государственными структурами. Достаточно посмотреть на совместные проекты в области образования, медицины, автоматизации муниципального и государственного управления. Особенного внимания заслуживает сотрудничество с военными разведывательными и правоохранительными органами [Черняк, 2013] которое объясняется не только стремлением крупных финансовых институтов путем сближения с властными структурами влиять на регулирование цифровой отраслью. Степень и характер взаимодействия убедительно демонстрируют, что сами государственные институты заинтересованы в сотрудничестве с IT-корпорациями. В каком-то смысле заинтересованность государственных органов очевидна, ведь IT-технологии значительно повышают эффективность менеджмента. Какую бы область мы ни брали в качестве примера — медицину, образование, безопасность, — везде применение цифровых технологий приводит к значительному повышению эффективности управления.

Возвращаясь к вопросу концентрации власти цифровых технологий, при более внимательном анализе можно обнаружить, что государственным структурам в определенном смысле выгодно положение монополизированного рынка IT. Государственный контроль, имеющий целью управление, и сбор поведенческих данных частными компаниями для доминирования на рынках тесно связаны между собой. Монополизм цифровых платформ выступает ключом к успеху государственного менеджмента. Сотрудничающий с властями глобализированный технологический сектор — гораздо более удобный союзник в проекте государственного контроля, чем фрагментированный, состоящий из мелких независимых субъектов.

Особенно высокая степень взаимодействия обнаруживается в области сотрудничества с государственными системами безопасности. В связи с развитием монополизированного IT-сектора службы безопасности получили возможность сократить значительную часть расходов по сбору информации, что сделало их работу несопоставимо более эффективной. Благодаря цифровым платформам теперь устройства самих граждан собирают большую часть данных, которые государственные службы используют для наблюдения за ними [Zuboff 2018: 431]. Граждане сами оплачивают эти устройства, а также услуги, к которым подключены, и сами кропотливо вводят данные, связанные с регистрацией фактов об их жизни,

мнениях и предпочтениях. Склонность к симбиозу государственных структур и цифровых платформ объясняется тем, что Big Tech, собирая данные, может легко это делать для служб безопасности, а агентства безопасности — следить за тем, чтобы правительства не ограничивали деятельность Big Tech. Исполнительным органам удобно прибегнуть к возложению дорогостоящих государственных обязанностей на крупные технологические компании, например автоматическое выявление нарушения авторских прав, обнаружение террористического и экстремистского контента. Дilemma состоит в том, что можно пытаться сократить власть цифровых гигантов, лишив их монопольного права владения рынками, либо возлагать на них контроль над выявлением нежелательных действий пользователей, но проблематично сделать и то и другое одновременно¹¹.

Анализ взаимодействия технологического сектора со структурами демократического государства показывает амбивалентный характер этих отношений, проявляющийся в том, что, с одной стороны, демократическое государство стоит на защите индивидуальной свободы и прав человека, с другой стороны, заинтересовано в эффективном управлении и обеспечении безопасности граждан. Говоря о демократическом государстве, мы прежде всего подразумеваем наличие в нем разделения независимых ветвей власти: исполнительной, законодательной и судебной. Наличие этих институтов эссециально для демократического строя. Применительно к нашему анализу концентрации власти в эпоху цифрового капитализма следует отметить, что исполнительная ветвь более склонна к «коллаборационизму» с монополизированным технологическим сектором. Для восстановления баланса властных полномочий следует учитывать этот эффект и делать опору прежде всего на гражданское общество, судебную и законодательную ветви власти.

Рассматривая китайский сектор рынка информационных технологий, следует отметить, что, хоть он и отделен от остального интернета великой огненной стеной, все же эффекты монополизации проявлены здесь не меньше, чем на мировом рынке. Аналогом американского GAFAM здесь выступает тройка крупнейших игроков: Tencent, Alibaba, Baidu, в которой сконцентрирован основной ресурс китайской IT-индустрии. В отличие от демократических стран, частные компании Китая находятся в совершенно других отношениях с государством. Имея относительную свободу в области разработки продукта, китайские техногиганты находятся под полной зависимостью от политической воли коммунистической партии. Правительство народной республики непосредственно определяет стратегию развития страны, в том числе глобальные планы по внедрению IT-технологий в общественную жизнь. Осознавая, что использование технологий искусственного интеллекта открывает возможность централизованного и эффективного управления огромной страной, правительство народной республики рассчитывает, что в перспективе суперкомпьютерам будет передана роль создателей генерального плана развития экономики.

Говоря о роли, которую цифровые технологии играют в формировании общественной жизни Китая, имеет смысл обратиться к примеру создания системы

¹¹ Doctorow C. How to Destroy Surveillance Capitalism // OneZero. 2020. August 26. URL: <https://onezero.medium.com/how-to-destroy-surveillance-capitalism-8135e6744d59> (дата обращения: 17.02.2021).

социального кредита. [Климович, 2020б]. Цель системы состоит в том, чтобы осуществлять всесторонний контроль за людьми посредством сбора и обработки их персональных данных. В рамках реализации системы предполагается анализ общественно-политического поведения частных лиц, компаний и других организаций для определения их «социальной репутации», на основе которой будет проводиться поощрительная и санкционная политика регулирования¹². В этом проекте государство активно задействует ресурсы частных компаний, привлекая весь арсенал влияния. Данные о гражданах, собираемые IT-компаниями, по указанию правительства используют в расчетах оценки социального кредита¹³. Крупнейший оператор поисковых систем Китая Baidu указан как один из центральных разработчиков платформы. Также в разработку системы рейтинга вовлечены компании Alibaba и Tencent.

Примечательно, что основным мотивом создания тотального цифрового контроля над населением правительство называет заботу о безопасности граждан. Поясняя цели государственной политики, главный идеолог суверенного интернета Китая Ван Хунин сказал, что в обязанности руководства входит содействие безопасности, наведение порядка и создание надежного, стабильного и процветающего цифрового мира [Scheuer, 2018]. Отметим, что какими убедительными ни казались бы соображения о цифровом суверенитете, вынесение решений о санкционировании за рамки судебной системы и передача их на алгоритмический уровень вступают в противоречие с основными ценностями свободы личности, лежащими в основании современной демократии и международного права.

Заканчивая краткий обзор специфики взаимодействия IT-гигантов с государственными органами Китая, отметим, что характер действия Пекинского правительства в вопросе распределения власти не уникален. Принципы, на которых строится политика Китая, в той или иной степени свойственны всем авторитарическим системам. Они подразумевают контроль за гражданами и манипуляцию их мнением. Специфической особенностью Китайской народной республики является сильно развитый сектор цифровой индустрии, обеспечивающий эффективный контроль населения, способствующий повышению и без того высокого уровня концентрации власти в руках немногочисленной государственной элиты.

Выводы

Подытоживая проделанный в статье анализ рисков изменения властных отношений, можно прийти к заключению, что тренд концентрации власти вокруг крупнейших IT-компаний в ближайшее время будет сохраняться. Продолжению этой тенденции, очевидно, способствует сама бизнес-модель техногигантов, построенная на сборе персональных данных в целях оптимизации продукта частных компаний. Стоит ожидать, что государственные органы отчасти будут способство-

¹² State Council Guiding Opinions concerning Establishing and Perfecting Incentives for Promise-Keeping and Joint Punishment Systems for Trust-Breaking, and Accelerating the Construction of Social Sincerity // China Copyright and Media. 2016. October 18. URL: <https://chinacopyrightandmedia.wordpress.com/2016/05/30/state-council-guiding-opinions-concerning-establishing-and-perfecting-incentives-for-promise-keeping-and-joint-punishment-systems-for-trust-breaking-and-accelerating-the-construction-of-social-sincer/> (дата обращения: 17.02.2021).

¹³ Botsman R. Big Data Meets Big Brother as China Moves to Rate Its Citizens //Wired. URL: <https://www.wired.co.uk/article/chinese-government-social-credit-score-privacy-invasion> (дата обращения: 17.02.2021).

вать этому процессу, но в демократических странах такая активность будет исходить в основном от сектора безопасности. Законодательная и судебная ветви, по всей видимости, должны стать опорой противостояния процессам дальнейшей концентрации власти цифровых гигантов. Однако как в демократических, так и в авторитарных обществах главная стратегия решения вопроса баланса власти — это опора на гражданское общество. Именно оно в первую очередь заинтересовано в защите собственных прав и является бенефициаром согласованного распределения властных полномочий между общественно контролируемыми институтами власти. Для того чтобы повысить сознательную гражданскую активность, необходимо довести до общественности суть процессов, лежащих в основе экономических моделей IT-гигантов, раскрыть природу концентрации власти в руках новоиспеченных лидеров эпохи цифрового капитализма.

В этом процессе особая роль ложится на представителей академической среды. Возможно, наступило время разработки проекта анализа цифровых буржуазно-капиталистических отношений, направленного на раскрытие механизмов господства и подавления в эпоху информационных технологий. Оираясь на опыт, предпринятый авторами критических теорий прошлого века, академический мир должен стать в авангарде просвещения гражданского общества цифровой эпохи. Благодаря критическому мышлению общество отстояло демократические свободы в эпоху индустриального капитализма прошлого века. То же самое предстоит сделать сегодняшнему поколению по отношению к цифровому капитализму.

Список литературы (References)

- Быков И. А., Халл Т. Э. Цифровое неравенство и политические предпочтения интернет-пользователей в России // Полис. Политические исследования. 2011. № 5. С. 151—163. URL: <https://www.politstudies.ru/article/4473> (дата обращения: 26.02.2021).
- Bykov I.A., Hall T. E. (2011) Digital Divide and the Internet-Users Political Preferences in Russia. *Polis. Political Studies*. No. 5. P. 151—163. URL: <https://www.politstudies.ru/article/4473> (accessed: 26.02.2021). (In Russ.)
- Дюркгейм Э. О разделении общественного труда. Метод социологии / пер. с фр. и послесловие Л. Б. Гофмана. М.: Наука, 1991.
- Durkheim É. (1991) *De la division du travail social*. Moscow: Nauka. (In Russ.)
- Маркузе Г. Одномерный человек: исследование идеологии развитого индустриального общества / пер. с англ. М.: ACT, 2002.
- Marcuse H. (2002) One-Dimensional Man: Studies in the Ideology of Advanced Industrial Society. Moscow: AST. (In Russ.)
- Мerton Р. К. Эффект Матфея в науке, II: накопление преимуществ и символизм интеллектуальной собственности // THESIS. 1993. Вып. 3. С. 256—276.
- Merton R. K. (1993) The Matthew Effect in Science, II: Cumulative Advantage and the Symbolism of Intellectual Property. THESIS. Vol. 3. P. 256—276. (In Russ.)
- Мerton Р. К. Социальная теория и социальная структура / пер. с англ. Е. Н. Егоровой. М.: ACT, Хранитель, 2006.

Merton R. K. (2006) Social Theory and Social Structure. Moscow: AST; Khranitel. (In Russ.)

Талер Р. Санстейн К. (2017). Nudge. Архитектура выбора. Как улучшить наши решения о здоровье, благосостоянии и счастье. М.: Манн, Иванов и Фербер.

Thaler R., Sunstein C. (2017) Nudge: Improving Decisions About Health, Wealth, and Happiness. Moscow: Mann, Ivanov and Ferber. (In Russ.)

Хабермас Ю. Структурное изменение публичной сферы: Исследования относительно категории буржуазного общества. М.: Весь Мир, 2016.

Habermas J. (2016) Strukturwandel der Öffentlichkeit: Untersuchungen zu einer Kategorie der bürgerlichen Gesellschaft. Moscow: Ves Mir. (In Russ.)

Хоркхаймер М., Адорно Т., Диалектика Просвещения. Философские фрагменты. М.; СПб.: Медиум: Ювента, 1997.

Horkheimer M., Adorno T. (1997) Dialektik der Aufklärung. Philosophische Fragmente. Moscow, St. Petersburg: Medium; Yuventa.

Черняк Л. Аналитика двойного назначения // Открытые системы — СУБД. 2013. № 10. URL: <https://www.osp.ru/os/2013/10/13039073> (дата обращения: 26.02.2021).

Chernyak L. (2013) Dual-Use Analytics. Otkrytye sistemy — SUBD. No. 10. URL: <https://www.osp.ru/os/2013/10/13039073> (accessed: 26.02.2021). (In Russ.)

Dolata U. (2019) Internet Plattformen Regulierung. Koordination von Märkten und Kuratierung von Sozialität im Internet. *Berliner Journal für Soziologie*. Vol. 129. P. 179—206. <https://doi.org/10.1007/s11609-020-00403-9>.

Gillespie T. (2010) The Politics of “Platforms”. *New Media & Society*. Vol. 12. No. 3. P. 347—364.

Habermas J. (1988) Theorie des kommunikativen Handelns. In 2 Bände. Band 2. Zur Kritik der funktionalistischen Vernunft. Frankfurt am Main: Suhrkamp.

Jeffrey R. (1974) A Theory of Interdependent Demand for a Communications Service. *Bell Journal of Economics and Management Science*. Vol. 5. No. 1. P. 16—37.

Langley P., Leyshon A. (2016) Platform Capitalism: The Intermediation and Capitalization of Digital Economic Circulation. *Finance and Society*. Vol. 2. No. 1. URL: https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwjH0ZSk0PvAhVdAhAIHc1rBygQFjAAegQIARAD&url=http%3A%2F%2Ffinanceandsociety.ed.ac.uk%2Fojjs-images%2Ffinanceandsociety%2FFFS_EarlyView_LangleyLeyshon.pdf&usg=A0vVaw2jLzeFjuJtaZgILCAV7tPA (accessed: 23.02.2021).

Lobel O. (2016) The Law of the Platform. *San Diego Legal Studies Paper*. No. 16—212. *Minnesota Law Review*. URL: https://papers.ssrn.com/sol3/Delivery.cfm/SSRN_ID2742380_code514132.pdf?abstractid=2742380&mirid=1 (accessed: 23.02.2021).

Moore M., Tambini D. (2018) Digital Dominance. The Power of Google, Amazon, Facebook, and Apple. New York, NY: Oxford University Press.

Scheuer S. (2018) Der Masterplan. Chinas Weg zur Hightech-Weltherrschaft. Verlag Herder GmbH, Freiburg im Breisgau.

Schwartz P. M. (1989) The Computer in German and American Constitutional Law: Towards an American Right of Informational Self-Determination. *American Journal of Comparative Law*. Vol. 37. No. 4. P. 675—701. <https://doi.org/10.2307/840221>.

Srnicek N. (2016) Platform Capitalism (Theory Redux). Cambridge: Polity.

Staab Ph. (2019) Digitaler Kapitalismus Markt und Herrschaft in der Ökonomie der Unknappheit. Berlin: Suhrkamp Verlag.

Zuboff Sh. (2018) Das Zeitalter des Überwachungs-Kapitalismus. Frankfurt: Campus.

Zysman J, Kenney M. (2016) The Next Phase in the Digital Revolution Platforms, Abundant Computing, Growth and Employment. BRIE Working Paper.

DOI: [10.14515/monitoring.2021.1.1807](https://doi.org/10.14515/monitoring.2021.1.1807)



Т. С. Мартыненко, Д. Е. Добринская

СОЦИАЛЬНОЕ НЕРАВЕНСТВО В ЭПОХУ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА: ОТ ЦИФРОВОГО К АЛГОРИТМИЧЕСКОМУ РАЗРЫВУ

Правильная ссылка на статью:

Мартыненко Т.С., Добринская Д. Е. Социальное неравенство в эпоху искусственного интеллекта: от цифрового к алгоритмическому разрыву// Мониторинг общественного мнения: экономические и социальные перемены. 2021. № 1. С. 171—192. <https://doi.org/10.14515/monitoring.2021.1.1807>.

For citation:

Martynenko T.S., Dobrinskaya D. E. (2021) Social Inequality in the Age of Algorithms: From Digital to Algorithmic Divide. *Monitoring of Public Opinion: Economic and Social Changes*. No. 1. P. 171–192. <https://doi.org/10.14515/monitoring.2021.1.1807>. (In Russ.)

СОЦИАЛЬНОЕ НЕРАВЕНСТВО В ЭПОХУ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА: ОТ ЦИФРОВОГО К АЛГОРИТМИЧЕСКОМУ РАЗРЫВУ

МАРТЫНЕНКО Татьяна Сергеевна — кандидат социологических наук, старший преподаватель кафедры современной социологии социологического факультета, МГУ имени М. В. Ломоносова, Москва, Россия

E-MAIL: ts.martynenko@gmail.com
<https://orcid.org/0000-0002-5101-2381>

ДОБРИНСКАЯ Дарья Егоровна — кандидат социологических наук, доцент кафедры современной социологии социологического факультета, МГУ имени М. В. Ломоносова, Москва, Россия

E-MAIL: darya.dobrinskaya@gmail.com
<https://orcid.org/0000-0002-8691-9637>

Аннотация. Статья посвящена анализу проблемы социального неравенства в эпоху алгоритмов, которые являются одним из элементов технологической инфраструктуры цифрового общества. Социологический интерес к алгоритмам объясняется их активным использованием в процессах коммуникации, принятия решений, управления, контроля и в конечном итоге их ролью в создании ландшафта искусственной социальной среды. Последствия внедрения алгоритмов находят отражение в разных сферах общественной жизни. Негативные эффекты алгоритмизации рассматриваются с точки зрения неравенства возможностей и социально значимых результатов. Скрытый характер воздействия алгоритмов, непрозрачность их работы, нелинейность, непредсказуемость в принятии решений определяют их влияние на систему

SOCIAL INEQUALITY IN THE AGE OF ALGORITHMS: FROM DIGITAL TO ALGORITHMIC DIVIDE

Tatiana S. MARTYNENKO¹ — Cand. Sci. (Soc.), Senior Lecturer at the Department of Modern Sociology, Faculty of Sociology
E-MAIL: ts.martynenko@gmail.com
<https://orcid.org/0000-0002-5101-2381>

Daria E. DOBRINSKAYA¹ — Cand. Sci. (Soc.), Associate Professor at the Department of Modern Sociology, Faculty of Sociology

E-MAIL: darya.dobrinskaya@gmail.com
<https://orcid.org/0000-0002-8691-9637>

¹ Lomonosov Moscow State University, Moscow, Russia

Abstract. The article analyses social inequality in the age of algorithms integral to digital society's technological infrastructure. Algorithms are of interest to sociologists due to their essential role in communication processes, decision-making, management, control, and consequently in creating a landscape of the artificial social environment. The introduction of algorithms affects various aspects of social life. Negative impacts are considered in terms of inequality of opportunities and socially important results. The latent character of the effects, lack of transparency, non-linear nature, and indeterminism in decision-making define the influence of algorithms on the system of social inequality. The article describes algorithm-related mechanisms (system of social ranking, filter bubbles, unequal access to algorithmic awareness).

му социального неравенства. В статье анализируются связанные с функционированием алгоритмов механизмы этого явления (системы социального ранжирования, пузыри фильтров, неравный доступ к алгоритмической осведомленности) и рассматриваются вопросы методологического характера, связанные с концептуализацией понятия «социальное неравенство» в эпоху алгоритмов.

Ключевые слова: социальное неравенство, цифровизация, цифровое неравенство, алгоритмы, алгоритмизация, цифровой разрыв

Keywords: social inequality, digitalization, digital inequality, algorithms, algorithmization, digital divide

Влияние технологий на общество — «горячая» тема большого числа исследований в области социальных наук. Никогда еще в человеческой истории социальный порядок не был в такой зависимости от технологического прогресса¹. Многие повседневные практики сегодня воспроизводятся и регулируются посредством технических устройств, функционирующих на базе программного обеспечения. С помощью алгоритмов навигационного приложения на смартфоне возможно проложить оптимальный маршрут из точки А в точку Б, выбрать музыку для спортивной тренировки или найти подходящий ресторан для романтического ужина. Все эти тенденции не только трансформируют социальные отношения, предполагая новые формы взаимодействия, трудовой деятельности, участия в управлении и многое другое, но и формируют в научном дискурсе новое междисциплинарное поле, где разворачиваются серьезные методологические дискуссии, в центре которых стоит вопрос «что происходит на пересечении собственно социальных и технологических изменений?» [Резаев, Стариков, Трегубова, 2020: 8].

Цифровизация и датификация создали условия для оформления искусственной социальной среды цифрового общества [Резаев, Стариков, Трегубова, 2020; Резаев, Трегубова, 2019], где значимую роль играют алгоритмические системы, коммуникационные сети, платформы и технологии больших данных. Цифровые устройства и сетевые системы, которые, как правило, функционируют на базе алгоритмов, стали частью рутинизированного опыта большинства², а способность алгоритмов выносить четкие и последовательные решения на основе больших массивов данных делает их «новыми посредниками власти в обществе» [Diakopoulos, 2013: 2].

¹ СоциоДиггер. 2020. Т. 1. Вып. 3 «Цифровизация и искусственный интеллект». М.: ВЦИОМ. URL: https://wciom.ru/fileadmin/file/nauka/podborka/wciom_sociodigger_102020.pdf (дата обращения: 11.02.2021).

² Это утверждение справедливо по отношению к населению тех государств, где активно идут процессы компьютеризации, информатизации и цифровизации. К сожалению, более 40 % населения Земли все еще не имеют выхода в сеть, что свидетельствует о серьезной проблеме цифрового разрыва в доступе к информационно-коммуникационным технологиям и о цифровом неравенстве в целом.

В последние годы алгоритмы проникают во все сферы жизни и находятся в фокусе внимания не только ученых, но и общественности. Европейское исследование, опубликованное в начале 2019 г., показывает, что население Европы достаточно плохо знакомо с работой алгоритмов: около 48 % до сих пор не знают, что такое алгоритмы, и примерно такая же доля опрошенных не осведомлена о том, что алгоритмы уже активно используются в жизненно важных сферах — например, при принятии решений о выдаче кредита, при выборе кандидата работодателем или в медицинской диагностике. Около 46 % европейцев положительно оценивают работу алгоритмов, и только 20 % озабочены негативными последствиями алгоритмического управления и принятия решений [Grzymek, Puntschuh, 2019]. В российском обществе ситуация аналогичная. По результатам опроса ВЦИОМ, проводившегося в том же 2019 г., большинство граждан (75 %) слышали о технологии искусственного интеллекта (далее — ИИ), однако только 29 % от общего числа опрошенных смогли объяснить, что такое ИИ. Большая доля россиян положительно (48 %) или, по крайней мере, нейтрально (31 %) относятся к распространению технологий искусственного интеллекта: 68 % не опасаются замещения технологиями ИИ специалистов своей профессии, а 87 % считают, что государство должно способствовать развитию этих технологий³. Мы видим достаточно низкий уровень осведомленности россиян о современных технологиях, специфике их работы и последствиях распространения. Объяснением может служить как недостаточное освещение данного вопроса в российских СМИ и научных публикациях, так и все еще не столь интенсивное и очевидное их воздействие на общественные процессы. Кроме того, причина может быть в том, что в общественном мнении отсутствует понимание угроз и негативных последствий внедрения технологий, в том числе алгоритмов, в повседневную жизнь.

Несмотря на такой технологический оптимизм [Нестик, 2019: 13], алгоритмы, становясь важной частью общественных отношений, оказывают все большее влияние на нашу жизнь, не только делая ее комфортнее, но и создавая новые риски и угрозы. Парадоксальным образом технологический прогресс, который, как казалось когда-то, сможет ликвидировать или, по крайней мере, сгладить многие социальные проблемы, на деле их только усиливает. Так, цифровизация, датификация и алгоритмизация качественно преобразовывают существующую систему социального неравенства. Новые его формы, источником которых в том числе являются неодинаковый доступ и разные возможности использования цифровых технологий, не существуют обособленно, но «накладываются» на традиционные (например, экономическое, гендерное и др. неравенства).

Негативные последствия распространения алгоритмов необходимо изучать в перспективе общей ситуации усиления социального неравенства. Обычно влияние технологий на социальное неравенство рассматривается как проблема цифрового разрыва. В социологической литературе [Dobrinskaya, Martynenko, 2019; Добринская, Мартыненко, 2020] принято трехуровневое членение цифрового разрыва: первый связан с наличием или отсутствием доступа к цифровым технологиям, второй — результат дифференциации во владении цифровыми навыками,

³ Искусственный интеллект: угроза или возможность? // ВЦИОМ. Аналитический обзор. 2020. 27 января. URL: <https://wciom.ru/index.php?id=236&uid=10132> (дата обращения: 12.02.2021).

а третий разрыв свидетельствует о неравных возможностях и жизненных шансах как результате отсутствующего или затрудненного доступа индивидов к цифровым технологиям, а также низкого уровня цифровых навыков. Третий уровень цифрового разрыва фиксирует не только проявление социальной дифференциации, но и возникновение новой формы социального неравенства, получившего названия цифрового. Эта новая форма проявления неравенства ограничивает возможности осуществления определенной жизненной стратегии индивида, полноценного участия в жизни цифрового общества, реализации своих гражданских прав и свобод, успешного карьерного продвижения, личностного роста и расширения социальных связей и отношений, а также демократического участия. Довольно часто влияние алгоритмов на социальное неравенство описывается как следующий уровень цифрового разрыва. Некоторые исследователи предлагают рассматривать его как «алгоритмический разрыв» в контексте проявления нового цифрового неравенства [Ragnedda, 2020], другие называют его «четвертым цифровым разрывом» [Stefanija, 2017] или «новым цифровым разрывом» [Gran, Booth, Bucher, 2020].

При всем многообразии позиций по вопросу влияния алгоритмов на общество, систематических исследований о проявлении социального неравенства в эпоху алгоритмизации пока еще не так много. В данной статье делается попытка очертить контуры этой проблемы, отталкиваясь от общего понимания социального неравенства как «разницы между неравным распределением желаемых результатов жизни (таких как здоровье, счастье, успех в образовании или материальные блага) и неравным распределением возможностей (доступ к власти и жизненным шансам, которые способствуют достижению желаемых результатов)» [Carter, Reardon, 2014: 3]. Влияние алгоритмов на социальное неравенство будет рассматриваться здесь с точки зрения неравного распределения возможностей и желаемых результатов. Мы обозначим основные подходы к определению алгоритмов, а также представим те факторы и механизмы, которые, на наш взгляд, оказывают наибольшее влияние на рост социального неравенства. В отношении неравенства возможностей внимание будет уделено доступу к информации, ее поиску, проблеме социального ранжирования, а также роли цифровых навыков в алгоритмической осведомленности [Gran, Booth, Bucher, 2020]. При обсуждении неравенства желаемых результатов будут затронуты трудовая сфера, сфера политического и гражданского участия, а также дискриминация и маргинализация отдельных индивидов и социально-демографических групп. В заключении обсудим некоторые методологические вопросы, связанные с изучением социального неравенства в условиях алгоритмизации.

Что такое алгоритмы?

Алгоритмы имеют доступ к данным по финансовым операциям, банковским вкладам, страховым взносам и пенсионным счетам. Уже сегодня они могут оказывать влияние на то, кому будет сделан телефонный звонок, каковы шансы человека на пересадку жизненно важных органов, и даже могут помочь в принятии решения о выборе будущего супруга или супруги [Steiner, 2012].

В самом общем виде алгоритм представляет собой набор шагов, организованных специальным образом, чтобы обработать инструкции или данные для

достижения определенной цели. Как пишет ирландский исследователь Р. Китчин, с развитием технологий понятие алгоритма эволюционирует, отражая технический, вычислительный, математический, политический, культурный, экономический, контекстуальный, материальный, философский, этический аспекты работы алгоритмов [Kitchin, 2017: 16]. Комплексный характер их функционирования объясняет и различные варианты использования термина «алгоритм» техническими специалистами, социальными исследователями и обычными пользователями [Gillespie, 2014a].

С технологической точки зрения речь идет о специфике решения конкретной задачи. В таком контексте алгоритм представляет собой совокупность программируемых этапов, приводящих к определенному результату работы технической системы.

В социальных науках алгоритм рассматривается как сложная система, возникающая в результате взаимодействия технологии и человека. Социальная значимость алгоритмов раскрывается через описание социальных явлений и процессов, движущей силой которых являются алгоритмические системы. Эти системы включают в себя не только сами алгоритмы, но и вычислительные сети, в которых они функционируют, людей, которые их проектируют и используют, данные (и пользователей), на которых они воздействуют, а также субъектов (индивидуальных и корпоративных), которые предоставляют услуги с помощью алгоритмов. Здесь на первый план выходят вопросы ответственности за разработку и внедрение алгоритмов, а также проблема контроля над результатами их работы, зачастую непреднамеренными и заранее непрогнозируемыми [Willson, 2017]. С социологической точки зрения интерес представляет изучение особенностей коммуникации с алгоритмическими системами, а также анализ социального контекста и его роли при программировании эффективных социальных алгоритмов [Esposito, 2017: 250] в условиях искусственной социальности⁴.

Рядовые пользователи, как правило, убеждены в объективности и беспристрастности алгоритмических систем⁵. Результаты, получаемые с помощью алгоритмов, обладают иной легитимностью (нечто сравнимое с научным методом, здравым смыслом или Божьим словом [Gillespie 2014b: 168]), нежели субъективная легитимность отдельного эксперта. Доверие человека автоматизированным расчетам и статистическим данным наделяет алгоритмы особым культурным авторитетом, что приводит к распространению метрической культуры [Beer, 2016].

Алгоритмы становятся новыми социальными агентами [Esposito, 2017], выступая партнерами по коммуникации (боты, голосовые помощники и т. п.). Будучи способными создавать и переносить информацию, алгоритмы также производят искусственную коммуникацию. В качестве партнера по коммуникации алгоритм не понимает содержания, смысла сообщений и не способен к их интерпретации [Esposito, 2017]. И здесь мы снова возвращаемся к вопросу о социальной ответ-

⁴ В качестве рабочего определения искусственной социальности мы используем предложенное А. В. Резаевым и Н.Д. Трегубовой: «Искусственная социальность представляет собой эмпирический факт участия агентов ИИ в социальных взаимодействиях в качестве активных посредников или участников этих взаимодействий» [Резаев, Трегубова, 2019: 43].

⁵ Logg J. M., Minson J. A., Moore D. A. Do People Trust Algorithms More Than Companies Realize? *Harvard Business Review*. 2018. 16 October. URL: <https://hbr.org/2018/10/do-people-trust-algorithms-more-than-companies-realize> (дата обращения: 14.02.2021).

ственности за результаты работы алгоритмов в случае их возможной поломки или ошибок в принятии решений. Эффекты алгоритмов не всегда однозначны и предсказуемы. Как полагает Т. Гиллеспи, это происходит по нескольким причинам. Во-первых, алгоритмы включены в более широкую сеть отношений, которые так или иначе опосредуют и трансформируют их работу (например, в результате плохого качества исходных данных). Во-вторых, работа алгоритмов может иметь побочные эффекты и непредвиденные последствия, а отсутствие контроля или наблюдения за их работой чревато неожиданными и не всегда предсказуемыми действиями с их стороны. В-третьих, алгоритмы могут быть «предвзятыми» или допускать ошибки из-за неправильного кодирования или погрешностей при их программировании. Наконец, они могут изменяться в результате адаптации их функционирования под запросы пользователей. Поэтому алгоритмы — результат работы и программистов, и пользователей [Gillespie, 2014b].

Алгоритмы разрабатываются и внедряются для создания дополнительной ценности и увеличения капитала; для регулирования поведения индивидов и формирования их предпочтений; для определения, ранжирования и классификации индивидов по заданным критериям [Kitchin, 2017]. Именно поэтому их нельзя рассматривать как объективную и беспристрастную форму знания или как нейтрального и независимого агента действия. Это позволяет говорить о возникновении так называемого алгоритмического управления [Katzenbach, Ulbricht, 2019]. Вопрос в том, в чьих руках находится эта власть алгоритмов и кто принимает решение о ее реализации. Выделяют три группы субъектов алгоритмической власти. Во-первых, это создатели программного кода алгоритма (программисты и другие технические специалисты). Во-вторых, это заказчики, для которых специалисты создают алгоритмы (как правило, это частные компании). В-третьих, это те, кто участвует в работе алгоритма [Mau, 2019]. Можно наблюдать две противоположные тенденции в управлении: с одной стороны, алгоритмы делают его более мощным, агрессивным и всеобъемлющим — и здесь мы сталкиваемся с проблемой тотального контроля. С другой стороны, управляемые механизмы становятся более инклузивными, что позволяет обеспечить большее социальное разнообразие, а также персонализированный подход к управлению [Katzenbach, Ulbricht, 2019].

Посредством механизмов алгоритмического управления формируется иная картина мира, где реализуются новые властные и дисциплинарные возможности, создаются дополнительные механизмы принуждения и контроля. В ее основе лежат системы программного обеспечения, функционирование которых имеет глубокие социальные последствия [Kitchin, Dodge, 2011]. Алгоритмы производят культурные объекты и проявляют все большую активность в создании смыслов и интерпретации культуры, что наиболее заметно в работе поисковых систем [Roberge, Melançon, 2017]. Алгоритмы влияют на различные сферы общественной жизни, реконструируют целые отрасли, полностью изменяя их организацию и функционирование, а также и весь рынок труда (к примеру, когда в результате автоматизации происходит массовая потеря рабочих мест) [Kitchin, 2017]. При этом, как правило, отсутствует понимание того, каким образом и посредством чего алгоритмы принимают то или иное решение. Именно поэтому получил распространение термин «общество черного ящика» (black box society) [Pasquale, 2015].

Насущной социальной дилеммой становится вопрос о принятии побочных эффектов алгоритмизации, усиливающей непрозрачность многих социальных процессов. Алгоритмы также изменяют способы идентификации, отслеживания, профилирования или оценивания индивидов, осуществляя эти процедуры, как правило, без ведома пользователя в режиме реального времени. Автоматизированные алгоритмические системы фиксируют поведение человека, осуществляют скрининг его эмоций, а также оценивают состояние здоровья. Особенность социальных механизмов алгоритмических систем в том, что они, в отличие от традиционных механизмов социализации и социального контроля, невидимы и сложно определяемы в качестве участников взаимодействия, как сложно определяема и нормативная база их функционирования. Несомненно, эти процессы бросают вызов понятию «непохожести» и индивидуальности, поскольку одна из основных функций алгоритмов — категоризация и типизация поступающих к ним данных по определенным признакам. При этом возникает вопрос о возможности существования реального публичного пространства, где доступно выражение разных мнений и противоположных позиций [Broadbent, Lobet-Maris, 2015: 114].

Люди повсюду контактируют с алгоритмами, благодаря которым внедряются новые формы управления и создаются возможности для новых форм накопления капитала. Распространенность, социальные эффекты и все больший объем полномочий автономных и автоматизированных алгоритмических решений требуют углубленного изучения последствий проявления этих тенденций, а основные характеристики алгоритмов, такие как непрозрачность, непредсказуемость, нелинейность, скрытый характер воздействия и наличие сопутствующих эффектов, определяют их влияние на проблему социального неравенства.

Влияние алгоритмов на неравенство возможностей

Доступ к информации как фактор неравенства возможностей

В последние месяцы 2020 г. все чаще появлялась информация о связанных с крупными медиакомпаниями судебных разбирательствах в США. Так, представленные в недавнем докладе [Nadler, Cicilline, 2020] результаты расследования конгресса США в отношении деятельности крупнейших технологических корпораций (Amazon, Apple, Google, Facebook) показали, что эти компании занимают монопольное положение в ключевых сегментах рынка и злоупотребляют своим доминированием. Негативные последствия этого доминирования ощущают на себе обычные пользователи, для которых алгоритмы ограничивают выбор информации и снижают качество поиска, попутно собирая данные о характеристиках самих пользователей⁶.

Одним из механизмов социального неравенства в условиях алгоритмизации становятся так называемые пузыри фильтров [Паризер, 2012]. Алгоритмы сайтов поисковиков определяют, какая информация была бы интересна для конкретного пользователя, на основании данных о его месторасположении, истории

⁶ Amazon, Apple, Google и Facebook обладают монопольной властью — расследование конгресса США // Digital Russia. 2020. 7 октября. URL: <https://d-russia.ru/amazon-apple-google-i-facebook-obladajut-monopolnoj-vlastju-rassledovanie-kongressa-ssha.html> (дата обращения: 20.10.2020).

предыдущих поисковых запросов, перемещениях мыши по экрану, длительности посещения веб-страниц других сайтов и т. п. В результате пользователь получает только ту информацию, которая согласуется с его или ее прошлым опытом, а нерелевантная не отображается. И если раньше речь шла преимущественно о барьерах в доступе, связанных с физическим пространством [Вершинина, 2018], то сегодня они в большом количестве возникают и в пространстве цифровом.

Пузыри фильтров создают для пользователя персонализированную среду, состоящую только из «близкого неизвестного» — «мелких фактов из спортивной жизни или политических знаков препинания, которые на самом деле не потрясают до основания наши схемы, а лишь ощущаются как новая информация» [Паризер, 2012: 96]. Эта среда «способна понять, кто вы, что вам нравится и чего вы хотите» [там же: 236], но не дает возможности узнать то, что находится вне вашего поля зрения. Алгоритмическая персонализация в большинстве своем невидима и неподконтрольна пользователям. Например, большинство американцев (74%) уверены, что контент в социальных сетях не дает адекватного представления о том, как в обществе воспринимаются важнейшие социальные проблемы, и только четверть опрошенных убеждены в обратном. Причина состоит в модерации контента с помощью алгоритмов платформ социальных медиа⁷.

Рассмотренный ранее пример работы поисковых систем показывает, что именно от алгоритмов, которые ранжируют наиболее востребованные поисковые запросы, в конечном счете зависит то, какие источники для пользователей более важны и полезны. Поэтому результаты поисковых запросов рассматриваются не просто как «информация», но и как «социальные данные» [Lupton, 2014]. Более того, совершенствование способности алгоритмов обрабатывать текстовую информацию, которую содержат поисковые запросы, расширяет возможности для использования цифровых следов пользователей [О'Нил, 2018: 107]. Речь идет не только об ограничении поиска и доступности информации, но и о дискриминации посредством механизма ранжирования результатов поиска. В настоящее время предпринимаются попытки создать более справедливую систему функционирования поисковых запросов и представления пользователям информации. Основная задача этой технологии сводится к необходимости преодолеть «эффект Матфея» [Morik, 2020], когда наиболее популярные результаты поиска становятся еще более востребованными, а новые или непопулярные сайты фактически не отображаются у большинства пользователей. Этот эффект связан с тем, что алгоритмы нацелены на упрощение поиска и получение пользователем максимально быстрого ответа на свой запрос, но фактически это приводит к неравным возможностям не только в поиске информации, но и при позиционировании себя и своей деятельности в интернете. Несмотря на то, что эта технология только разрабатывается, уже сегодня очевидно, что она может дать новые возможности пользователям в сети. Но это не гарантирует, что не появятся новые механизмы, которые позволят каким-то группам или индивидам занимать более привилегированное положение.

⁷ Smith A. Public Attitudes Toward Computer Algorithms // Pew Research Center. 2018. 16 November. URL: <https://www.pewresearch.org/internet/2018/11/16/public-attitudes-toward-computer-algorithms/> (дата обращения: 12.02.2021).

Социальные рейтинги и системы социального ранжирования как фактор неравенства возможностей

Технологии Web 2.0 укрепили веру в возможности «цифровой демократии»: считалось, что многочисленные отзывы и комментарии пользователей способны выделить среди всего многообразия товаров и услуг лучшие, а голос каждого будет услышан. Однако функционирование алгоритмов ранжирования не полностью объективно, поскольку в них изначально заложена человеческая предвзятость [Hassan, 2020: 166].

Механизмы социального ранжирования и социальных рейтингов определяются процессом квантификации. Квантификация включает в себя распространение, разработку и внедрение новых показателей, использование различных механизмов оценивания (например, скоринговые системы), а также практик самоизмерения — селф-трекинга. Социальные отношения, социальный статус, поведение и даже эмоции оцениваются количественно. Подобная «объективность» делает весьма проблематичными любые попытки оспорить имеющиеся показатели. Получая в свое распоряжение все большие объемы данных, общество приближается к тому, чтобы стать полностью «управляемым данными», иными словами, «обществом мониторинга, контроля и оценки, которое больше не верит ни во что, если это не выражается количественно» [Mau, 2019: 24]. С социологической точки зрения количественные самоописания такого рода не просто отражают существовавшую ранее социальную реальность, а могут рассматриваться как генерирующий метод конструирования социальных различий с помощью новых механизмов (лайков, звезд, рекомендаций, отзывов, рейтингов и т. п.). Эти системы социального ранжирования, возникающие в связи с действием алгоритмов, существуют не только в цифровой среде, но и выходят за ее пределы. Как отмечают социологи, «широкое внедрение таких алгоритмов приводит к изменению реальности, в частности формированию поведения, нацеленного именно на получение высоких оценок, или неявного ограничения доступности источников информации из-за особенностей их ранжирования» [Суворова..., 2019: 105—106]. Распространение получают репутационные системы, поощряющие определенное поведение индивида, а также многочисленные рейтинговые механизмы, внедряемые на разных уровнях (рейтинг водителя такси, интернет-магазинов, отелей и др.).

В этой связи важно обратить внимание на три социально значимых аспекта квантификации социального в новом «метрическом обществе» [Mau, 2019: 6]. Во-первых, язык чисел меняет привычные представления о социальной ценности и социальном статусе. Количественные измерения, или метрики, институционализируют определенные «порядки ценности», которые создают ориентиры и обосновывают выставляемые оценки, внедряя новые нормативные принципы и установки. Этот аспект отчетливо проявляется в различных рейтингах образовательных учреждений, системе научометрических показателей, рейтингов стран, регионов, городов по различным показателям и т. п.

Во-вторых, квантификация социального стимулирует конкуренцию, поскольку наличие количественной информации усиливает тенденцию к сравнению и, как следствие, к соперничеству [Mau, 2019]. С социальной точки зрения сравнения, как правило, более эффективны и убедительны, когда они выходят за границы

субъективного восприятия и индивидуальной интерпретации и основаны на объективных критериях (или критериях, которые считаются таковыми).

В-третьих, наблюдается тенденция ко все большей социальной иерархизации, когда представления в виде таблиц, графиков, ранжированных списков или оценок способны перевести качественные различия в количественные неравенства, подвергая то, что ранее было несопоставимым, единой системе оценивания.

Важен не только сам акт квантификации, но и то, каким образом происходит оценивание, а также кто осуществляет оценку. Если традиционно ключевую роль в процессах оценки играли эксперты и институты, то сейчас любой член метрического общества может быть экспертом, поскольку любое суждение, основанное на цифрах и числовых показателях, представляется более объективным и потому сложно оспоримым. Ценность в данном контексте рассматривается как «социально производимое», а не «внутренне присущее». Именно поэтому «квантификацию» следует рассматривать как стратегию «переописания мира» с помощью языка чисел [Mau, 2019]. Единственный способ оставаться за пределами квантификации — не оставлять цифровых следов, что подразумевает добровольный выход из соответствующих контекстов цифровых коммуникаций и сетевого взаимодействия [Mau, 2020], который могут себе позволить далеко не все. Глубокая медиатизация социальной жизни [Couldry, Hepp, 2017] сформировала новую «нормальность», где индивиды по умолчанию готовы делиться большими объемами информации о себе [Neff, Nafus, 2016]. Более того, логика оптимизации, рационализации и повышения производительности, которую неолиберализм навязал практически каждому аспекту жизни, ведет к борьбе за лучшие показатели [Espeland, Stevens, 2008; Mennicken, Espeland, 2019]. Квантификация социальной жизни стимулирует появление новых форм социального неравенства, где все подвергаются сравнению, социальному ранжированию и оценке, а индивиды стремятся продемонстрировать лучшие показатели.

Алгоритмическая (не)осведомленность как фактор неравенства возможностей

Алгоритмы приводят в действие механизм «ловушки бедности». С одной стороны, они углубляют социальный разрыв, основывающийся на присущих индивидам социально-демографических (например, возраст, уровень образования, наличие социальных связей) и социально-экономических характеристиках. С другой стороны, создают новые формы неравенства, обусловленные необходимостью владеть специфическими алгоритмическими навыками, доступ к получению которых ограничен.

Возможности индивидов понимать особенности функционирования алгоритмов и социальные последствия их влияния тесно связаны с наличием или отсутствием у них цифровых навыков. Цифровой разрыв в настоящее время — результат мощных и зачастую невидимых процессов ранжирования и маргинализации, которые осуществляют алгоритмические системы. Некоторые ошибочно полагают, что алгоритмы снижают роль информационных, коммуникативных, даже стратегических навыков, а также навыков создания контента, но фактически они требуют дополнительных алгоритмических навыков. Получается, что отсутствие навыков алгоритмической осведомленности [Gran, Booth, Bucher, 2020] делает работу ал-

горитмов непрозрачной и зачастую предвзятой для пользователей. В результате цифровые медиа и интернет вещей посредством внедрения алгоритмических систем усиливают существующее неравенство [van Dijk, 2020].

Непонимание того, как работают алгоритмы, исследователи определяют как новый цифровой разрыв [*ibidem*]⁸. В этой связи предлагается пересмотр определения цифровой грамотности, которая сегодня представляет собой понимание и прогнозирование последствий постоянного подключения к сети⁹. Алгоритмические системы — своего рода «черные ящики» не только для пользователей, но даже для самих разработчиков, что как раз и объясняется отсутствием соответствующих цифровых навыков¹⁰. Поэтому на повестке дня стоит вопрос нормативного регулирования прозрачности работы алгоритмических систем. Так, в рамках недавно утвержденного в Европейском союзе «Генерального регламента ЕС о защите персональных данных» люди имеют «право на объяснение» критериев, которые алгоритмы используют в своих решениях относительно обработки персональных данных¹¹. Однако прозрачность может быть полезна только пользователям, достаточно хорошо разбирающимся в технической специфике работы алгоритмов¹².

Роль алгоритмов сегодня необходимо учитывать не только в отношении работы с поисковыми системами, но и при использовании социальных сетей, сайтов интернет-магазинов, новостных агрегаторов и других сервисов. Те, кто понимает, что алгоритмы определяют, какой контент они видят, могут как скорректировать свои ожидания, так и использовать стратегии поиска контента таким образом, чтобы обойти ограничения, налагаемые платформами и алгоритмическими программами. Те же, кому не хватает специальных знаний и понимания роли алгоритмов, больше всего подвергаются манипуляциям и зависят от того, какие сайты доступны им в первую очередь [Hargittai, Micheli, 2019]. Следовательно, их картину мира формируют алгоритмы. Все чаще социологи подчеркивают, что для работы с алгоритмами необходимы специфические «алгоритмические навыки» и обладание алгоритмической осведомленностью, которые позволили бы выявлять последствия их воздействия.

Таким образом, основными механизмами, усиливающими социальное неравенство в условиях алгоритмизации, становятся пузыри фильтров, системы социального ранжирования и социальных рейтингов, а также неравная алгоритмическая осведомленность. Эти механизмы влияют на возможности отдельных индивидов

⁸ См. также Susarla A. The New Digital Divide Is Between People Who Opt Out of Algorithms and People Who Don't. *Fast Company*. 2019. 18 April. URL: <https://www.fastcompany.com/90336381/the-new-digital-divide-is-between-people-who-opt-out-of-algorithms-and-people-who-dont> (дата обращения: 13.02.2021).

⁹ Ibidem, см. также: Rainie L., Anderson J. Code Dependent: Pros and Cons of the Algorithm Age // Pew Research Center. 2017. 8 February. URL: <https://www.pewresearch.org/internet/2017/02/08/code-dependent-pros-and-cons-of-the-algorithm-age/> (дата обращения: 12.02.2021).

¹⁰ Susarla A. The New Digital Divide Is Between People Who Opt Out of Algorithms and People Who Don't. *Fast Company*. 2019. 18 April. URL: <https://www.fastcompany.com/90336381/the-new-digital-divide-is-between-people-who-opt-out-of-algorithms-and-people-who-dont> (дата обращения: 13.02.2021).

¹¹ Генеральный регламент о защите персональных данных // Европейская служба внешних связей. 2019. 23 мая. URL: https://eeas.europa.eu/headquarters/headquarters-homepage_ru/45073/ (дата обращения: 12.02.2021).

¹² Susarla A. The New Digital Divide Is Between People Who Opt Out of Algorithms and People Who Don't. *Fast Company*. 2019. 18 April. URL: <https://www.fastcompany.com/90336381/the-new-digital-divide-is-between-people-who-opt-out-of-algorithms-and-people-who-dont> (дата обращения: 13.02.2021).

и целых социально-демографических групп получать выгоды и преимущества в различных сферах общественной жизни. Действие этих механизмов определяется рядом социально-демографических характеристик, среди которых возраст (у старших возрастных групп обычно ниже алгоритмическая осведомленность и уровень владения цифровыми навыками), образование (в зависимости от уровня и специализации), место жительства (развитость инфраструктуры в городе и сельской местности, а также разных регионах и странах), специфика профессиональной деятельности и др. Способность эффективно использовать эти механизмы создает возможности карьерного роста, экономического благополучия, реализации гражданских и политических прав и свобод, а также накопления социального капитала.

Роль алгоритмов в неравенстве социальных результатов

В настоящее время алгоритмы не только существенно ограничивают возможности, но и оказывают влияние на распределение конкретных желаемых результатов, затрагивая фактически все сферы общественной жизни. В качестве неравенства желаемых социальных результатов производится анализ проблем неравного доступа к образованию, рабочим местам, услугам, а также гражданскому и политическому участию.

В последние годы наиболее заметным становится использование алгоритмов в экономической сфере. Так, функционирование алгоритмов сегодня включено во многие аспекты трудовой деятельности: от подбора персонала до контроля на рабочем месте. Для коммерческих организаций использование алгоритмов создает возможности извлечения прибыли посредством автоматизации, оптимизации и рационализации бизнес-процессов. Зачастую наемные работники попадают в уязвимое положение, которое связано с возможной потерей рабочих мест, усилением контроля и предвзятого отношения со стороны работодателя.

Сегодня многие компании активно используют алгоритмы в процессе подбора персонала, особенно на этапе первичного собеседования с кандидатом. Согласно данным опроса компании HeadHunter, проведившемся в 2018 г., около трети представителей российских предприятий убеждены в изменении роли HR-служб в связи с распространением технологий ИИ и роботизации. При этом более 60 % респондентов предполагают, что к 2050 г. подбор персонала будут осуществлять роботы, а 47 % опрошенных заявили, что ИИ займется привлечением соискателей¹³. При использовании алгоритмов для подбора персонала возникает вопрос о допустимости подобного отбора. Согласно одному из свойств алгоритмов, они с необходимостью упрощают все многообразие атрибутов, присущих человеку.

В зарубежной практике периодически возникают ситуации непреднамеренного расизма или национализма, когда соискателю отказывают в рабочем месте из-за особенностей его фамилии (К. О'Нил описывает на примере США случаи исключения из списка соискателей всех претендентов латиноамериканского происхождения вне зависимости от их образования или квалификации [О'Нил, 2018]). Выявить причины отказов фактически невозможно из-за непрозрачности алго-

¹³ Как искусственный интеллект изменит HR? Презентация Headhunter. 2018. URL: <https://hhcdn.ru/file/16676170.pdf> (дата обращения: 12.02.2021).

ритмов. Коммерческие организации не открывают программные коды, объясняя это уязвимостью своих алгоритмических систем перед манипуляциями со стороны хакеров, конкурентов и т. п. Пользователи же, которые знают об их существовании, могут применять стратегии, направленные на устранение некоторых ограничений, создаваемых этими системами.

Исследовательский центр Пью (США) в отчете о проведенном в 2017 г. опросе о влиянии алгоритмов показывает серьезную обеспокоенность людей по поводу использования алгоритмов для конкретных целей. 57 % респондентов в США высказали такие же опасения по поводу автоматизированного просмотра резюме кадровыми агентствами и отделов по набору персонала внутри организаций¹⁴. По данным HeadHunter, в России каждая третья компания использует автоматизированные системы по подбору персонала. При этом лишь 8 % респондентов считают, что алгоритм может принимать окончательное решение о найме сотрудника¹⁵. Неравные возможности доступа к информации об особенностях функционирования алгоритмов приводят к неравенству социальных позиций, например работодателя и работника, крупных компаний и субъектов малого предпринимательства и др.

Вся повседневная деятельность человека в цифровой среде становится для технологических гигантов источником важнейшей информации, которая в дальнейшем не только подвергается коммерциализации, но и позволяет влиять на большие группы людей, корректируя их вкусы, интересы и убеждения [Hassan, 2020]. Однако сбор этих поведенческих данных, их анализ и использование доступны далеко не всем. Как следствие, в условиях новой искусственной социальности в привилегированном положении оказываются обладающие соответствующими техническими или материальными возможностями.

Неравенство возможностей преобразуется в экономической сфере в конкретные проявления неравного распределения желаемых социальных результатов: дохода, доступа к рабочим местам и др. Система социального ранжирования и создания рейтингов постепенно монополизирует высокий статус одних продавцов товаров и услуг, вытесняя остальных и оставляя им лишь незначительную долю покупателей. Агентам малого бизнеса не удается преодолеть пузыри фильтров при продвижении товара или услуги. Тем самым алгоритмы воспроизводят существующие отношения социального неравенства и создают новые барьеры для наименее защищенных групп населения.

Другой сферой, в функционирование которой активно внедряются алгоритмы, стало пространство гражданского участия. Здесь действия алгоритмов подрывают представления о демократических основах цифрового пространства, серьезно ограничивая информационное поле пользователей. Свобода слова и самовыражения находится под контролем алгоритмов, часто дискриминирующих пользователей по расовому и гендерному принципу [Noble, 2018]. Модерирование

¹⁴ Rainie L., Anderson J. Code Dependent: Pros and Cons of the Algorithm Age // Pew Research Center. 2017. 8 February. URL: <https://www.pewresearch.org/internet/2017/02/08/code-dependent-pros-and-cons-of-the-algorithm-age/> (дата обращения: 12.02.2021).

¹⁵ «Не так страшен черт, как его малют», или что сейчас происходит с автоматизацией подбора персонала? Презентация HeadHunter. 2019. URL: <https://hhcdn.ru/file/16861694.pdf> (дата обращения: 12.02.2021).

контента, в котором обвиняют сегодня крупные медиакомпании¹⁶, приводит к тому, что получение полной картины социальной реальности становится невозможным. Компании, разрабатывающие алгоритмы, фактически становятся своего рода «привратниками» (*«gatekeepers»*), сортирующими и отбирающими специфическую информацию.

В некотором смысле алгоритмы начинают выступать как «этикopolитические существа», поскольку отдают предпочтение одним свойствам перед другими и переопределяют границы добра и зла, хорошего и плохого [Amoore 2020: 5]. При этом с учетом возможности алгоритма к изменению возникает необходимость постоянного мониторинга их деятельности для своевременного выявления ошибок, а также подробного анализа того, как именно они учатся действовать, что лежит в основе их решений и как формируется их представление о мире на основе полученных ими данных.

Следует рассмотреть и то, как алгоритмы становятся источником дискриминации и маргинализации отдельных социально-демографических групп. Эти процессы базируются на цифровых следах пользователей и их атрибутивных характеристиках. В первом случае непрерывный сбор информации позволяет алгоритмам принимать персонализированные решения о стоимости товаров, услуг и др. Их результатом уже сегодня становится дифференцированная система страхования, включающая, помимо стандартных, большое количество дополнительных критериев, на основе которых используется индивидуализированное формирование стоимости страховки и процентной ставки по кредиту.

Во втором случае речь идет о ситуациях, когда отношение к индивиду выстраивается на основе, например, его/ее местоположения. Так, в США проживание в неблагополучном районе приводит к тому, что алгоритм требует от полицейского более тщательного досмотра, а от судьи — вынесения более серьезного наказания даже за незначительные правонарушения. В то же время жители элитных районов фактически остаются вне поля зрения. «В результате мы криминализировали бедность, полагая, что инструменты, которые мы при этом использовали, — не только научные, но и справедливые» [О'Нил, 2018: 126]. Для обозначения совокупности всех этих последствий используется понятие «алгоритмическая дискриминация» [Winter, 2015]. В связи с утратой анонимности в интернет-пространстве пользователи регулярно сталкиваются с проявлениями дискриминации, но ее очень сложно выявить. В этом отношении действия алгоритмов иногда лишь отражают социальную реальность, когда речь идет о расовой, гендерной или национальной дискриминации.

Постепенно алгоритмы внедряются и в систему образования, затрагивая следующие аспекты: поступление в учебные заведения, проведение экзаменов учащихся и аттестации учителей. В августе 2020 г. широко обсуждался скандал в британских школах, связанный с использованием алгоритмов для подсчета итоговых оценок. Предполагалось, что с помощью набора критериев, заложенных в алгоритм оценки, будет решена задача максимальной точности и объективности

¹⁶ Amazon, Apple, Google и Facebook обладают монопольной властью — расследование конгресса США // Digital Russia. 2020. 7 октября. URL: <https://d-russia.ru/amazon-apple-google-i-facebook-obladajut-monopolnoj-vlastju-rassledovanie-kongressa-ssha.html> (дата обращения: 14.02.2021).

оценивания. Однако особенности работы этих алгоритмов привели к тому, что около 40% школьников посчитали свои оценки заниженными, что спровоцировало массовые выступления против результатов экзаменов. Связано это было с тем, что при выставлении оценок учитывался в том числе средний уровень успеваемости в школе и ее образовательный рейтинг¹⁷.

Кроме того, алгоритмы используются и для составления рейтингов абитуриентов высших учебных заведений. Эти системы учитывают не только объективные данные о школьной успеваемости или достижениях в спорте, но и на основе других критериев определяют вероятность поступления. Доступ к информации об этих дополнительных критериях становится важным ресурсом для поступления. В итоге «образовательная система в основном обслуживает привилегированные слои» [О'Нил, 2018: 98].

Рассмотренные выше проявления неравенства возможностей (ранжирование и система рейтингов, неравный доступ к информации и алгоритмической осведомленности) становятся механизмами воспроизведения социального неравенства, которые встречаются сегодня фактически во всех сферах общественной жизни. Алгоритмы используются сегодня для принятия решений полицией, вынесения судебных решений, при трудоустройстве, приеме в учебные заведения, при предоставлении различных социальных услуг [Eubanks, 2018; Noble, 2018; О'Нил, 2018].

Алгоритмы и социальное неравенство: новые направления научного поиска

Проведенный анализ позволяет сделать ряд выводов и поставить новые вопросы для исследователей неравенства. Налицо усиление влияния технологий на проблему цифрового неравенства. Схема трехуровневого членения цифрового разрыва не способна в полной мере отразить и учесть: 1) все технологическое разнообразие современных средств, обеспечивающих цифровое присутствие; 2) репертуар цифровых навыков и компетенций, которые варьируются в зависимости от целей использования цифровых технологий; 3) все потенциальные возможности и жизненные шансы, отражающие включение в процессы цифровизации через доступ и навыки применения.

Как следствие работы пузырей фильтров, социального ранжирования и неравных возможностей повышения алгоритмической осведомленности, возникают новые проявления дискриминации и маргинализации во многих сферах, ограничиваются возможности политического и гражданского участия, эффективной экономической деятельности и т. п. Автоматизированное принятие решений алгоритмическими системами обычно благоприятствует людям и коллективам, которые уже находятся в привилегированном положении, дискриминируя маргинализированных людей [Noble, 2018], что действительно представляет собой серьезную проблему, с которой в процессе растущей автоматизации общества необходимо бороться. Тем не менее это явление нельзя назвать новым, и не только алгоритм в нем виноват — предвзятые наборы данных и алгоритмические правила прини-

¹⁷ Алгоритм не сработал: как школьники добились отмены несправедливой системы оценок. *Коммерсант United Kingdom*. 2020. 17 августа. URL: <https://www.kommersant.uk/articles/algoritm-ne-srabotal-kak-shkolniki-dobilis-otmeny-nespravedlivoy-sistemy-otsenok> (дата обращения: 12.02.2021).

тия решений также порождают дискриминацию. Скорее, можно говорить о том, что любая технологическая и бюрократическая процедура материализует такие классификации, как пол, социальный класс, географическое пространство, раса. Они не берут свое начало в этих системах, они просто отражают преобладающие в обществе предрассудки и стереотипы [Katzenbach, Ulbricht, 2019].

Возникает вопрос, является ли все то, что было описано выше, новым, четвертым уровнем цифрового разрыва? Будучи непрозрачными, не поддающимися контролю и пониманию, алгоритмы управляют нашей жизнью, составляя профили пользователей, собирая данные, принимая решения о том, предоставить ли доступ или нет к информации, услугам и т. п. [Stefanija, 2017]. Создают ли пузыри фильтров, механизмы социальных рейтингов и ранжирования алгоритмический разрыв?

Безусловно, технологии делают проблему социального неравенства в целом и цифрового неравенства в частности более комплексной. Однако будет ли обоснованным и методологически верным вводить новые понятия для каждой новой технологии, учитывая скорость их разработки и внедрения? Одним из выходов могла бы стать продуктивная кодификация уже имеющихся понятий и категорий по проблематике цифрового разрыва и цифрового неравенства.

Попытка решения данной методологической проблемы представлена в концепции стека цифрового неравенства [Robinson et al., 2020a]. Метафора стека (термин, используемый в программировании и обозначающий многоуровневое взаимосвязанное представление данных), по мнению авторов концепции, способна отразить всю сложность множественных проявлений цифровых неравенств на разных уровнях. В результате конструируется стек цифрового неравенства, где каждый слой — доступа (1), навыков (2) и использования (3) содержит ряд компонентов, одновременно влияя и находясь под влиянием других слоев стека. Например, доступ как базовый слой всего стека объединяет подключенность к сети, доступ к цифровым устройствам и программному обеспечению, а каждый из этих компонентов, в свою очередь, охватывает множество других показателей (например, подключение рассматривается с точки зрения качества, продолжительности, устойчивости соединения и т. д.).

Авторы подчеркивают комплексный характер стека, включая в него и влияние множества социально-демографических характеристик. Но наиболее важна возможность данной концепции отразить масштабы влияния технологий на воспроизведение социальных различий. В результате социальная поляризация постоянно воспроизводится и увеличивается [Осипова, 2019; Полякова, 2018]. На примере функционирования платформ, больших данных, алгоритмов, автоматизации, киберпреступности и т. д. концептуальная схема стека цифрового неравенства фиксирует возникновение новых его проявлений. Некоторые из них свидетельствуют о радикальных сокращениях роли отдельных пользователей и усилиении власти технологических гигантов и создателей технологий как наиболее сильных социальных акторов [Robinson et al., 2020b]. Тем не менее сильные стороны концепции стека — комплексность и всеохватывающий характер — превращаются в основные ее недостатки: применить эту концепцию на практике представляется весьма затруднительным в силу необходимости одновременного учета большого числа переменных на разных уровнях анализа (от отдельного индивида до обще-

ства в целом). Поэтому вопрос о методологии изучения социального неравенства в условиях алгоритмизации остается открытым.

В заключение обратим внимание и на другой аспект проявления цифрового неравенства в эпоху алгоритмов. Возможно, отказ от алгоритмического вмешательства станет роскошью и символом изобилия, доступной лишь немногим избранным¹⁸. Проблема состоит в том, насколько серьезные последствия ожидают людей, которые находятся не на той стороне цифрового разрыва.

Список литературы (References)

- Вершинина И. А. Репрезентация власти в городском пространстве: концепция Йорана Терборна // Вестник Российской университета дружбы народов. Серия Социология. 2018. Т. 18. № 2. С. 226—237. <https://doi.org/10.22363/2313-2272-2018-18-2-226-237>.
- Vershinina I. A. (2018) Representation of Power in the Urban Space: Göran Therborn's Theory. *RUDN Journal of Sociology*. Vol. 18. No. 2. P.226—237. <https://doi.org/10.22363/2313-2272-2018-18-2-226-237>. (In Russ.)
- Добринская Д. Е., Мартыненко Т. С. Возможно ли цифровое равенство? (о книге Я. ван Дейка «Цифровой разрыв») // Социологические исследования. 2020. № 10. С. 158—164. <https://doi.org/10.31857/S013216250009459-7>.
- Dobrinskaya D. E., Martynenko T. S. (2020) Is Digital Equality Possible? (On the Book "The Digital Divide" by J. van Dijk). *Sociological Studies*. No. 10. P. 158—164. <https://doi.org/10.31857/S013216250009459-7>. (In Russ.)
- Нестик Т. А. Перспективы психологических исследований внедрения технологий искусственного интеллекта // Цифровое общество в культурно-исторической парадигме / под ред. Марцинковской Т. Д., Орестовой В. Р., Гавриченко В. Р. М.: МПГУ, 2019. С. 13—22.
- Nestik T. A. (2019) Perspectives of Psychological Research of Artificial Intelligence Technologies Implementation. In: Marzinkovskaya T. D., Orestova V. R., Gavrichenko V. R. (eds.) *Digital Society in Cultural-Historical Paradigm*. Moscow: MPGУ. P. 13—22. (In Russ.).
- О'Нил К. Убийственные большие данные. Как математика превратилась в оружие массового поражения. М.: ACT, 2018.
- O'Neil C. (2018) Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. Moscow: AST. (In Russ.)
- Осипова Н. Г. Социальное неравенство в современном мире // Вестник Московского университета. Серия 18. Социология и политология. 2019. Т. 25. № 4. С. 124—153. <https://doi.org/10.24290/1029-3736-2019-25-4-124-153>.
- Osipova N. G. (2019) Social Inequality in the Modern World. *Moscow State University Bulletin. Series 18. Sociology and Political Science*. Vol. 25. No. 4. P. 124—153. <https://doi.org/10.24290/1029-3736-2019-25-4-124-153>. (In Russ.)

¹⁸ Susarla A. The New Digital Divide Is Between People Who Opt Out of Algorithms and People Who Don't. *Fast Company*. 2019. 18 April. URL: <https://www.fastcompany.com/90336381/the-new-digital-divide-is-between-people-who-opt-out-of-algorithms-and-people-who-dont> (дата обращения: 13.02.2021).

Паризер Э. За стеной фильтров. Что Интернет скрывает от вас? М.: Альпина Бизнес Букс, 2012.

Pariser E. (2012) The Filter Bubble: What the Internet is Hiding from You. Moscow: Alpina Business Books. (In Russ.)

Полякова Н. Л. Оформление социального неравенства в практиках повседневности: историческая перспектива // Вестник Московского университета. Серия 18. Социология и политология. 2018. Т. 24. № 4. С. 7—25. <https://doi.org/10.24290/1029-3736-2018-24-4-7-25>.

Polyakova N. L. (2018) The Formation of Social Inequality in the Everyday Practices: The Historical Perspective. *Moscow State University Bulletin. Series 18. Sociology and Political Science*. Vol. 24. No. 4. P. 7—25. <https://doi.org/10.24290/1029-3736-2018-24-4-7-25>. (In Russ.)

Резаев А. В., Стариakov В. С., Трегубова Н. Д. Социология в эпоху «искусственной социальности»: поиск новых оснований // Социологические исследования. 2020. № 2. С. 3—12. <https://doi.org/10.31857/S013216250008489-0>.

Rezaev A. V., Starikov V. S., Tregubova N. D. (2020) Sociology in the Age of “Artificial Socialit”’: Search of New Bases. *Sociological Studies*. No. 2. P. 3—12. <https://doi.org/10.31857/S013216250008489-0>. (In Russ.)

Резаев А. В., Трегубова Н. В. «Искусственный интеллект», «онлайн-культура», «искусственная социальность»: определение понятий // Мониторинг общественного мнения: экономические и социальные перемены. 2019. № 6. С. 35—47. <https://doi.org/10.14515/monitoring.2019.6.03>.

Rezaev A. V., Tregubova N. D. (2019) Artificial Intelligence, On-line Culture, Artificial Sociality: Definition of the Terms. *Monitoring of Public Opinion: Economic and Social Changes Journal*. No. 6. P. 35—47. <https://doi.org/10.14515/monitoring.2019.6.03>. (In Russ.)

Суворова А. В., Бахитова А. А., Кузнецова А. Д., Гуляев П. Р. Социальные алгоритмы онлайн-сообществ: аналитический обзор // Информационное Общество. № 4—5. 2019. С. 105—112. URL: <http://infosoc.iis.ru/article/view/109> (дата обращения: 12.02.2021)

Suvorova A. V., Bakhitova A. A., Kuznetsova A. D., Gulyaev P. R. (2019) Social Algorithms for Online Communities: An Analytical Review. *Information Society*. No. 4—5. P. 105—112. URL: <http://infosoc.iis.ru/article/view/109> (accessed: 12.02.2021). (In Russ.)

Amoore L. (2020) Cloud Ethics: Algorithms and the Attributes of Ourselves and Others. Durham: Duke University Press. <https://doi.org/10.2307/j.ctv11g97wm>.

Beer D. (2016) Metric Power. London: Palgrave Macmillan. <https://doi.org/10.1057/978-1-37-55649-3>.

Broadbent S., Lobet-Marais C. (2015) Towards a Grey Ecology. In: Floridi L. (ed.) *The Onlife Manifesto*. Cham: Springer International Publishing. P. 111—124. https://doi.org/10.1007/978-3-319-04093-6_15.

Carter P. L., Reardon S. F. (2014) Inequality Matters. *William T. Grant Foundation Inequality Paper*. URL: <https://ed.stanford.edu/sites/default/files/inequalitymatters.pdf>. (accessed: 13.02.2021).

Diakopoulos N. (2013) Algorithmic Accountability Reporting: on the Investigation of Black Boxes. Columbia Journalism School. Tow Center for Digital Journalism. <https://doi.org/10.7916/D8ZK5TW2>.

Dobrinskaya D. E., Martynenko T. S. (2019) Defining the Digital Divide in Russia: Key Features and Trends. *Monitoring of Public Opinion: Economic and Social Changes Journal*. No. 5. P. 100—119. <https://doi.org/10.14515/monitoring.2019.5.06>.

Espeland W. N., Stevens M. L. (2008) A Sociology of Quantification. *European Journal of Sociology*. Vol. 49. No. 3. P. 401—436. <https://doi.org/10.1017/S0003975609000150>.

Esposito E. (2017) Artificial Communication? The Production of Contingency by Algorithms. *Zeitschrift für Soziologie*. Vol. 46. No. 4. P. 249—265. <https://doi.org/10.1515/zfsoz-2017-1014>.

Gillespie T. (2014a) Algorithm [draft] [#digitalkeyword]. *Culture Digitally*. URL: <http://culturedigitally.org/2014/06/algorithm-draft-digitalkeyword/> (accessed: 12.02.2021).

Gillespie T. (2014b) The Relevance of Algorithms. In: Gillespie T., Boczkowski P. J., Foot K. A. (eds.) *Media Technologies: Essays on Communication, Materiality, and Society*. Cambridge, MA: MIT Press. P. 167—194. <https://doi.org/10.7551/mitpress/9780262525374.003.0009>.

Gran A. B., Booth P., Bucher T. (2020) To Be or Not to Be Algorithm Aware: A Question of a New Digital Divide? *Information, Communication & Society*. <https://doi.org/10.1080/1369118X.2020.1736124>.

Grzymek V., Puntschuh M. (2019) What Europe Knows and Thinks about Algorithms. Results of a Representation Survey. *Discussion Paper Ethics of Algorithms*. No. 10. P. 1—37. URL: <http://aei.pitt.edu/102582/1/WhatEuropeKnowsAndThinkAboutAlgorithm.pdf> (accessed: 14.02.2021).

Hargittai E., Micheli M. (2019) Internet Skills and Why They Matter. In: Graham M., Dutton W. H. (eds.) *Society and the Internet: How Networks of Information and Communication Are Changing Our Lives*. Oxford: Oxford University Press. P. 109—124. <https://doi.org/10.1093/oso/9780198843498.003.0007>.

Hassan R. (2020) The Condition of Digitality: A Post-Modern Marxism for the Practice of Digital Life. London: University of Westminster Press.

Katzenbach C., Ulbricht, L. (2019) Algorithmic Governance. *Internet Policy Review*. Vol. 8. No. 4. P. 1—18. <https://doi.org/10.14763/2019.4.1424>.

Kitchin R. (2017) Thinking Critically About and Researching Algorithms. *Information, Communication & Society*. Vol. 20. No. 1. P. 14—29. <https://doi.org/10.1080/1369118X.2016.1154087>.

- Kitchin R., Dodge M. (2011) *Code/Space: Software and Everyday Life*. Cambridge, MA: MIT Press.
- Lupton D. (2014) *Digital Sociology*. Abingdon, Oxon.: Routledge.
- Mau S. (2019) *The Metric Society on the Quantification of the Social*. Cambridge; Medford, MA: Polity Press.
- Mau S. (2020) Numbers matter! The society of indicators, scores and ratings. *International Studies in Sociology of Education*. Vol. 29. No. 1—2. P. 19—37. <https://doi.org/10.1080/09620214.2019.1668287>.
- Mennicken A., Espeland W. N. (2019) What's New With Numbers? Sociological Approaches to the Study of Quantification. *Annual Review of Sociology*. Vol. 45. No. 1. P. 223—245. <https://doi.org/10.1146/annurev-soc-073117-041343>.
- Morik M., Singh A., Hong J., Joachims T. (2020) Controlling Fairness and Bias in Dynamic Learning-to-Rank. *SIGIR'20: Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*. July 2020. P. 429—438. <https://doi.org/10.1145/3397271.3401100>.
- Neff G., Nafus D. (2016) *Self-Tracking*. Cambridge, MA: MIT Press.
- Noble S. U. (2018) *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York, NY: New York University Press.
- Pasquale F. (2015) *The Black Box Society*. Cambridge, MA: Harvard University Press.
- Ragnedda M. (2020) *Enhancing Digital Equity. Connecting the Digital Underclass*. Cham: Palgrave Macmillan.
- Roberge J., Melançon L. (2017) Being the King Kong of Algorithmic Culture Is a Tough Job After All: Google's Regimes of Justification and the Meanings of Glass. *Convergence*. Vol. 23. No. 3. P. 306—324. <https://doi.org/10.1177/1354856515592506>.
- Robinson L., Schulz J., Blank G., Ragnedda M., Ono H., Hogan B., Mesch G. S., Cotten S. R., Kretchmer S. B., Hale T. M., Drabowicz T., Yan P., Wellman B., Harper M.-G., Quan-Haase A., Dunn H. S., Casilli A. A., Tubaro P., Carvath R., Chen W., Wiest J. B., Dodel M., Stern M. J., Ball C., Huang K.-T., Khilnani A. (2020a) Digital inequalities 2.0: Legacy inequalities in the information age. *First Monday*. Vol. 25. No. 7. <https://doi.org/10.5210/fm.v25i7.10842>.
- Robinson L., Schulz J., Dunn H. S., Casilli A. A., Tubaro P., Carvath R., Chen W., Wiest J. B., Dodel M., Stern M. J., Ball C., Huang K.-T., Blank G., Ragnedda M., Ono H., Hogan B., Mesch G. S., Cotten S. R., Kretchmer S. B., Hale T. M., Drabowicz T., Yan P., Wellman B., Harper M.-G., Quan-Haase A., Khilnani A. (2020b) Digital Inequalities 3.0: Emergent Inequalities in the Information Age. *First Monday*. Vol. 25. No. 7. <https://doi.org/10.5210/fm.v25i7.10844>.
- Stefanija A. P. (2017) The Fourth Digital Divide-Divide by Algorithm? On the Algorithms-Mediated Society and What to Do Next. URL: <http://popstefanija.com/the-fourth-digital-divide-divide-by-algorithm/> (accessed: 12.02.2021).

Steiner C. (2012) Automate This: How Algorithms Took Over Our Markets, Our Jobs, and the World. New York, NY: Portfolio/Penguin.

Nadler J., Cicilline D. N. (2020) Investigation of Competition in Digital Markets: Majority Staff Report and Recommendations. Washington, DC: Subcommittee on Antitrust, Commercial and Administrative Law of the Committee on the Judiciary. URL: https://judiciary.house.gov/uploadedfiles/competition_in_digital_markets.pdf?utm_campaign=4493-519 (accessed: 13.02.2021).

van Dijk J. (2020) The Digital Divide. Cambridge, Medford, MA: Polity Press.

Willson M. (2017) Algorithms (and the) Everyday. *Information, Communication and Society*. Vol. 20. No. 1. P. 137—150. <https://doi.org/10.1080/1369118X.2016.1200645>.

Winter J. (2015) Algorithmic Discrimination: Big Data Analytics and the Future of the Internet. In: Winter J., Ono R. (eds.) *The Future Internet. Public Administration and Information Technology*. Vol. 17. Cham: Springer International Publishing. P. 125—140. https://doi.org/10.1007/978-3-319-22994-2_8.

DOI: [10.14515/monitoring.2021.1.1808](https://doi.org/10.14515/monitoring.2021.1.1808)



А. М. Корбут

ОДОМАШНИВАНИЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА: УМНЫЕ КОЛОНКИ И ТРАНСФОРМАЦИЯ ПОВСЕДНЕВНОЙ ЖИЗНИ

Правильная ссылка на статью:

Корбут А. М. Одомашнивание искусственного интеллекта: умные колонки и трансформация повседневной жизни // Мониторинг общественного мнения: экономические и социальные перемены. 2021. № 1. С. 193—216. <https://doi.org/10.14515/monitoring.2021.1.1808>.

For citation:

Korbut A. M. (2021) Domestication of Artificial Intelligence: Smart Speakers and Transformation of Everyday Life. *Monitoring of Public Opinion: Economic and Social Changes*. No. 1. P. 193–216. <https://doi.org/10.14515/monitoring.2021.1.1808>. (In Russ.)

ОДОМАШНИВАНИЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА: УМНЫЕ КОЛОНКИ И ТРАНСФОРМАЦИЯ ПОВСЕДНЕВНОЙ ЖИЗНИ

КОРБУТ Андрей Михайлович — кандидат социологических наук, старший научный сотрудник Центра фундаментальной социологии, Национальный исследовательский университет «Высшая школа экономики», Москва, Россия
E-MAIL: akorbut@hse.ru
<https://orcid.org/0000-0003-3835-9863>

Аннотация. Появившись относительно недавно, в 2014 г., умные колонки (портативные колонки с интегрированным голосовым помощником) уже стали для многих пользователей привычной частью домашней обстановки. В данной работе рассматривается, какие трансформации происходят в социальных практиках в результате появления в домохозяйствах умных колонок. В статье анализируются два аспекта одомашнивания искусственного интеллекта при использовании умных колонок: 1) включение умных устройств в повседневную домашнюю культуру, 2) формы и способы взаимодействия с умными устройствами. Преобразование домашней социальности, к которому приводит появление умных колонок, заключается, с одной стороны, в изменении рутинных домашних практик (приготовления еды, уборки, прослушивания музыки, установки будильника, выяснения прогноза погоды и т.д.). В некоторых случаях умные колонки меняют только фон деятельности, но в других случаях трансформируют способ и частоту ее осуществления. С другой стороны, появление умных колонок приводит к новым формам коммуникации, которые ра-

DOMESTICATION OF ARTIFICIAL INTELLIGENCE: SMART SPEAKERS AND TRANSFORMATION OF EVERYDAY LIFE

*Andrei M. KORBUT¹—Cand. Sci. (Soc.), Senior Researcher at the Center for Fundamental Sociology
E-MAIL: akorbut@hse.ru
<https://orcid.org/0000-0003-3835-9863>*

¹ National Research University Higher School of Economics, Moscow, Russia

Abstract. Being a relatively recent invention (they appeared in 2014), smart speakers (speakers with an integrated voice assistant) have already become part of many users' home environment. This paper examines what kind of transformations occur in social practices due to the introduction of smart speakers to the domestic life. The article analyses two aspects of the 'domestication' of artificial intelligence when using smart speakers: (1) involvement of intelligent devices in everyday home culture and (2) forms and ways of interacting with smart devices. Transformation of home sociality that the introduction of smart speakers leads to, consists, on the one hand, in changing routine domestic practices (cooking, cleaning, music listening, setting an alarm clock, checking the weather report, etc.). In some cases, smart speakers only change the activity's background, but they transform the same way and frequency of activities in other cases. On the other hand, smart speakers' introduction leads to new forms of communication that were previously unusual in domestic life. For example, users increasingly shift to loud speech or avoid specific phrases and expressions to protect privacy. Relying on

нее были непривычны для домашней жизни. Например, пользователи чаще используют громкую речь, избегают определенных фраз и выражений для сохранения приватности или начинают фразы с обращения к собеседнику.

В статье с опорой на данные существующих исследований, автоэтнографические наблюдения автора и анализ фрагментов разговоров с умной колонкой показывается, что, несмотря на стремление разработчиков умных колонок сделать их использование «естественному», подобные разговаривающие устройства требуют от пользователей не просто применения к искусственному интеллекту привычного способа взаимодействия, а выработки новых форм взаимодействия и социальности.

Ключевые слова: умные колонки, социология повседневности, дом, искусственный интеллект, этнометодология

Благодарность. Статья подготовлена в сотрудничестве с Лабораторией геймификации Сбербанка.

the existing research data and his own auto-ethnographic observations and analysis of parts of conversations with the smart speaker, the author shows that, despite the inventor's desire to make the use of smart speakers 'natural', users do need to develop new forms of interaction and sociality to interact with AI.

Keywords: smart speakers, sociology of everyday life, home, artificial intelligence, ethnometodology

Acknowledgments. The article is prepared in collaboration with the SberBank Gamification Lab.

Введение

Сегодня многие цифровые технологии основываются на искусственном интеллекте (или машинном обучении, как предпочитают говорить специалисты, чтобы не создавать завышенных ожиданий относительно «интеллектуальности» подобных систем¹). С социологической точки зрения искусственный интеллект интересен как элемент культурного воображаемого, то есть в контексте того, как различные социальные группы представляют себе сущность искусственного интеллекта, его возможности и ограничения, как он изображается в медиа и становится элементом дискурсивных практик (см., например, [Cave, Dihal, Dillon, 2020]). Все это важные элементы социальных процессов, однако у искусственного интеллекта есть и более прагматическая сторона: чтобы оказывать то воздействие, которое от них ожидают разработчики, технологии искусственного интеллекта должны «оповедневничаться», рутинизироваться, становиться знакомой и привычной частью обыденных ситуаций, в которых ими будут пользоваться не только люди со специальными навыками, но и обычные пользователи, решающие рутинные профессиональные или бытовые задачи. Такое повседневное обживание технологий искусственного интеллекта происходит в обстоятельствах, где реальные практики обусловливаются структурами повседневной жизни и в меньшей степени поддаются контролю со стороны разработчиков, чем культурное воображаемое. Здесь на первый план выходит социология повседневной жизни как инструмент понимания того, каким образом люди вступают во взаимодействие с искусственным интеллектом и наделяют его смыслом.

Одомашнивание машин — давняя мечта, которая двигала и до сих пор движет разработчиками искусственного интеллекта. Хотя сегодня многие из них призывают не слишком доверять прогнозам о скором проникновении умных машин во все области человеческой жизни и даже отказаться от самой фразы «искусственный интеллект» как вводящей в заблуждение, мы по-прежнему сталкиваемся с обещаниями появления «домашних роботов», которыми сопровождается выпуск прототипов или действующих устройств и технологий, основанных на искусственном интеллекте. Некоторые из этих устройств уже покинули лаборатории и проникли в дома обычных людей. Самыми известными из них стали домашние роботы-пылесосы и умные колонки. И если роботы-пылесосы вряд ли могут считаться вершиной искусственного интеллекта, то умные колонки в полной мере соответствуют идеалу разработчиков, продавцов и поборников технологий искусственного интеллекта: это бытовые устройства, которые постоянно находятся дома, не требуют специальных навыков для пользования ими и взаимодействуют с людьми посредством естественного языка.

В данной статье мы рассмотрим умные колонки как новый элемент домашней социальности. Хотя обычные пользователи уже довольно давно сталкиваются с голосовыми помощниками (в смартфонах или компьютерах) и рекомендательными сервисами, разговаривают с телефонными чат-ботами, умные колонки вводят

¹ В данной статье мы будем использовать термин «искусственный интеллект», поскольку он получил наибольшее распространение в качестве собирательного названия определенного класса технологий, но при этом нас не будет интересовать, насколько такое название справедливо и действительно ли это «интеллект». Мы изучаем взаимодействие с соответствующими технологиями, а не то, моделируют ли они особенности какого-либо «естественного» интеллекта.

искусственный интеллект в парадигмально повседневную среду — домашнюю обстановку, и максимально упрощают взаимодействие с ним. Каким образом в этом случае происходит одомашнивание искусственного интеллекта? Как искусственный интеллект становится партнером по взаимодействию и опосредует рутинные домашние практики? Мы рассмотрим эти вопросы, отталкиваясь от ряда подходов в социологии повседневности, прежде всего — этнографии. Предметом нашего внимания будут исключительно бытовые практики. Мы остановимся на двух аспектах одомашнивания умных колонок — их включенности в домашнюю культуру и способах взаимодействия с ними². При этом задача будет заключаться в выделении и первичном описании ключевых социальных феноменов, связанных с этими практиками одомашнивания. В своем анализе мы будем отталкиваться от данных, полученных в рамках других исследований, от собственного опыта пользования умной колонкой и от данных, собранных в рамках осуществляемого нами с коллегами исследования особенностей взаимодействия с умными колонками.

Но прежде чем приступить к рассмотрению особенностей одомашнивания умных колонок, остановимся на специфике социологического интереса к взаимосвязи между домом и технологиями и, в частности, умными колонками.

Дом и технологии

Дом был важным, но парадоксальным объектом изучения в социологии на всем протяжении ее истории. С одной стороны, дом делает явными многие аспекты социальной жизни, связанные с центральными вопросами социальных наук: соотношение работы и досуга [Oakley, 2018], общественное разделение труда [Phizacklea, Wolkowitz, 1995], структуру гендерных [Pink, 2004; Scicluna, 2017] и классовых [Gilbert, 2016] отношений в обществе, процессы социализации [Bartlett, 1997], хозяйственную деятельность [Хокшилд, Мачун, 2020], родственные и семейные отношения [Smart, 2007], представление себя другим [Riggins, 1991], символическую организацию памяти и идентичности [Csikszentmihalyi, Rochberg-Halton, 1981; Chevalier, 1999]. При этом дом был своеобразным идеалом, в котором воплощались как обыденные, так и исследовательские представления о совершенной жизни [Chambers, 2020]. С другой стороны, собственно дом слишком легко не замечается, становясь для социологов экраном, на который проецируются другие, скрытые для невооруженного взгляда процессы и механизмы. Например, в известном приложении «Дом, или Перевернутый мир» к своей

² Это не означает, разумеется, что другие аспекты «оповедневивания» умных колонок не представляют интереса. Например, важную социологическую информацию дают социально-демографические характеристики пользователей умных колонок. Согласно данным Pew Research Center, в США портрет пользователя умных колонок выглядит следующим образом: молодой (32% пользователей в возрасте 18—29 лет) или средних лет (28% пользователей в возрасте 30—49 лет), со средним и высоким доходом, с высшим образованием, житель города или пригорода. См.: Auxier B. 5 Things to Know about Americans and Their Smart Speakers. URL: <https://www.pewresearch.org/fact-tank/2019/11/21/5-things-to-know-about-americans-and-their-smart-speakers/> (дата обращения: 17.01.2021).

Данные о российских пользователях умных колонок гораздо скучнее. Самое общее представление можно получить, например, из аналитического отчета компании Just AI. См.: Рынок умных колонок и голосовых ассистентов: Россия и мир. Прогнозы и тренды // Just AI. 2020. 7 сентября. URL: <https://just-ai.com/blog/rynek-umnyh-kolonok-i-golosovykh-assistentov-rossiya-i-mir-prognozy-i-trendy> (дата обращения: 17.01.2021).

Однако в любом случае в данной статье мы будем рассматривать только те изменения, которые умные колонки привносят в домашнюю социальность в качестве специфического (разговаривающего) объекта. Проблема пользователя, его характеристик и статуса требует отдельного обсуждения.

книге «Практический смысл» Пьер Бурдье рассматривает дом как «микрокосм, организованный в соответствии с теми же оппозициями, которые организуют универсум» [Бурдье, 2001: 528]. И даже если дом противопоставляется окружающему миру как особая область порядка, он все равно отражает этот мир, пусть и в перевернутом виде.

Социологи, не рассматривавшие дом как проекцию социальных кодов и практик, обращали внимание на то, что он представляет собой особое пространство повседневной жизни, где осуществляются практики, которые сложно или невозможно встретить в других местах [Pink et al., 2017]. В первую очередь это практики социальных взаимоотношений (интимных, родственных, дружеских; их принято относить к области «приватного», в противоположность «публичному»), но также — что важнее для нашего исследования, — технологические практики. Дом аккумулирует в себе определенные формы деятельности, а последние всегда имеют специфическое материальное, в том числе технологическое воплощение. Домашние практики заставляют инженеров и дизайнеров находить новые технические решения для их осуществления (например, разогрев еды и стирка белья радикально изменились в результате появления микроволновой печи и стиральной машины), но также сами оказываются производными от новых технологий, проникающих в дома и создающих соответствующие практики. Эта связь между домом и технологиями зачастую ускользает от внимания исследователей. Поставленный Дебби Хиндус в 1999 г. диагноз («...технологии в домах получают мало внимания в исследовательском сообществе» [Hindus, 1999: 199]) остается справедливым.

Цифровые технологии, разумеется, не обошли дом стороной. Когда начали появляться вычислительные машины, способные решать не только узкоспециализированные задачи, многие увидели их потенциал для радикального преображения домашней жизни. Наиболее полное воплощение эти ожидания получили в идее умного дома, которая до сих привлекает как разработчиков, так и обычных пользователей. Согласно определению Фрэнсис Олдрич, умный дом — это «жилище, оснащенное вычислительно-информационными технологиями, которые предвосхищают и реагируют на потребности жильцов, обеспечивая им комфорт, удобство, безопасность и развлечение посредством управления технологиями внутри дома и связи с внешним миром» [Aldrich, 2003: 17]. И хотя воплощение этой мечты оказалось не таким простым, как представлялось первоначально, различные элементы умного дома, включая умные терmostаты, умные лампочки, умные замки, умные жалюзи, умные чайники, умные холодильники и, конечно, умные колонки, уже с нами. Умный дом по своей природе социален, поскольку им следует управлять и он должен «реагировать на потребности жильцов», прежде всего через голосовые интерфейсы. До появления голосовых помощников это было почти невозможно. Сегодня вычислительные мощности и, главное, достижения в машинном обучении позволяют вести с различными устройствами полноценный (в техническом смысле) разговор — пользователи не просто дают голосовые команды, но обмениваются с «умными» устройствами репликами. В этом смысле современные умные устройства меняют социальность дома: в домашнем пространстве появляется еще один социальный объект, с которым можно разгово-

ривать и который при этом обладает рядом тревожных социальных способностей вроде способности постоянно слышать то, что произносится в доме.

Эта видоизмененная домашняя социальность требует социологического анализа, удобный объект для которого — использование умных колонок как распространенного домашнего устройства с голосовым интерфейсом³. Этот анализ особенно важен в ситуации, когда пандемия COVID-19 может существенно повлиять на формы организации домашней жизни и сместить или даже стереть границу между домом и работой.

Умная колонка как элемент домашней социальности

Дом — не только пространство, в котором отражается представление людей о том, кто они такие и какое положение в обществе они занимают, и которое играет принципиально важную роль в проведении границы между приватным и публичным. Дом — это также пространство повседневной жизни, где люди осуществляют специфические практики, связанные с организацией пространства и времени, управлением материальной средой, телесностью и коммуникацией с другими. Появление умных колонок в домах в той или иной степени затронуло все эти аспекты домашней социальности. У людей (что немаловажно, не обязательно владеющих смартфоном или компьютером) появился не только новый партнер по взаимодействию, но и новый инструмент осуществления повседневных практик, который способен менять традиционные формы домашней активности, а также предлагать новые возможности. Для социологии это важная и интересная ситуация, позволяющая наблюдать, с одной стороны, как происходит проникновение в повседневность новой технологии и с другой — как осуществляется одомашнивание искусственного интеллекта в той его форме, которая наиболее понятна, привычна и доступна, в форме разговорного искусственного интеллекта.

Основная характеристика умной колонки как домашнего устройства, разумеется, — возможность коммуникации с ней посредством голоса. Однако эта коммуникация предполагает ряд дополнительных условий, определяемых как техническими характеристиками устройства, так и особенностями среды, в которой оно должно использоваться.

Во-первых, умная колонка должна быть фоновым объектом, который всегда под рукой. Колонка должна быть «in hand», подручной (если воспользоваться термином М. Хайдеггера), и встраиваться в совокупность других домашних объектов. В отличие от, например, телевизора, который обычно предполагает длительное фокусирование на нем пользователя или пользователей (даже если телевизор работает в фоновом режиме), колонка должна быть малозаметной и только время от времени становится объектом внимания.

³ В 2019 г. по всему миру было продано 146,9 млн умных колонок. См.: Koksal I. The Sales of Smart Speakers Skyrocketed // Forbes. 2020. 10 March. URL: <https://www.forbes.com/sites/ilkerkoksal/2020/03/10/the-sales-of-smart-speakers-skyrocketed/?sh=6788f07538ae> (дата обращения: 17.01.2021).

В России, по данным «М. Видео-Эльдорадо», в первые семь месяцев 2020 г. было продано 1,3 млн устройств портативной акустики и умных колонок (причем продажи умных колонок выросли в 6,5 раз по сравнению с тем же периодом прошлого года). См.: Аналитика «М. Видео-Эльдорадо»: продажи умных колонок в России в 2020 году выросли в 7 раз // CNews. 2020. 16 сентября. URL: <https://cnews.ru/link/n515185> (дата обращения: 17.01.2021).

Во-вторых, умная колонка должна располагаться в доме так, чтобы она всегда была доступна для голоса. Поскольку умная колонка должна быть присоединена к источнику питания, что сильно ограничивает возможности ее перемещения внутри дома, с учетом конструкции современных жилищ это означает, что пользователям надо либо покупать несколько колонок, либо выбирать, в каком месте дома колонка наиболее нужна. Выбор места, в свою очередь, будет связан с тем, как колонка используется, поскольку любой дом представляет собой совокупность практических пространств, в которых люди осуществляют определенные виды деятельности (см. об этом ниже).

В-третьих, умная колонка должна служить домашним помощником. Это связано как с функционалом умной колонки (она может озвучить прогноз погоды, запустить музыку, сообщить нужный адрес, подсказать перевод или значение слова, рассказать анекдот и т. д.), так и с тем, что колонка представляет собой, по сути, физический интерфейс для голосового помощника (Alexa, Siri, Алисы или Маруси). В доме умная колонка исполняет служебную роль. Это выражается, среди прочего, в том, что она говорит, только когда к ней обращаются.

В-четвертых, так как умная колонка — домашнее устройство, она оказывается предметом коллективного пользования. Дом — это чаще всего место группового проживания, но даже в случае одиноко живущих людей колонка может быть объектом групповых действий (например, при посещении гостей). Эта особенность стимулируется определенными функциями, которые закладывают в колонку производители. Например, Amazon Echo, самая популярная в мире умная колонка, имеет функцию «Announcements», которая позволяет делать «объявления» для всего дома, то есть создавать систему «внутренней связи» (при наличии нескольких установленных колонок).

Эти социальные особенности использования умных колонок не остались без внимания ученых. Несмотря на новизну данной технологии, в ряде работ были выявлены важные особенности повседневного использования умных колонок в домах. Например, Э. Бенето и ее коллеги из Вашингтонского университета исследовали десять семей, которые в течение месяца впервые пользовались умной колонкой Amazon Echo Dot⁴ [Beneteau et al., 2019, 2020a, 2020b]. Авторов прежде всего интересовал вопрос внедрения и распространения инноваций. Они выяснили, что пользователи постепенно сокращают количество сценариев использования колонки и действуют только небольшой набор ее функций, что связано с тем, как они выявляют функциональность колонки. Чаще всего они спрашивали саму колонку о ее возможностях (хотя с течением времени отказывались от этого), использовали метод проб и ошибок или получали советы от родственников, друзей и знакомых и в гораздо меньшей степени полагались на информацию, предоставляемую производителем. Это исследование показывает, что, попадая в домашнюю среду, колонка становится элементом доминирующих там форм социальности, включая формы обучения и передачи знаний. Если предлагаемые колонкой формы социальности не соответствуют домашним, это существенно снижает ее возможности в качестве партнера по взаимодействию.

⁴ Amazon Echo Dot — более портативная версия умной колонки Amazon Echo. В отличие от Echo, Echo Dot имеет небольшие размеры и лишена отдельного дисплея.

Более интеракционный ракурс был выбран в исследовании М. Поршерона и его коллег [Porcheron et al., 2018; Fischer et al., 2019; Reeves, 2017]. С помощью конверсационного анализа они проанализировали взаимодействия пяти семей с умной колонкой Amazon Echo Dot в течение месяца. Их интересовало, каким образом взаимодействие с умной колонкой встраивается в другие домашние занятия вроде семейного ужина. Они показывают, что взаимодействие с колонкой часто разрушает параллельные формы деятельности, поскольку переключает внимание на само это взаимодействие, требующее обширных согласованных усилий членов семьи, которые при решении возникающих проблем руководствуются «политикой дома» — рутинными, устоявшимися способами домашнего разделения труда, прав и обязанностей.

Ф. Бентли и соавторы [Bentley et al., 2018] дополнили этот интеракционный анализ лонгитюдными количественными данными использования умных колонок в 88 домах в течение 110 дней. Они демонстрируют, что обращение к умным колонкам обусловливается суточными и недельными ритмами домашней жизни, связанными с рабочими графиками членов семьи и принятыми в доме формами досуга. Эта матрица домашней жизни настолько устойчива, что с течением времени способы использования умных колонок мало трансформируются. После первоначального периода тестирования возможностей колонки пользователи создают небольшой набор домашних рутин, в рамках которых эти колонки становятся необходимым элементом. Один из выводов из исследования Бентли и его коллег заключается в том, что использование умных колонок принципиально отличается от использования голосовых ассистентов, доступных на смартфонах через мобильные приложения. Умные колонки попадают в особую обстановку домашней жизни, материальная и социальная организация которой как стимулирует взаимодействие с умными колонками (ими пользуются гораздо чаще, чем голосовыми помощниками), так и создает препятствия для этого (например, содержание и форма взаимодействия с умной колонкой в значительной мере определяются местоположением колонки внутри дома).

Указанные исследования позволяют не только ответить на ряд эмпирических вопросов, касающихся особенностей использования умных колонок и их интеграции в домашнюю повседневность, но и поставить некоторые концептуальные вопросы, касающиеся отношений между пользователями подобного рода устройств и объектами, с которыми они взаимодействуют. Как отмечают М. Алач и соавторы, повсеместное распространение голосовых разговорных интерфейсов приводит к появлению «гиперсоциализированного мира, в котором все является социальным, и мы живем в кухнях, где с нами разговаривают даже тостеры» [Alač et al., 2020: 4]. Однако в основе технологий, обеспечивающих такую гиперсоциализацию, лежит представление об «индивидуальных Я как элементарных единицах социального» [ibidem]. Концентрация разработчиков на «пользователе» стабилизирует эти Я, в результате чего им приписываются устойчивые характеристики, ожидание которых закладывается в компьютерную систему [Grint, Woolgar 1997]. В результате «пользовательский опыт» не только становится в центр процесса создания голосовых интерфейсов, но и направляет внимание исследователей. Например, М. Поршерон с соавторами, аргументи-

руя необходимость отказа от используемого в индустрии термина «разговорные агенты» в пользу термина «голосовой пользовательский интерфейс», говорят, что «разговорное взаимодействие — ошибочное название для подобного типа человеко-компьютерной интеракции, поскольку оно смешивает взаимодействие с устройством в рамках разговора с самим разговором» [Porcheron et al., 2018: 9]. Иными словами, по мнению авторов, умная колонка, взаимодействие с которой они изучали, — это лишь то, что используется в разговоре, но не является партнером по взаимодействию, а значит, «двигателем» этой разговорной системы выступает пользователь, пытающийся решать с помощью соответствующего устройства свои практические задачи. М. Алач и ее коллеги предлагают отказаться от такой индивидуализированной концепции пользователя в пользу представления о разговаривающих устройствах как «элементах мира, к которым мы интеракционно открыты (а не орудиях, превращающих людей в более могущественных когнитивных агентов)» [Alač et al., 2020: 9].

В данной статье мы рассмотрим вторую составляющую интеракционной открытости, о которой говорят М. Алач и соавторы. Если они предлагают переопределить субъекта, взаимодействующего с разговорным искусственным интеллектом, то мы предлагаем проанализировать, какого рода объектом является разговаривающее устройство, в частности, умная колонка. Вопрос понимания объектов, с которыми взаимодействуют люди, — один из самых обсуждаемых в современной социологии. После долгого периода игнорирования роли объектов в социальном мире либо рассмотрения их как смысловых субстратов, меняющихся в зависимости от того, какой смысл им приписывают люди⁵, ряд исследователей предложили радикально «пересобрать» социальное, чтобы за объектами или вещами была признана исполняемая ими конститутивная роль [Ерофеева, 2017; Латур, 2014; Molotch, 2011]. Этот социологический «объективизм», однако, не позволяет понять, каков интеракционный статус объектов в социальном мире, который радикально меняется, когда объекты сами начинают разговаривать с людьми. Для этого нужны другие понятия и подходы, которые мы предложим ниже. В этом отношении мы продолжаем линию анализа «объектов как практических достижений» [Nevile et al., 2014: 4], то есть того, как объекты возникают во взаимодействии, но при этом не конструируются пользователями. Мы распространим эту логику анализа на специфические объекты — разговаривающие машины. На наш взгляд, они позволяют поставить ряд вопросов, которые были невозможны до появления этих объектов в повседневных ситуациях.

Указанные выше исследования дают возможность выделить два ключевых аспекта взаимодействия с умными колонками: их интеграцию в домашнюю повседневную культуру и собственно коммуникацию с ними.

Умные колонки и привычные домашние практики

Дом — специфическое социальное пространство. Его специфика определяется не только тем, что дом в определенном смысле противостоит рабочему месту как пространство досуга (хотя в последние годы это меняется и сегодня дом все чаще

⁵ Наиболее последовательный теоретик такого подхода, вероятно, Герберт Блумер [Блумер, 2017].

становится местом работы [Oakley, 2018]), но и тем, что дом — это пространство приватное, то есть такое, доступ к которому регулируется особыми социальными правилами (как юридическими, так и неформальными) и которое выступает своеобразным продолжением человека: его/ее личности, идентичности и даже телесности. С точки зрения социологии повседневности дом обладает следующими характеристиками:

1) Дом — это опривыченное пространство. Длительное нахождение в доме, его относительно небольшая территория и стабильный набор действий, совершаемых в рамках ежедневных и еженедельных рутин, приводят к формированию ряда привычек, связанных как с перемещениями внутри дома, так и с формами активности в нем. Например, есть ряд привычных действий, которые совершаются людьми дома после пробуждения или перед тем, как лечь спать. Эти привычки запечатлеваются в теле настолько глубоко, что мы можем перемещаться внутри дома и осуществлять различные действия, уделяя этому гораздо меньше внимания, чем в любой другой обстановке. Мы можем попасть в нужное место дома или найти нужную вещь с закрытыми глазами.

2) Чтобы эти привычки работали, необходимо определенным образом преобразовать домашнюю среду: она должна стать «слепком» домашних действий. В результате регулярного совершения определенных действий вещи внутри дома распределяются так, что какие-то из них оказываются на периферии привычных практик, а какие-то — всегда под рукой. В этом смысле пространство дома — это аккумулированное пространство. Вещи в нем «оседают» в различных местах в зависимости от того, в какие практики они включены. Одни вещи оказываются у задней стенки шкафа на верхней полке, другие — на раковине в ванной. Появление новых вещей порождает своеобразные «волны» в этом аккумулированном пространстве, заставляя переконфигурировать сложившиеся кластеры вещей. (Эти «волны» могут быть как очень небольшими, например, когда мы немного меняем положение продуктов в холодильнике, чтобы добавить туда пакет молока, так и довольно значительными, например, когда нам приходится перебирать кладовку, чтобы положить туда чемодан.) И практически в любом доме существуют также рутинные способы избавления от вещей (например, с помощью ведра для мусора). При этом аккумуляция определяется не только частотой использования тех или иных вещей, но и структурой самого пространства, вмещающего только вещи определенного размера, и локальной историей дома (например, тем, что в доме могут в разное время проживать разные люди, в результате действий которых те или иные вещи оказываются в тех или иных местах).

3) Еще одна важная характеристика дома, тесно связанная с двумя предыдущими, — зональность. В любом доме есть различные пространства действия (или «центры активности» [Crabtree, 2003: 12]), которые различаются между собой особой материальной конфигурацией, отражающей и обеспечивающей совершаемые там практики. Отчасти эта зональность определяется инфраструктурой жилища (в нем есть «кухня», «туалет», «ванная», «спальня» и т. д.), но в любом случае эта инфраструктура должна быть включена в привычные домашние практики, чтобы можно было говорить о ее влиянии. «Кухню» делает

«кухней» не то, как спроектирован дом, а то, используют ли жильцы это пространство в качестве кухни. При этом зоны внутри дома имеют пространственно иtempорально подвижные границы. Например, потребление еды или чтение могут происходить в разных местах. В этом смысле зоны внутри дома — скорее полюса притяжения, а не выделенные территории для определенных практик. Тем не менее можно говорить, что в определенных местах конкретные формы деятельности совершаются чаще. Например, телевизор чаще располагается в гостиной, чем в других местах дома, и, соответственно, связанные с телевизором индивидуальные и коллективные практики чаще осуществляются именно там.

Что происходит, когда в это привычное, аккумулированное, зонированное пространство дома попадает умная колонка? Оказывает ли ее появление одинаковое воздействие на три выделенных аспекта дома или какие-то из них трансформируются сильнее? Чтобы ответить на эти вопросы, необходимо внимательнее посмотреть на то, что именно меняется в привычных домашних практиках после появления умной колонки.

Во-первых, меняется фон некоторых практик. Все действия внутри дома имеют определенный материальный фон — это может быть как некоторая конфигурация вещей, так и звуковой фон. Часто этот фон управляется самими акторами. Поскольку умная колонка предназначена для воспроизведения звуков, она меняет звуковой фон: например, определенные действия вроде приготовления еды начинают сопровождаться музыкой, проигрываемой с помощью колонки. Однако проигрывание музыки — не единственный способ, которым умные колонки влияют на фон деятельности. Во многих случаях умная колонка представляет собой фоновый источник информации, к которому периодически обращаются при осуществлении какой-то другой деятельности. Скажем, за семейным ужином участники могут задавать колонке вопросы, касающиеся значения слов или сведений по определенной теме. Такое фоновое информационное присутствие умной колонки имеет и обратную сторону — поскольку умная колонка не только сообщает, но и собирает информацию, она является постоянным фоновым слушателем всего, что говорят люди дома, причем слушателем, который подключен к интернету и может отправлять услышанное за пределы дома. У многих пользователей такое фоновое присутствие умной колонки вызывает беспокойство по поводу сохранения приватности происходящего в доме. Результатом этого беспокойства может быть как появление определенных практик пользования колонкой (например, выключение ее в определенные моменты), так и избегание тем, которые считаются пользователями непубличными.

Во-вторых, присутствие умной колонки может менять само содержание некоторых практик. В качестве примера можно привести прием гостей. Колонка может становиться не только фоновым проигрывателем музыки, но и инструментом развлечения: хозяева демонстрируют гостям, что умеет и не умеет делать искусственный интеллект, а гости устраивают «тесты», задавая колонке сложные вопросы или давая веселые задания. В этом случае застольные или дружеские практики частично меняют свое содержание, и среди знакомых практик обмена новостями, рассказывания шуток и т. д. появляется еще одна неосуществимая

без умной колонки и использующая как ее возможности, так и ее ограничения в качестве ресурсов. При этом коммуникативные сбои в общении с колонкой могут становиться поводом для развлечения и веселья, а не раздражения и недовольства.

В-третьих, умные колонки могут менять частоту осуществления ряда привычных практик. Например, пользователи могут начать чаще пользоваться таймером, узнавать прогноз погоды и новости или прослушивать музыку, чем это было раньше. Отчасти увеличение частоты этих практик связано с ускорением процесса: получить информацию о погоде у искусственного интеллекта быстрее, чем взять подходящее устройство и запустить специальную программу или открыть веб-сайт. Но изменение частоты также связано с тем, что сама умная колонка по мере ее использования начинает ассоциироваться с определенными практиками, которые можно с ее помощью осуществлять, и эти практики могут «открепляться» от других материальных объектов (вроде смартфонов) при их осуществлении дома.

Выделенные особенности трансформации повседневных домашних практик, вызываемые появлением в доме умных колонок, позволяют утверждать, что они оказывают неодинаковое влияние на три указанных выше аспекта домашней жизни. Наиболее заметно они влияют на устоявшиеся практики, меняя их содержание или фон, а также приводя к формированию новых домашних привычек. Кумулятивность домашнего пространства умные колонки затрагивают минимально или вообще не влияют на нее, поскольку умные колонки включены скорее в коммуникативные домашние практики и лишь косвенно связаны с организацией материальной среды в доме. Что касается зональности дома, то умные колонки здесь оказывают определенное воздействие, но не такое сильное, как в случае привычных практик. Хотя есть определенные технические обстоятельства, определяющие, в каком месте дома будет находиться колонка,— например, расположение розеток, длина провода, близость стен (если звук в колонке регулируется датчиком движения, колонку нельзя ставить близко к стенам),— с социологической точки зрения важны, прежде всего, практические социальные обстоятельства, обуславливающие выбор места. Во-первых, умная колонка должна быть размещена в том месте, которое кажется жильцам наиболее подходящим, в зависимости от того, как они используют колонку и как они воспринимают ее возможности. Если колонка используется для проигрывания фоновой музыки, она может оказаться, например, на кухне. Если она используется для получения общей информации и игр, она будет в гостиной. Если важна возможность ставить будильник, колонка будет расположена в спальне. Во-вторых, колонка должна быть размещена так, чтобы она была доступна для голоса, но при этом не мешала осуществлять другие действия. Чаще всего колонки оказываются на периферии жизненного пространства той комнаты, в которой они устанавливаются.

Например, у автора настоящей статьи умная колонка после нескольких попыток «закрепления» на кухне оказалась в основной комнате, исполняющей одновременно роль спальни и гостиной (см. рис. 1).

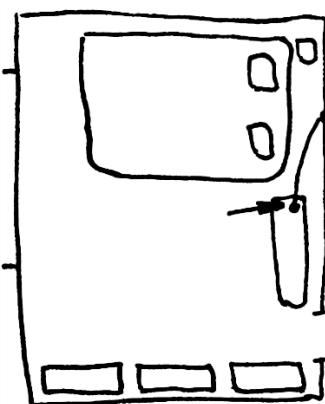


Рис. 1. Расположение умной колонки в доме автора

Такое ее расположение (на невысокой тумбе возле кровати) связано в первую очередь с доступностью для голоса людей, использующих эту комнату, и со спецификой пользования колонкой: чаще всего она сообщает информацию о погоде. В то же время колонка находится далеко от наиболее часто используемых мест в комнате: она не создает помех для потоков вещей и действий, не связанных с ней, и необходимость ее перемещения возникает только в процессе уборки.

Одомашнивание искусственного интеллекта, воплощенного в умной колонке, не только предполагает выработку особых форм коммуникации с ним и приводит к трансформации повседневных практик в доме, но и порождает новые практики, которые могут менять привычный прагматический ландшафт дома. До появления умной колонки граница между приватным и публичным проходила по границе дома, которая представляла собой, по выражению Эрвина Гофмана, своеобразную «мембрну» [Goffman, 1961: 65—66]. Эта мембрана не была абсолютно непроницаемой (у большинства домов есть двери и окна, и обеспечить их полную свето- и звукоизоляцию невозможно), но существенно ограничивала возможности перемещения людей, вещей, информации и звуков через нее. Появление умной колонки меняет эту конфигурацию приватного/публичного. Теперь граница между тем и другим проходит внутри дома и в пространстве коммуникации. Несмотря на то что все производители умных колонок заверяют покупателей, в гарантированном сохранении приватности данных, наличие в доме устройства, которое постоянно слушает разговоры, может существенно влиять на то, что и как говорится в доме. Люди могут избегать конфиденциальных разговоров в присутствии колонки или стараться не разглашать чувствительную информацию, например, отказываясь совершать покупки через умные колонки [Lau, Zimmerman, Schaub 2018; Malkin et al., 2019; Voit et al., 2020]. Отчасти такое избегающее поведение обусловливается тем, что при взаимодействии с умной колонкой посредством голоса пользователи неизбежно должны полагаться на привычное чередование ролей слушающего и говорящего, лежащее в основе любой коммуникации при помощи естественного языка. Поэтому пользователи знают, что умная колонка в принципе способна быть слушателем. Другая причина заключается в том, что

сам факт постоянного присутствия колонки в доме и ее готовности к коммуникации приводит к приписыванию ей «подслушивающей» позиции. Если раньше приватность регулировалась физической мембранный границы дома, то теперь она регулируется доступностью или недоступностью высказываний для искусственного интеллекта.

Однако соображения приватности — не единственное обстоятельство, которое обуславливает появление новых практик внутри дома под влиянием умных колонок. Другая причина связана с коллективным использованием умных колонок. У колонки может быть несколько пользователей, конкурирующих между собой за управление ею. Колонка провоцирует такую конкуренцию, делая всех пользователей равнозначными. Это может приводить к тому, что пользователи иногда управляют не столько действиями колонки, сколько действиями других пользователей в отношении колонки (например, корректируют чужие запросы) [Porcheron et al., 2018].

Еще один аспект одомашнивания искусственного интеллекта — возможность с ним «играть». Здесь уместно воспользоваться социологическим различием ситуативной игры (*play*) и формальной игры (*game*) [Mead, 2015]. С умными колонками возможны обе, но в данном случае они имеют свою специфику, что позволяет говорить о них как новых практиках, связанных с появлением умных колонок в домах. Что касается ситуативной игры, то, конечно, умные колонки — не единственный объект, с которым осуществляется такой тип игр. Взрослые, например, играют в ситуативные игры с детьми, домашними животными и друг с другом. Однако разные типы ситуативных игр предполагают разные коммуникативные способности партнеров по игре. Поэтому в той мере, в какой о разговорном искусственном интеллекте можно говорить как о партнере по взаимодействию, его коммуникативные особенности будут требовать особых типов ситуативных игр. Например, пользователи могут пытаться заставить искусственный интеллект произнести определенное слово в ситуации, когда он запрограммирован не исполнять просьбу типа «Х, скажи Y» (в собранных нами данных есть несколько таких ситуаций). Эти же коммуникативные особенности искусственного интеллекта могут использоваться в формальных играх, где пользователи должны выполнять определенные действия, предусмотренные правилами. Так, некоторые из подобных игр предполагают, что пользовать должен с помощью одной фразы заставить умную колонку произнести определенное слово, не называя это слово или его производные.

Коммуникация с умной колонкой

Главное, с чем приходится сталкиваться пользователю умной колонки, — это коммуникативные особенности разговорного искусственного интеллекта как партнера по взаимодействию. Во многом процесс одомашнивания заключается в том, что пользователи приспосабливаются к тому, как искусственный партнер говорит, делает паузы, отвечает на вопросы, задает вопросы и т.д. Несмотря на значительный прогресс, достигнутый в последние годы, разговорный искусственный интеллект все еще демонстрирует способности, не позволяющие вести с ним «нормальный» разговор и приводящие к многочисленным сбоям во взаимодей-

ствии (хотя, как указывалось выше, эти сбои могут быть не только источником недовольства, но и поводом для развлечения).

В данном случае отправной точкой в анализе коммуникации с разговорным искусственным интеллектом лучше брать не дихотомическую шкалу «человеческая коммуникация — все другие формы коммуникации», поскольку и сама человеческая коммуникация осуществляется в очень разнообразных формах, и коммуникация с другими партнерами по взаимодействию не всегда оценивается с точки зрения того, является ли собеседник человеком. Лучше исходить из того, что существует коммуникативный спектр, в котором разговорный искусственный интеллект занимает свое место (или даже несколько мест).

Взаимодействие с умной колонкой разворачивается как последовательность высказываний, в которой действуют стандартные ожидания относительно диалога: что реплики будут сменять друг друга без длинных пауз между ними и что собеседники не будут говорить одновременно [Сакс, Шеглофф, Джейферсон, 2015]. При этом в каждой своей ответной реплике говорящий должен демонстрировать понимание сказанного ранее и делать уместные высказывания. Современные умные колонки с трудом могут соблюдать эти условия, что побуждает пользователей находить способы корректировки своих действий и действий колонки. Наиболее распространенные способы исправления возникающих проблем таковы:

- просодическая модификация (повышение громкости, замедление речи, интонационные подчеркивания);
- семантическая модификация (например, замена слов и сокращение длины фраз);
- фонетическая модификация (например, пользователи замолкают после фразы, чтобы колонке было легче распознать сказанное).

В рамках диалога с умной колонкой важное значение имеет то, как пользователи интерпретируют действия искусственного интеллекта. В этом отношении важны два момента:

1) Молчание колонки. Поскольку взаимодействие с колонкой ведется в разговорном формате, для пользователей любое молчание со стороны колонки приобретает смысл в связи с текущей ситуацией и чаще всего интерпретируется либо как время, когда колонка «думает» (то есть совершает какие-то операции), либо как свидетельство проблемы, с которой столкнулась колонка.

В качестве примера ориентации на молчание как «значащее» действие в разговоре с умной колонкой можно привести следующий эпизод⁶:

- 40 П1 x:::: а: какой тебе нравится фильм
41 (.) больше всего?
42 (4.3)
43 П2 (всё сломалась)

⁶ В нижеследующих транскриптах применяется конверс-аналитическая система транскрибирования речи, разработанная Гейл Джейферсон. Используются следующие символы: П — пользователь, А — Алиса, (0.0) — паузы в целых и десятых долях секунды, (.) — пауза около 0,1 секунды, ::— растягивания звука (количество двоеточий отражает продолжительность растягивания), слово — акцентирование, ↑ ↓ — повышение и понижение тона, * * — понижение громкости в выделенном фрагменте, [— момент наложения реплик, (слово) — неуверенное распознавание, < > — более медленное произнесение. Все записи, транскрипты которых приводятся ниже, были собраны автором и командой исследователей, изучающих взаимодействие пользователей с умными колонками.

44 (0.5)

45 П1 алло алиса привет ты здесь

Длительное молчание голосового помощника (строка 42) после прямого вопроса создает у пользователя 2 впечатление, что помощник «сломалась». Термин «сломалась» может означать разные вещи, однако реакция пользователя 1 в строке 44 показывает, что для него проблема заключается не в физической поломке, а в коммуникативном сбое: Алиса перестала разговаривать и ее надо «расшевелить», для чего применяются типичные «стартовые» выражения, которые чаще всего используются для начала разговора или начала фразы («алло», «привет»). В этом случае молчание колонки расценивается пользователями как свидетельство возникшей проблемы, и эта проблема заключается не в особенностях формулировки вопроса, а в выпадении голосового помощника из взаимодействия.

2) Ответы колонки представляют собой ресурс для ведения диалога и выявления проблем: для пользователя все действия колонки являются индикатором того, что происходит сейчас, происходило ранее и может происходить далее в разговоре. Но иногда информация, которую пользователи извлекают из действий голосового помощника, не соответствует действительному положению дел, и пользователи могут видеть проблему там, где ее нет, либо не замечать существующую проблему. Например, в следующем фрагменте из другого разговора пользователь «пропускает» неуместный ответ голосового помощника:

65 П ну давай анекдот какойнибудь расскажи
66 (3.6)
67 А в недрах тундры (.) выдра в гетрах (.)
68 тырит в ведра (.) ядра кедров.
69 (1.4)
70 П *понятно* это кто написал такой стишок?
71 (4.9)
72 А на сол точка ру (.) есть такой ответ (0.4)
73 стихи пишутся о главном (.) о любви,
74 дружбе, родине, людях, детях.
75 (1.4)
76 П всё ясно (.) *м* какая твоя любимая
77 песня битлз?

«Ответ» Алисы на вопрос, заданный в строке 70, не соответствует этому вопросу. Несмотря на разрушение связности разговора, пользователь реагируют подтверждающей фразой «все ясно» и задает следующий вопрос. В этом случае резкая смена темы (от «стишка» к «песням Битлз») может означать, что пользователь решил не обращать внимание на неуместность ответа Алисы, но независимо от того, осознает он этот сбой в связности разговора или нет, в самом взаимодействии проблема игнорируется.

Важный аспект коммуникативных способностей разговорного искусственного интеллекта — контекстуальная нечувствительность, которая проявляется как в ходе самого разговора, так и в предшествующих ему обстоятельствах. Во время разговора быстро обнаруживается, что умная колонка не может отслеживать локальную историю разговора, и поэтому каждую новую реплику лучше формули-

ровать так, как если бы она была первой репликой и ей не предшествовали другие. Для пользователей это всегда непростая задача, поскольку в обычной коммуникации мы полагаемся на предшествующие реплики как на ресурс понимания текущей реплики, доступный всем участникам разговора. О том, что пользователи ориентируются на контекстуальную нечувствительность голосовых помощников, свидетельствует частое использование активирующих фраз в начале реплик. Все современные умные колонки предполагают использование активирующей фразы для начала разговора. Это может быть словосочетание («Окей, Гугл») или имя («Алиса»). Однако пользователи часто используют активирующую фразу не только в начале разговора, но и в начале отдельных реплик внутри разговора. Как показывает следующий фрагмент, далеко не всегда это связано с технической необходимостью «разбудить» колонку.

- 15 П Алиса скажи (0.3) <как> (0.6)
16 по: таджики будет школа.
17 (2.2)
18 А школа на таджикском (.) мақта:(б)е
19 (1.0)
20 П .ф пасиба Алиса
21 (1.9)
22 А я стараюсь
23 П ↑Алиса а какая твоя любимая книга.

В строке 23 пользователю нет необходимости использовать активирующую фразу, поскольку после окончания реплики голосового помощника прошло непродолжительное время, однако пользователь все равно начинает свою реплику с «Алиса», не столько «активируя» голосового помощника, сколько начиная новую тему. Мы видим, что даже в отсутствие необходимости «пробуждения» колонки пользователи обращаются к голосовому помощнику по имени, не полагаясь на поддержание связности разговора за счет удержания искусственным интеллектом предыдущего хода взаимодействия.

При этом, сколь бы уникальными ни были активирующие фразы, люди могут произносить или сами эти слова, или слова, похожие на них, в ходе обычных разговоров, тем самым активируя умную колонку тогда, как они этого не хотят. Это нечаянное начало разговора приводит к тому, что люди вынуждены прерывать взаимодействие (это одна из причин, по которой самое частое слово, используемое при взаимодействии с умными колонками,— «стоп»⁷). Чтобы избегать такого неуправляемого начала разговора, пользователи могут стараться не произносить ключевые слова, когда они, например, рассказывают кому-то о колонке в ее присутствии.

Другой аспект коммуникативных сбоев, с которыми сталкиваются пользователи умных колонок, заключается в том, что, если взаимодействие с колонкой включается в какую-то другую деятельность, сбой в этом взаимодействии может приводить к разрушению основной деятельности. Например, как показано в работе М. Поршерона и соавторов [Porcheron et al., 2018], если во время семейного ужина его участники пытаются заставить умную колонку поиграть с ними в игру или

⁷ По данным [Sciuto et al., 2018] и [Bentley et al., 2018].

проиграть музыкальную композицию, но колонка их не понимает, это на какое-то время приводит к разрушению семейного ужина как специфической практики. Участники начинают фокусироваться на колонке, оставляя другие формы деятельности и переставая коммуницировать друг с другом.

Наконец, стоит отметить такую важную особенность одомашнивания искусственного интеллекта, как проявление вежливости. Речь идет о том, каким образом пользователи обращаются к умным колонкам, как они начинают и заканчивают разговор и реагируют на действия искусственного интеллекта. Помимо «официального» обращения по имени пользователи могут обращаться к колонке разными другими способами, в том числе теми, которые при взаимодействии с людьми сочли бы «невежливыми». Например, как показывают наши данные, в моменты коммуникативных сбоев пользователи могут называть колонку дурой. В то же время пользователи часто избегают привычных проявлений вежливости при взаимодействии с искусственным интеллектом. Это может быть связано с отсутствием ожидания, что искусственный интеллект будет «обижаться» не отсутствие «спасибо», «пожалуйста» и «до свидания», поскольку не способен понимать даже простые фразы. Но гораздо важнее, что искусственный интеллект воспринимается как компьютерная система, взаимодействие с которой носит более прагматический характер. Столкнувшись несколько раз с тем, что умная колонка неправильно понимает высказывание, пользователи избавляются от «непрагматических» компонентов взаимодействия вроде форм вежливости. Когда задача — произнести высказывание так, чтобы искусственный интеллект понял сказанное, это высказывание должно быть выстроено таким образом, чтобы максимально облегчить искусственному интеллекту задачу понимания: лучше говорить громко, медленно, с большими паузами и сообщать только то, что позволит добиться желаемого.

Но при рассмотрении роли вежливости во взаимодействии с умными колонками не следует забывать, что речь идет о специфическом типе домашних отношений, частью которых становится умная колонка. Как отмечалось выше, в доме колонка играет сервисную роль, «выполняя» просьбы пользователей. Эти сервисные отношения могут побуждать пользователей оценивать как свои реплики, так и реплики голосового помощника с точки зрения того, каким образом они отражают эти отношения. Рассмотрим следующий фрагмент:

- 04 П понятно. алиса скажи какой компьютер
05 (.) лучше всего купить сейчас
06 (2.0)
07 А ответ есть на выборовед точка ру (0.3)
08 читаю (.) приобрести хороший компьютер
09 (.) можно тремя способами. (0.2)
10 купить готовый (.) заказать сборку (.)
11 или сделать ее самому
12 (0.9)
13 П очень информативно пасиба (0.6)
14 .x::: а как ты думае[шь-
15 А [поняла (.) исправлюсь
16 (0.5)

17 П нет нет всё всё хорошо всё хорошо
18 я (...) удовлетворен ответом (1.7)

Реплика Алисы в строке 15 становится для пользователя поводом для оценки того, как могла быть воспринята его реакция (строка 13) на ответ голосового помощника (строки 7—11). Судя по тому, как пользователь в строках 17 и 18 пытается исправить «неправильное» впечатление от своей реакции, он полагает, что Алиса сочла его реплику в строке 13 выражением неудовольствия полученным ответом. Независимо от того, насколько ироничной была фраза «очень информативно», пользователь при починке разговора привлекает контекст сервисных отношений, в соответствии с которыми его реплики выражают удовлетворенность предоставленной услугой (информацией). Подобная забота о том, чтобы «не обидеть» голосового помощника, свидетельствует, что контекст сервисных отношений способен «возвращать» определенные формы вежливого поведения во взаимодействие с голосовыми помощниками.

Заключение

Мы рассмотрели общие механизмы одомашнивания искусственного интеллекта, выделив ключевые аспекты домашних практик, частью которых становятся умные колонки. Мы показали, что умные колонки — это не просто новый объект, на который распространяются старые, привычные практики взаимодействия, а то, что трансформирует домашнюю социальность и побуждает изобретать новые формы действия. Дом, разумеется, не единственная среда, в которую будут проникать и уже проникают технологии искусственного интеллекта. Социальная организация каждой из этих сред накладывает свой отпечаток на то, как пользователи взаимодействуют с этими технологиями в своей повседневной жизни. Однако в той мере, в какой искусственный интеллект будет все больше предполагать использование голоса в качестве канала коммуникации с различными устройствами, между различными социальными обстановками будут обнаруживаться определенные сходства. Некоторые из них были отмечены в этой статье.

Как мы видели, одомашнивание искусственного интеллекта происходит не так, как предполагают разработчики. Социальные ситуации имеют собственную организацию, определяющую «естественность» взаимодействия с умными устройствами. Без учета этой организации, которую социологи выявляют, анализируя повседневные практики, технологии искусственного интеллекта будут в лучшем случае забавной игрушкой, а не полезной частью повседневной жизни. Социологи способны предоставить два типа описаний, которые могут лежать в основу разработки более социально ориентированных алгоритмов. С одной стороны, это детальный анализ различных повседневных культур (например, домашних⁸). С другой стороны, это детальный анализ практик взаимодействия с алгоритмами в естественных ситуациях. На наш взгляд, именно в этих двух направлениях должна развиваться социология искусственного интеллекта.

⁸ В качестве примера такого рода социологического анализа можно привести серию исследований группы британских социологов, занимавшихся изучением особенностей домашней среды в рамках проекта по разработке умного дома [Chamberlain, Crabtree, 2016; Crabtree, 2003; Crabtree, Rodden, 2004; Crabtree, Tolmie, 2016; Crabtree et al. 2015; Harper, 2003, 2011; Nilsson et al., 2019; Tolmie et al., 2008].

Список литературы (References)

- Блумер Г. Символический интеракционизм: перспектива и метод / пер. с англ. А. М. Корбута. М.: Элементарные формы, 2017.
- Blumer H. (2017) Symbolic Interactionism: Perspective and Method. Moscow: Elementary Forms Press. (In Russ.)
- Бурдье П. Практический смысл / пер. с фр. А. Т. Бикбова, К. Д. Вознесенской, С. Н. Зенкина, Н. А. Шматко; под ред. Н. А. Шматко. СПб.: Алетейя, 2001.
- Bourdieu P. (2001) Le sens pratique. Saint Petersburg: Aletheia. (In Russ.)
- Ерофеева М. А. Проблематизация объекта в современной теоретической социологии // Социологические исследования. 2017. № 7. С. 13—23.
- Erofeeva M. A. (2017) The Problem of the Object in Contemporary Social Theory. *Sociological Studies*. No. 7. P. 13—23. (In Russ.)
- Латур Б. Пересборка социального: введение в акторно-сетевую теорию / пер. с англ. И. Полонской; под ред. С. Гавриленко. М.: Изд. дом Высшей школы экономики, 2014.
- Latour B. (2014) Reassembling the Social: An Introduction to Actor-Network-Theory. Moscow: The Higher School of Economics Publishing House. (In Russ.)
- Сакс Х., Щеглофф Э. А., Джефферсон Дж. Простейшая систематика организации очередности в разговоре / пер. с англ. А. М. Корбута // Социологическое обозрение. 2015. Т. 14. № 1. С. 142—202. URL: <https://sociologica.hse.ru/2015-14-1/147275640.html> (дата обращения: 22.02.2021).
- Sacks H., Schegloff E. A., Jefferson G. (2015) A Simplest Systematics for the Organization of Turn-Taking for Conversation. *Russian Sociological Review*. Vol. 14. No. 1. P. 142—202. (In Russ.) URL: <https://sociologica.hse.ru/2015-14-1/147275640.html> (accessed: 22.02.2021).
- Хокшилд А., Мачун Э. Вторая смена: работающие семьи и революция в доме / пер. с англ. И. Кушнаревой; под ред. А. Космарского. М.: Изд. дом Высшей школы экономики, 2020.
- Hochschild A. R., Machung A. (2020) The Second Shift: Working Parents and the Revolution at Home. Moscow: The Higher School of Economics Publishing House. (In Russ.)
- Alač M., Gluzman Y., Aflatoun T., Bari A., Jing B., Mozqueda G. (2020) Talking to a Toaster: How Everyday Interactions with Digital Voice Assistants Resist a Return to the Individual. *Evental Aesthetics*. Vol. 9. No. 1. P. 3—53. URL: https://eventalaesthetics.net/wp-content/uploads/2020/05/EAV9N1_2020_Alac_Toaster_3_53.pdf (accessed: 18.01.2021).
- Aldrich F. K. (2003) Smart Homes: Past, Present and Future. In: Harper D. (ed.) *Inside the Smart Home*. London: Springer. P. 17—39. https://doi.org/10.1007/1-85233-854-7_2.
- Bartlett S. N. (1997) Housing as a Factor in the Socialization of Children: A Critical Review of the Literature. *Merrill-Palmer Quarterly*. Vol. 43. No. 2. P. 169—198.

- Bentley F., LuVogt Ch., Silverman M., Wirasinghe R., White B., Lottridge D. (2018) Understanding the Long-Term Use of Smart Speaker Assistants. In: *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*. Art. No. 91. <https://doi.org/10.1145/3264901>.
- Beneteau E., Richards O. K., Zhang M., Kientz J. A., Yip J., Hiniker A. (2019) Communication Breakdowns Between Families and Alexa. In: *CHI'19: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. Paper No. 243. <https://doi.org/10.1145/3290605.3300473>.
- Beneteau E., Guan Y., Richards O. K., Zhang M. R., Kientz J. A., Yip J., Hiniker A. (2020a) Assumptions Checked: How Families Learn About and Use the Echo Dot. In: *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*. Art. No. 3. <https://doi.org/10.1145/3380993>.
- Beneteau E., Boone A., Wu Y., Kientz J. A., Yip J., Hiniker A. (2020b) Parenting With Alexa: Exploring the Introduction of Smart Speakers on Family Dynamics. In: *CHI'20: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. <https://doi.org/10.1145/3313831.3376344>.
- Cave S., Dihal K., Dillon S. (eds.) (2020) AI Narratives: A History of Imaginative Thinking About Intelligent Machines. Oxford: Oxford University Press. <https://doi.org/10.1093/oso/9780198846666.001.0001>.
- Chamberlain A., Crabtree A. (2016) Searching for Music: Understanding the Discovery, Acquisition, Processing and Organization of Music in a Domestic Setting for Design. *Personal and Ubiquitous Computing*. Vol. 20. No. 4. P. 559—571. <https://doi.org/10.1007/s00779-016-0911-2>.
- Chambers D. (2020) Cultural Ideals of Home: The Social Dynamics of Domestic Space. London: Routledge.
- Chevalier S. (1999) The French Two-Home Project: Materialization of Family Identity. In: Cieraad I. (ed.) *At Home: An Anthropology of Domestic Space*. Syracuse: Syracuse University Press. P. 83—94.
- Crabtree A. (2003) The Social Organization of Communication in Domestic Settings. <http://www.cs.nott.ac.uk/~pszaxc/work/IEMCA03.pdf> (accessed: 18.01.2021).
- Crabtree A., Rodden T. (2004) Domestic Routines and Design for the Home. *Computer Supported Cooperative Work*. Vol. 13. No. 2. P. 191—220. <https://doi.org/10.1023/b:cosu.0000045712.26840.a4>.
- Crabtree A., Rodden T., Tolmie P., Mortier R., Lodge T., Brundell P., Pantidi N. (2015) House Rules: The Collaborative Nature of Policy in Domestic Networks. *Personal and Ubiquitous Computing*. Vol. 19. No. 1. P. 203—215. <https://doi.org/10.1007/s00779-014-0771-6>.
- Crabtree A., Tolmie P. (2016) A Day in the Life of Things in the Home. In: *CSCW'16: Proceedings of the 19th Conference on Computer Supported Cooperative Work (San*

Francisco (CA), USA, 27 February—2 March 2016). New York: ACM Press. P. 1738—1750. <https://doi.org/10.1145/2818048.2819954>.

Csikszentmihalyi M., Rochberg-Halton E. (1981) The Meaning of Things: Domestic Symbols and the Self. Cambridge: Cambridge University Press.

Fischer J. E., Reeves S., Porcheron M., Sikveland O. (2019) Progressivity for Voice Interface Design. In: *CUI'19: Proceedings of the 1st International Conference on Conversational User Interfaces*. Art. No. 26. <https://doi.org/10.1145/3342775.3342788>.

Gilbert P. (2016) Classes, Gender, and Lifestyles in Domestic Space. *Actes de la recherche en sciences sociales*. Vol. 215. No. 5. P. 4—15.

Goffman E. (1961) Encounters: Two Studies in the Sociology of Interaction. Indianapolis: Bobbs-Merrill.

Grint K., Woolgar S. (1997) The Machine at Work: Technology, Work and Organization. Cambridge: Polity Press.

Harper D. (ed.) (2003) Inside the Smart Home. London: Springer.

Harper D. (ed.) (2011) The Connected Home: The Future of Domestic Life. London: Springer.

Hindus D. (1999) The Importance of Homes in Technology Research. In: *Cooperative Buildings. Integrating Information, Organizations, and Architecture: Proceedings of the Second International Workshop, CoBuild'99 (Pittsburgh, PA, USA, October 1—2, 1999)*. Berlin: Springer. P. 199—207. https://doi.org/10.1007/10705432_18.

Lau J., Zimmerman B., Schaub F. (2018) Alexa, Are You Listening? Privacy Perceptions, Concerns and Privacy-Seeking Behaviors With Smart Speakers. In: *Proceedings of the ACM on Human-Computer Interaction*. Art. 102. <https://doi.org/10.1145/3274371>.

Malkin N., Deatrick J., Tong A., Wijesekera P., Egelman S., Wagner D. (2019) Privacy Attitudes of Smart Speaker Users. *Proceedings on Privacy Enhancing Technologies*. No. 4. P. 250—271. <https://doi.org/10.2478/popets-2019-0068>.

Mead G. H. (2015) Mind, Self, and Society. Chicago: University of Chicago Press.

Molotch H. (2011) Objects in Sociology. In: Clarke A. J. (ed.) *Design Anthropology: Object Culture in the 21st Century*. Vienna: Springer. P. 100—116. https://doi.org/10.1007/978-3-7091-0234-3_8.

Nevile M., Haddington P., Heinemann T., Rauniomaa M. (2014) On the Interactional Ecology of Objects. In: Nevile M., Haddington P., Heinemann T., Rauniomaa M. (eds.) *Interacting with Objects: Language, Materiality, and Social Activity*. Amsterdam: John Benjamins. P. 3—26. <https://doi.org/10.1075/z.186.01int>.

Nilsson T., Crabtree A., Fischer J., Koleva B. (2019) Breaching the Future: Understanding Human Challenges of Autonomous Systems for the Home. *Personal and Ubiquitous Computing*. Vol. 23. No. 2. P. 287—307. <https://doi.org/10.1007/s00779-019-01210-7>.

- Oakley A. (2018) *The Sociology of Housework*. Bristol: Policy Press.
- Phizacklea A., Wolkowitz C. (1995) *Homeworking Women: Gender, Racism and Class at Work*. London: Sage.
- Pink S. (2004) *Home Truths: Gender, Domestic Objects and Everyday Life*. Oxford: Berg.
- Pink S., Leder Mackley K., Moroșanu R., Mitchell V., Bhamra T. (2017) *Making Homes: Ethnography and Design*. London: Bloomsbury Academic. <https://doi.org/10.4324/9781003085942>.
- Porcheron M., Fischer J. E., Reeves S., Sharples S. (2018) Voice Interfaces in Everyday Life. In: *CHI'18: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (Montreal QC, Canada, April 21–26, 2018)*. Paper No. 640. <https://doi.org/10.1145/3173574.3174214>.
- Reeves S. (2017) Some Conversational Challenges of Talking with Machines. In: *Talking With Conversational Agents in Collaborative Action: Workshop at the 20th ACM Conference on Computer-Supported Cooperative Work and Social Computing (CSCW'17) (Portland, Oregon, USA, February 25–March 1, 2017)*. URL: <https://nottingham-repository.worktribe.com/output/844412/some-conversational-challenges-of-talking-with-machines> (accessed: 18.01.2020).
- Riggins S. H. (1990) The Power of Things: The Role of Domestic Objects in the Presentation of Self. In: Riggins S. H. (ed.) *Beyond Goffman: Studies on Communication, Institution, and Social Interaction*. Berlin: Mouton de Gruyter. P. 341—368. <https://doi.org/10.1515/9783110847291.341>.
- Scicluna R. M. (2017) Home and Sexuality: The “Other” Side of the Kitchen. London: Palgrave Macmillan.
- Sciuto A., Saini A., Forlizzi J., Hong J. I. (2018) “Hey Alexa, What’s Up?”: Studies of In-Home Conversational Agent Usage. In: *DIS'18: Proceedings of the 2018 Designing Interactive Systems Conference (Hong Kong, China, June 2018)*. New York: ACM Press. P. 857—868. <https://doi.org/10.1145/3196709.3196772>.
- Smart C. (2007) *Personal Life*. Cambridge: Polity Press.
- Tolmie P., Crabtree A., Rodden T., Benford S. (2008) “Are You Watching This Film or What?”: Interruption and the Juggling of Cohorts. In: *CSCW'08. Proceedings of the ACM Conference on Computer Supported Cooperative Work (San Diego (CA), USA, 8—12 November 2008)*. New York: ACM Press. P. 257—266. <https://doi.org/10.1145/1460563.1460605>.
- Voit A., Niess J., Eckert C., Ernst M., Weingärtner H., Woźniak P. W. (2020) “It’s Not a Romantic Relationship”: Stories of Adoption and Abandonment of Smart Speakers at Home. In: Cauchard J., Löchtefeld M. (eds.) *MUM 2020: 19th International Conference on Mobile and Ubiquitous Multimedia*. New York: ACM. P. 71—82. <https://doi.org/10.1145/3428361.3428469>.

DOI: [10.14515/monitoring.2021.1.1797](https://doi.org/10.14515/monitoring.2021.1.1797)



А. М. Давыдова, М. А. Солянова, К. Соренсен

ДИСЦИПЛИНАРНЫЕ ПРАКТИКИ ЦИФРОВОГО СЕЛФ-ТРЕКИНГА: МЕЖДУ ЭМАНСИПАЦИЕЙ И КОНТРОЛЕМ

Правильная ссылка на статью:

Давыдова А. М., Солянова М. А., Соренсен К. Дисциплинарные практики цифрового селф-трекинга: между эмансипацией и контролем // Мониторинг общественного мнения: экономические и социальные перемены. 2021. № 1. С. 217—240. <https://doi.org/10.14515/monitoring.2021.1.1797>.

For citation:

Davydova A. M., Solyanova M. A., Sorensen K. (2021) Disciplinary Digital Self-Tracking Practices: Between Emancipation and Control. *Monitoring of Public Opinion: Economic and Social Changes*. No. 1. P. 217–240. <https://doi.org/10.14515/monitoring.2021.1.1797>. (In Russ.)

ДИСЦИПЛИНАРНЫЕ ПРАКТИКИ ЦИФРОВОГО СЕЛФ-ТРЕКИНГА: МЕЖДУ ЭМАНСИПАЦИЕЙ И КОНТРОЛЕМ

ДАВЫДОВА Александра Михайловна — студентка образовательной программы «Социология», Национальный исследовательский университет «Высшая школа экономики», Москва, Россия
E-MAIL: amdaydova_1@edu.hse.ru
<https://orcid.org/0000-0003-1922-1017>

СОЛЯНОВА Мария Алексеевна — студентка образовательной программы «Социология», Национальный исследовательский университет «Высшая школа экономики», Москва, Россия
E-MAIL: solyanova.m.a@gmail.com
<https://orcid.org/0000-0001-9392-1076>

СОРЕНСЕН Кристина — студентка образовательной программы «Социология», Национальный исследовательский университет «Высшая школа экономики», Москва, Россия
E-MAIL: ksorensen@edu.hse.ru
<https://orcid.org/0000-0002-7052-2985>

Аннотация. В связи с растущей цифровизацией общества селф-трекинг через приложения и устройства стремительно набирает популярность. Несмотря на его обширную экспансию в разные сферы жизни, сам феномен имеет противоречивую природу. С одной стороны, селф-трекинг способствует расширению возможностей для самопознания с целью дальнейшего преобразования себя, с другой — мобильные технологии вторгаются в личное пространство индивида, устанавливая контроль над практиками повседневной жизни. Именно такой парадоксальный характер цифрового селф-трекинга подталкивает к его

DISCIPLINARY DIGITAL SELF-TRACKING PRACTICES: BETWEEN EMANCIPATION AND CONTROL

Alexandra M. DAVYDOVA¹ — Sociology Program Student
E-MAIL: amdaydova_1@edu.hse.ru
<https://orcid.org/0000-0003-1922-1017>

Maria A. SOLYANOVA¹ — Sociology Program Student
E-MAIL: solyanova.m.a@gmail.com
<https://orcid.org/0000-0001-9392-1076>

Kristina SORENSEN¹ — Sociology Program Student
E-MAIL: ksorensen@edu.hse.ru
<https://orcid.org/0000-0002-7052-2985>

¹ National Research University Higher School of Economics, Moscow, Russia

Abstract. Due to ever-growing digitalisation, self-tracking devices and apps are gaining popularity. Despite broad penetration into various spheres of life, self-tracking itself is a contradictory phenomenon. On the one hand, it enhances self-awareness opportunities for further self-transformations; on the other hand, mobile technologies penetrate individual's private sphere, establishing control over everyday practices. This very paradoxical character of digital self-tracing requires its detailed analysis and investigation. A series of interviews with users of different digital trackers was undertaken to find the basic control mechanisms. They involve several disci-

далнейшему и более детальному изучению.

В рамках исследования проведена серия интервью с пользователями различных цифровых трекеров. В результате были выявлены основные механизмы осуществления контроля со стороны мобильных приложений. Они выражаются в наборе дисциплинарных (уведомления, визуализация данных пользователя) и социальных практик (шеринг во внутренних и внешних социальных сетях). Также в фокусе оказалась специфика взаимоотношений «пользователь — трекер», характеризующаяся не только стилями потребления и ролями, ранее описанными в литературе, но и механизмами легитимации контроля и трансформацией использования. Все перечисленные аспекты цифрового селф-трекинга связаны с тем, как пользователи проводят границу между контролем и эмансиацией.

Ключевые слова: продукты цифрового селф-трекинга, дисциплинарные практики, квантификация телесности, мобильные приложения, контроль со стороны трекеров, пользователь цифровых трекеров

Благодарность. Статья подготовлена в сотрудничестве с Лабораторией геймификации Сбербанка.

Введение

Селф-трекинг — измерение и анализ повседневных действий, привычек и поведения [Ajana, 2017: 1]. Этот феномен и, в частности, цифровой селф-трекинг, инструментами которого выступают технологические устройства или мобильные приложения, стремительно набирает популярность в связи с цифровизацией общества, сопровождающейся повышением доступности мобильных телефонов со встроенными или скачиваемыми приложениями и появлением специальных

disciplinary (notifications, data visualisation) and social (sharing in internal and external social media) practices. The study's focus was also 'user-tracker' relationship marked not only by specific consumption patterns and roles earlier described in the literature but also by legitimising control and transformation of the tracker's use. All the above-mentioned aspects of digital self-tracking are related to how users draw the line between control and emancipation.

Keywords: disciplinary practices, products of digital self-tracking, quantification of body, apps, tracker control, user of digital self-tracking

Acknowledgments. The article is prepared in collaboration with the SberBank Gamification Laboratory.

устройств-трекеров. На рынке представлен ряд продуктов цифрового селф-трекинга с новыми функциями, позволяющими измерять не только метрики здоровья и активности, но и показатели более специфических сфер человеческой жизнедеятельности (отношения, финансы, психология и т. д.). Таким образом, самонаблюдение проникает из более «количествоенных» сфер в более «качественные», или даже можно сказать «интимные», которые нельзя измерить напрямую с помощью количественных показателей.

Предметное расширение сфер воздействия селф-трекинга позволяет трансформировать все большее число аспектов жизни, открывает новые возможности для саморазвития и самопознания. Такие продукты позволяют построить уникальную персонализированную программу развития в зависимости от индивидуальных целей пользователя. Более того, инструменты цифрового селф-трекинга «освобождают» пользователей, снимая часть ответственности за прогресс индивида. Так, помимо построения персонализированной программы развития, устройства и приложения напоминают и дополнительно мотивируют выполнять ее с помощью элементов геймификации [Hassan, Dias, Hamari, 2019: 153; Rapp, 2018: 1]. Вместе с этим расширяется и влияние устройств цифрового селф-трекинга на жизнь пользователей: они проникают в личное пространство индивида и дисциплинируют его.

Продукты цифрового селф-трекинга функционируют на стыке двух дискурсов: освобождающего и принуждающего. С одной стороны, специализированные устройства помогают пользователю в саморазвитии — следят за его прогрессом и напоминают выполнять рекомендации для достижения желаемых результатов [Ним, 2018: 176]. С другой стороны, такие приложения и устройства оказывают дисциплинирующее воздействие и контролируют поведение пользователя [там же: 177].

Ключевым для настоящей работы стал вопрос, как пользователи продуктов цифрового селф-трекинга определяют для себя границу между контролем со стороны устройств и возможностями для саморазвития, которые эти устройства предоставляют.

Селф-трекинг: столкновение дискурсов эманципации и контроля

Селф-трекинг — феномен совсем не новый: идея «познай самого себя» широко известна по меньшей мере со времен Древней Греции. Измерение показателей индивидов из пространства публичного (примером могут служить весы в парках, которыми за небольшую плату может воспользоваться любой желающий) перешло в приватное. С развитием цифровых технологий отслеживание своих показателей снова приобретает аспект публичности — пользователи приложений для селф-трекинга получают возможность делиться своими результатами с другими. Несмотря на видоизменение форм селф-трекинга, фундаментальная идея познания себя с целью улучшения собственной жизни прослеживается до сих пор [Crawford, Lingel, Karppi, 2015: 481].

На этой идеи основывается и крайне популярное движение Quantified Self (количественное измерение себя/измеряемый «Я»/количественная самость) [Ramirez, 2014]. Оно объединяет пользователей и создателей устройств для цифрового селф-трекинга, в рамках которого используются самые разные инструменты —

от фитнес-браслетов и умных часов до мобильных приложений. Девиз Quantified Self — «самопознание через цифры» — указывает на становление такой количественной самости, в рамках которой мониторинг собственной деятельности открывает новые возможности для самоопределения и самосовершенствования. Происходит сдвиг в сторону рассмотрения себя и своего тела в качестве проекта, который нуждается в постоянном развитии и инвестициях. О сходстве с «проектом» или же с «мини-корпорацией» [Martin, 2000] также свидетельствуют показатели, основываясь на которых, пользователи делают выводы о своей результативности и эффективности [Ajana, 2017: 10]. И в рамках корпораций, и в рамках селф-трекинга такие выводы принимаются на базе объективных метрик, будь то объемы выручки, или ключевые показатели эффективности организации (KPI), или же количество пройденных человеком шагов.

Во многом благодаря цифровым технологиям, которые, в отличие от предшествующих, позволяют получать данные непрерывно и незамедлительно, происходит дальнейшее развитие цифрового селф-трекинга. Во-первых, пользователей продуктов трекинга становится все больше и больше: на 2019 г. каждый пятый взрослый житель США регулярно пользовался трекерами или мобильными приложениями для поддержания здоровья¹. Во-вторых, трекинг охватывает все больше сфер деятельности человека и помимо базовых количественных характеристик (таких как рост, вес, количество пройденных шагов, скорость пробежки) начинает включать в себя и более «качественные» метрики, связанные с субъективным опытом пользователя (например, настроение или осознанность).

Именно эта особенность современного трекинга позволяет говорить уже не о количественной (quantified), а качественной (qualified) самости [Swan, 2015: 93]. Качественная самость заключается в использовании количественных измерений для описания качественных аспектов жизни — таких как настроение или продуктивность. Это упрощает пользователям задачу внесения данных. Например, можно дать субъективную оценку какой-либо из характеристик по шкале от 1 до 10 вместо текстового описания собственного состояния. Следовательно, открывается возможность для более точной аналитики. Более того, качественная самость отражается в возможности индивида стать лучшей версией себя, на что зачастую и бывает направлен селф-трекинг. Для этого, в свою очередь, необходимо представить пользователям инструментов цифрового селф-трекинга качественную обратную связь и рекомендации о возможных направлениях изменения поведения.

В рамках и количественной, и качественной самости селф-трекинг становится инструментом познания и, впоследствии, улучшения себя. Таким образом, метрики, полученные с помощью инструментов трекинга, воспринимаются в качестве источника формирования такого знания пользователей о себе: в процессе самосовершенствования смешается фокус от интуиции и самоощущения к объективным данным, предоставляемым трекерами [Crawford, Lingel, Karppi, 2015: 492]. Стоит заметить, что имеющиеся технологии несовершенны и, несмотря на заявленную точность, обладают погрешностью, в связи с чем возникает вопрос о доверии к собранным данным.

¹ McCarthy J. One in Five U.S. Adults Use Health Apps, Wearable Trackers. *Gallup*. 2019. 11 December. URL: <https://news.gallup.com/poll/269096/one-five-adults-health-apps-wearable-trackers.aspx> (дата обращения: 27.01.2021).

Несмотря на существование этой проблемы, большинство пользователей признают трекеры и приложения полезными для достижения целей, связанных с улучшением или поддержанием состояния здоровья². Более того, инструменты цифрового трекинга берут на себя часть ответственности за развитие пользователя: напоминают о задачах и тренировках или даже помогают сделать необходимую для пользователя активность более увлекательной с помощью инструментов геймификации (визуализации прогресса, системы вознаграждений или уровней) [Whitson, 2013: 169]. В этом отчасти заключается эмансипирующий или, другими словами, освобождающий дискурс, в терминах которого описывается селф-трекинг.

Тем не менее существует и другой, противоположный эмансипирующему дискурс — критический. В его рамках ключевую роль играют идеи Мишеля Фуко о субъективации и техниках себя.

В процессе субъективации у индивида развивается внутренняя связь с самим собой, конструируется отношение к себе, с помощью которого формируется его «Я» как субъекта действия [Фуко, 1998]. Техники себя, в свою очередь, определяются М. Фуко как действия, которые индивид совершает над собой, своим телом, мыслями и поступками для того, чтобы изменить себя [Савельева, 2015: 7].

Таким образом, в процессе субъективации человек самостоятельно определяет, в некотором роде «познает себя», а также обретает способность изменять себя в соответствии с тем, как именно он хочет себя определить. Учитывая это, мобильные приложения и устройства для селф-трекинга можно назвать инструментом субъективации и, следовательно, «техникой себя»: гаджеты для селф-трекинга (такие как фитнес-браслеты) и мобильные трекинг-приложения позволяют получить достаточно полное представление о некоторых аспектах своей жизни (будь то частота пульса или спортивные достижения). Более того, если потребуется, они помогают пользователям изменять свое поведение. Например, такие приложения могут напоминать пользователю о необходимости проявлять физическую активность или поддерживать водный баланс.

Субъективация в цифровом селф-трекинге происходит через опору пользователя на данные и цифровые показатели, предоставляемые трекером. Квантifiцируя пользователя, трекер объективирует его, представляя в виде совокупности показателей — хотя, как уже было отмечено ранее, существует недоверие к объективности рассчитываемых показателей. Сама же субъективация в таком случае может происходить, если пользователь самостоятельно использует предоставленные данные для самоопределения и дальнейшего преобразования себя. Расплывчатая граница между объективацией («определение» пользователя трекером) и субъективацией («определение» пользователем самого себя с помощью трекера) заставляет задуматься о значимости активной роли пользователя в цифровом селф-трекинге.

Селф-трекингу присущи и некоторые ключевые характеристики техник себя. В прошлом формированию «самости» во многом способствовало социальное окружение, поскольку для процесса субъективации необходимо было сущест-

² McCarthy J. One in Five U.S. Adults Use Health Apps, Wearable Trackers. *Gallup*. 2019. 11 December. URL: <https://news.gallup.com/poll/269096/one-five-adults-health-apps-wearable-trackers.aspx> (дата обращения: 27.01.2021).

вование некоторого, хотя бы виртуального «другого», в качестве которого могут выступать как отдельные индивиды, так и институты [Голенков, 2007: 63]. Это отражается и во многих аспектах цифрового селф-трекинга: практически все устройства и приложения позволяют не только видеть результаты других пользователей, но и делиться своими собственными данными с друзьями в социальных сетях, что лишний раз стимулирует пользователей достигать поставленной цели.

«Техники себя» и процесс субъективации раскрывают проблематику властных отношений и дисциплинарных практик. С одной стороны, самостоятельные и свободные действия индивида предполагают внутренний аспект субъективации, который заключается в усилиях самого индивида, направленных на получение знания о себе. С другой стороны, в процессе субъективации задействованы и механизмы контроля над индивидом, точнее — дисциплинарные практики. Инструменты цифрового селф-трекинга в некотором роде дисциплинируют пользователей, отчасти даже получают власть над привычками и действиями индивидов. Примером дисциплинарных практик, связанных с цифровым селф-трекингом, может считаться, в частности, система оповещений, включенная практически во все мобильные приложения, в том числе и те, которые нацелены на самопознание и саморазвитие пользователей.

Рассмотренные ранее аспекты еще раз доказывают, что цифровой селф-трекинг и его инструменты существуют и действуют на пересечении двух противоречащих друг другу дискурсов. С одной стороны, эти механизмы освобождают пользователей, предоставляя им новые, расширенные возможности для саморазвития и самопознания. С другой стороны, такие устройства подчиняют пользователей, навязывая систему дисциплинарных практик.

Роль пользователя в цифровом селф-трекинге

Учитывая то, что цифровой селф-трекинг — явление многоаспектное, важно понимать, каким образом его инструменты могут влиять на пользователя. Поскольку селф-трекинг воплощает в себе концепцию количественной (и неразрывно с ней связанной качественной) самости, предполагается активное и добровольное участие пользователя в процессе самонаблюдения, то есть сотрудничество с инструментами трекинга [Berson, 2015: x].

Вероятно, именно поэтому фигуре пользователя уделяется большое внимание в существующих исследованиях цифрового селф-трекинга. Например, Д. Кристенсен и М. Рукенштейн вводят концепцию «лаборатории самонаблюдения», которая рассматривает технологию селф-трекинга как источник ресурсов и возможностей для саморазвития пользователя [Kristensen, Ruckenstein, 2018: 3626]. Другая важная идея в рамках взаимодействия человека и селф-трекинга — концепция «возникающего себя», которая заключается в том, что персональная аналитика данных во многом сформирована самими технологиями селф-трекинга, способными предопределять идеи и практики пользователей [Ruckenstein, 2014: 70]. Кроме того, в рамках исследований отношений между пользователями и трекерами отмечаются разные роли, которые могут играть инструменты трекинга в восприятии индивидов. В зависимости от целенаправленности и значимости для пользователя трекеры могут восприниматься пользователями в качестве наставника, инструмента или игрушки [Lyall, Robards, 2018: 113].

Существующие исследования позволяют получить достаточно полное представление о цифровом селф-трекинге и его особенностях, но один важный аспект этого явления остается неизученным. Инструменты цифрового селф-трекинга предоставляют пользователям широкие возможности для саморазвития, но при этом «вторгаются» в их жизнь и навязывают дисциплинирующие практики. Именно на последних и фокусируется настоящее исследование. В нем мы подробно рассмотрим становление дисциплинарных практик и критерии их эффективности. Также внимание будет уделено тому, как пользователи определяют для себя границу между возможностями для саморазвития, которые предоставляют инструменты трекинга, и контролем со стороны трекеров и приложений. Наконец, будут рассмотрены варианты взаимодействия между человеком и инструментами селф-трекинга, выделяемые в рамках разного отношения пользователей к контролю со стороны гаджетов и приложений.

Методы исследования

Ключевой вопрос исследования — определение границы между контролем и эманципацией, которую выстраивают пользователи в процессе взаимодействия с продуктами цифрового селф-трекинга. Основной метод сбора данных — полуформализованные интервью с владельцами различных трекеров (см. Приложение). Часть используемых устройств измеряют более «качественные» сферы повседневной жизни, отражая экспансию, о которой говорилось выше.

На первом этапе двухступенчатой выборки с помощью онлайн-опроса была собрана ключевая информация о потенциальных участниках интервью: социально-демографические характеристики (пол, возраст), используемые продукты селф-трекинга (название, опыт использования). Всего в опросе приняло участие 63 респондента. По результатам были выявлены три основные группы пользователей: те, кто использует только трекеры (фитнес-браслеты, умные часы); те, кто использует исключительно мобильные приложения; а также те, кто использует оба типа продуктов цифрового селф-трекинга. Именно эти группы легли в основу рекрутинга респондентов для второго этапа исследования, а именно интервьюирования. Еще одним важным элементом выборки стало достижение гетерогенности в каждой из перечисленных групп по полу, возрасту (младше 30 лет, 30—45 лет, старше 45 лет) и опыту использования (до 3 месяцев, от 3 месяцев до года и более 1 года). Всего на втором этапе было проведено 18 полуформализованных интервью.

В процессе анализа данных мы опирались на процедуры построения обоснованной теории. Было проведено кодирование в несколько этапов. На первом этапе выделены тематические фрагменты интервью, связанные с отдельными анализируемыми аспектами селф-трекинга (эмоции, контроль). На втором этапе было произведено первичное кодирование данных каждым участником исследовательского коллектива. Важно отметить, что первичное кодирование каждый исследователь осуществлял независимо, что позволило обеспечить методическую триангуляцию. На заключительном этапе первичные коды концептуализировались и интерпретировались в кодах последующих уровней совместно всем исследовательским коллективом.

В данной статье будут подробно рассмотрены три аспекта цифрового селф-трекинга: дисциплинарные практики, соотношение дискурсов контроля и эмансипации, специфика взаимодействия пользователя и продуктов трекинга в контексте динамики пользовательского опыта.

Результаты

1. Как пользователи переживают ситуацию контроля

Один из аспектов цифрового селф-трекинга — столкновение двух противоречящих друг другу дискурсов: дискурса контроля и дискурса эмансипации. В этой части анализа будут выявлены и описаны основные механизмы реализации данного противоречия непосредственно в практиках цифрового селф-трекинга, а также условия, при которых эти механизмы присутствуют в опыте пользователя.

Большая часть наших информантов была склонна отрицать наличие контроля со стороны используемого селф-трекера. На этом тема насаждаемого внешнего контроля продуктами цифрового селф-трекинга могла бы быть закрыта из-за отсутствия такового в опыте пользователя. Однако в интервью часто фигурировал сюжет давления и принуждения, которые на них оказывают продукты трекинга:

...Я чувствую на себе какое-то давление, что я сегодня ничего не делаю, вообще халевщица сплошная. И это заставляет переживать... (Р2, жен., 19 лет)

Наверное, это немного дисциплинирует. Они заставляют меня лично больше пройти, больше проехать и так далее. (Р17, муж., 50 лет)

Все это свидетельствует о том, что приложения активно используют дисциплинарные практики, направленные на установление контроля над деятельностью и привычками человека. И как мы видим из цитат респондентов, приложения действительно оказывают воздействие на поведение человека, принуждая его действовать определенным образом. Тогда почему сами пользователи не идентифицируют такое воздействие и влияние в терминах контроля и вмешательства?

В этой ситуации работает механизм легитимации контроля со стороны приложения или трекера. Пользователи оправдывают наличие такого рода вмешательства в их повседневную жизнь и деятельность. На данных нами были выделены два основания, которые позволяют пользователям оправдывать применение дисциплинарных практик трекерами: *объективность* и *забота*.

1.1. Достоверность

Первое основание легитимности контроля — «*объективность*» продуктов цифрового селф-трекинга. Тема достоверности данных, полученных с помощью трекера, уже поднималась в литературе [Crawford, Lingel, Karppi, 2015]. Данные трекера воспринимаются как источник более точного знания о себе. Несмотря на то, что некоторые респонденты замечают погрешности в измерениях, эти данные кажутся им более четкими, нежели самоощущения. Как следствие, происходит сдвиг от интуиции и самоощущения к «*объективным*» данным трекера:

Пока статистику не соберешь, самому себя трудно оценить... (Р6, муж., 56 лет)

Ну, они скорее просто предоставляют мне плюс-минус объективную картину. (Р8, муж., 23 года)

Из-за того, что люди не способны с точностью трекера измерять показатели, гаджет становится единственным релевантным источником достоверных данных, что наделяет его особой властью во взаимоотношениях «пользователь — трекер». Под этим предлогом контроль деятельности со стороны продукта в глазах пользователей становится оправданным.

Стоит отметить, что «объективность» данных также влияет на уровень удовольствия и радости от соответствия заданным нормам. Достижение абстрактной нормы разнообразных метрик заставляет пользователя чувствовать гордость за себя, что объясняется восприятием этих данных как достоверных и точных.

У нас была подготовка к важному мероприятию в рамках проекта, и я как бы мог работать день, мог работать ночь, потом работать день, и там набегало большое количество часов, я был собой очень горд и не очень хорошо себя физически чувствовал, но я был невероятно горд. (Р8, муж., 23 года)

Таким образом, достоверность становится основанием, которое позволяет продуктам селф-трекинга устанавливать контроль над деятельностью человека и насаждать свою «волю» пользователям. Более того, достоверность трекера усиливает эмоциональные переживания пользователя как в положительном, так и в отрицательном ключе.

1.2. Забота

Второе основание, которое позволяет трекеру вмешиваться в повседневность пользователя, — постулирование заботы о нем. Трекер становится «заботливым наставником», который направляет человека на правильный путь, помогает избежать негативных последствий «неправильного» образа жизни, рекомендуя придерживаться конкретных норм. И пользователи действительно убеждены, что такое вмешательство трекера благоприятно влияет на их самочувствие и здоровье:

*Ну не то какой я должна быть, а то, что есть некие **ориентиры** по активности, и как бы в среднем лучше их соблюдать, чтобы быть здоровым. А если я их регулярно не соблюдаю, то я в определенный момент столкнусь с последствиями: плохой физической формой, лишним весом, каким-то не очень хорошим самочувствием, да каким-то недовольством собой. (Р18, жен., 32 года)*

Респонденты верят, что наличие трекера и соответствие его «нормам» позволит им улучшить свою жизнь, здоровье и самочувствие. Такие положительные эффекты, с позиции пользователя, оправдывают наличие вмешательства и контроля извне. В данном случае забота тесно сопряжена с темой здоровья. От пользователей лишь требуется соответствовать нормам, «насаждаемым» устройством.

Трекер заботится также и о том, чтобы пользователь выполнял взятые на себя обязательства. В данном случае сами нормы уже не «насаждаются» устройством, пользователь выбирает их сам (обычно они превосходят автоматически заданные). Теперь трекер становится «вежливым надзирателем», который следит, чтобы пользователи выполняли свои ежедневные задачи. Если этого не происходит, он мягко, но настойчиво напоминает пользователю о них:

...Приложение тебе говорит, что ты вроде планировал столько-то спать в день, но последние две недели столько не спал, а потом начинаешь удивляться, почему какая-то вялость, утомляемость постоянно и так далее. (Р6, муж., 56 лет)

Таким образом, забота о пользователе в разных ее проявлениях позволяет легитимировать контроль и вмешательство в его жизнь. Субъективно воспринимаемое положительное воздействие оправдывает деятельность селф-трекера.

Из всего вышеописанного можно заключить, что контроль легитимируется за счет определенных характеристик и функций трекеров, которые рассматриваются пользователями как исключительно благоприятные и полезные для них самих. Именно они позволяют оправдать вмешательство трекера в жизнь человека. В случаях легитимации пользователи предпочитают не замечать факты принуждения и давления со стороны трекеров, оправдывая их разговорами о полезности и заботе.

1.3. Избегание контроля при отсутствии легитимации

Не все пользователи готовы игнорировать присутствие внешнего контроля на основаниях объективности и заботы. Это может быть связано с тем, что факторов общей полезности использования продуктов оказывается недостаточно для оправдания проникновения в личное пространство пользователя. Как пользователь оценивает вмешательство трекеров при отсутствии легитимации?

Ощущение внешнего контроля в этих условиях воспринимается болезненно. Острота восприятия зачастую связана с возникающим ощущением внутренней злости по отношению к источнику дискомфорта. Об этом свидетельствуют очень яркие и наполненные эмоциями высказывания информантов:

...Просто приложения всегда страдают какими-то очень надоедливыми push-уведомлениями, когда они пишут: «вы уже два дня не медитировали, срочно помедитируйте и повысьте свой уровень». И ты думаешь, нет, отстань от меня, тупая машина. (Р13, жен., 29 лет)

Чрезмерное «вмешательство» в повседневную жизнь человека воспринимается достаточно агрессивно и приносит пользователю негативные эмоции. Ощущение контроля начинает доставлять пользователям дискомфорт, от которого хочется избавиться. Для этого пользователи могут использовать разные способы. На наших данных были выделены две стратегии ухода от излишнего принуждения со стороны трекеров.

Первый вариант — отказ от использования гаджета. Такой способ чаще всего применялся по отношению к трекерам, которые, по мнению пользователей, уста-

навливали чрезмерный контроль над деятельностью человека, вызывая ощущение тревоги и угнетенности:

Вот после того, как начала — негативные, мне кажется, потому что это отнимало время, я чувствовала себя каким-то параноиком, который каждую еду взвешивает и записывает, мне это не очень нравилось. (Р15, жен., 26 лет)

Когда такой тотальный контроль становился невыносимым, люди принимали решение отказаться от трекеров. Интересно, что чаще всего в таком ключе описывали опыт использования приложений для подсчета калорий.

Второй стратегией становится изменение личного подхода к реализации «техники себя». Этот вариант используется, когда предложенные трекерами суточные нормативы оказываются слишком высокими и недостижимыми. Осознание невозможности соответствовать завышенным требованиям трекера демотивирует пользователей. В такой ситуации выбор совершается в пользу прекращения попыток полностью соответствовать нормативам. Люди продолжают использовать трекер, но становятся более снисходительны к себе, не требуя ничего сверх собственных возможностей.

Я думала: «Ну как же так?». Винила себя. Вначале мы обсуждали, что была зависимость небольшая. Но сейчас хватку я подрасслабила и решила подобнее к себе относиться. Поэтому сейчас я думаю: «Не смогла, так не смогла». (Р1, жен., 32 года)

Этот вариант позволяет избежать чрезмерных переживаний от несоответствия нормам, установленным трекерами. В рамках данной стратегии многие перестают обращаться к приложениям ежедневно и трансформируют практику их использования в эпизодическую. После такой трансформации люди начинают «проще» относиться к замечаниям приложения, часто просто игнорируя постоянные уведомления и напоминания.

В рамках этой стратегии пользователи провозглашают свою «независимость» от трекера, которая проявляется в самостоятельности принятия решений. Создается своеобразная зона свободы действия, которая позволяет снизить силу давления и внешнего контроля. Пользователь «дистанцируется» от трекера.

Рассмотренные стратегии хоть и способствуют минимизации переживаний и ощущения контроля, но также минимизируют или вовсе ликвидируют само взаимодействие между пользователем и продуктом селф-трекинга.

Таким образом, «эмансипирующий» дискурс тесно связан с «дисциплинирующим». Именно оценка пользователем соотношения этих дискурсов в деятельности трекера определяет, как будет восприниматься дисциплинирующий аспект трекинга и какие практики будут использованы.

Если пользователи высоко оценивают освобождающую силу используемого продукта, его полезность и важность в своей жизнедеятельности, в восприятии трекера превалирует «эмансипирующий» дискурс. Именно при таком восприятии может срабатывать механизм легитимации контроля. Пользователи рассматривают функции продукта как исключительно благоприятные и полезные для них самих, что и становится основанием для легитимации.

Если же в восприятии пользователей доминирует критический дискурс, они склонны воспринимать внешний контроль и вмешательство. Именно здесь могут возникать стратегии, призванные минимизировать ощущение надзора.

Важно отметить, что в процессе использования соотношение дискурсов в восприятии пользователя может меняться. Трансформация соотношения может быть связана с особенностями деятельности трекера, которые начинают восприниматься как излишнее вмешательство или, наоборот, оцениваться как исключительно полезные. Таким образом, в опыте пользователя могут быть представлены оба варианта соотношения и, соответственно, разные механизмы переживания ситуации контроля.

2. Типология дисциплинарных практик в цифровом селф-трекинге

Одна из задач настоящего исследования — поиск дисциплинарных практик, присущих инструментам цифрового селф-трекинга. Такие практики, в сущности, представляют собой механизмы контроля, задействованные в процессе субъективации, то есть в процессе определения индивидом самого себя. Выделение блока механизмов контроля и его влияния на пользователя — ключевой элемент для определения того, куда все-таки склоняется баланс между эмансинацией и контролем.

Для определения дисциплинарных практик, существующих в инструментах цифрового селф-трекинга, из базы данных были выделены те фрагменты интервью, в которых речь так или иначе шла о конкретных инструментах, побуждающих пользователя к регулярному использованию приложений или устройств (браслетов, умных часов) для отслеживания собственных характеристик. После этого мы (1) составили подробные схемы классификации дисциплинарных практик, а также (2) выделили факторы позитивного/негативного восприятия применительно к некоторым из указанных механизмов.

2.1 Классификация дисциплинарных практик

По результатам проведенного анализа мы выделили 15 дисциплинарных практик, некоторые были объединены в группы. Классификация представлена на рис. 1.

Приведем в пример одну из самых распространенных дисциплинарных практик — уведомления, которые регулярно напоминают пользователям о приложении:

Тебе «Здоровье» пишет, что ты, там, сегодня не ходил, и ты такая «угу, надо проявить какую-то активность». (Р2, жен., 19 лет)

Заслуживает отдельного пояснения такой инструмент, как «полоса». В сущности, она представляет собой обозначение в приложении непрерывности его ежедневного использования:

Когда видишь, что количество дней подряд растет — это приятно. (Р11, жен., 30 лет)



Рис. 1. Классификация дисциплинарных практик

Такой механизм контроля зачастую направлен• именно на стимуляцию ежедневного использования инструмента цифрового трекинга и, следовательно, на возведение действия, отслеживаемого пользователем, в разряд привычки.

Следующая дисциплинарная практика — сопровождение — заключается в активной роли инструмента селф-трекинга в процессе выполнения пользователем отслеживаемого действия. Она характерна в основном для трекинга физической активности (например, тренировок или бега). Сам инструмент трекинга в рамках данного механизма зачастую «очеловечивается» пользователями:

Дамочка, которая озвучивает, что ты должен на данный момент сделать. (Р2, жен., 19 лет)

Оценивание как дисциплинарная практика включает в себя все проявления замечаний оценочного характера со стороны инструмента селф-трекинга — как позитивных, так и негативных:

Ты такой: «Ура, меня там похвалила компьютерная штука». (Р3, муж., 27 лет)

Награды как дисциплинарная практика представляют собой некоторую систему поощрений, существующую в приложениях для трекинга. Это может быть система достижений, которую пользователи открывают в процессе использования трекеров (медали, специальные значки, которые становятся доступны пользователям лишь после выполнения определенных действий), или измеримые поощрения пользователей (монеты, бриллианты, которые пользователи могут использовать для разблокировки возможностей внутри приложения).

Уровни как дисциплинарная практика, в свою очередь, отражают возможность пользователя открывать новые ступени, достигать их или повышать свой уровень в процессе использования инструмента для трекинга. Эта практика может быть

связана с очками опыта как наградой, присваиваемой пользователям за выполнение определенных действий.

Стоит отдельно остановиться на особенностях «цели» как одного из типов дисциплинарных практик. Согласно идеям движения Quantified Self, измерение показателей активности и их отслеживание должно быть не самостоятельной целью, а шагом на пути к самосовершенствованию [Ajana, 2017: 6]. Учитывая это, неудивительно, что постановка цели в рамках использования инструментов цифрового селф-трекинга — одна из самых распространенных дисциплинарных практик, выделяемых пользователями.

Важно отметить, что люди более мотивированы достичь цели, если она была постановлена ими самостоятельно. Если же цель установил трекер, у пользователей возникают фрустрации и чувство дискомфорта, связанные с непониманием механизма расчета желаемых показателей.

...Бред такой вообще невероятный, на мой взгляд. То есть усредненные показатели среднестатистического человека со среднестатистическим весом, среднестатистическим пульсом и со среднестатистическим образом жизни. Настолько усредненные цифры, что они ни о чем вообще. (Р4, муж., 36 лет)

Также информанты выделяли дисциплинарные практики, которые были объединены в группы, — визуализация и социальные механики³. В первую были включены все практики, в рамках которых инструменты селф-трекинга наглядно предоставляли пользователям информацию об их активности, а во вторую — все практики, связанные с наличием другого актора.

2.2 Факторы позитивного или негативного восприятия дисциплинарных практик

Рассмотренные выше дисциплинарные практики, заключенные в функционал селф-трекинговых приложений, воспринимаются пользователями по-разному. По этой причине стоит отдельно осветить те факторы, от которых зависит восприятие пользователем ситуации вмешательства трекера как скорее положительного или скорее отрицательного опыта. Положительные факторы отмечены знаком +, отрицательные — знаком -.

+Интенция

Чаще всего пользователи, исходя из своих индивидуальных целей, устанавливают напоминания самостоятельно, поэтому они не воспринимаются как вторжение в их личную жизнь. Иными словами, пользователи выделяют приложению легитимную сферу контроля, предоставляя согласие на вмешательство. В таком случае напоминания воздействуют на них положительно, побуждая их к действию.

³ Термин «социальные механики» используется нами в том смысле, которым оперируют исследователи игр, игровых механик и геймификации. Социальные механики, таким образом, определяются нами как категория инструментов геймификации, в рамках которой для вовлечения пользователя в процесс селф-трекинга используются возможности взаимодействия с другими людьми.

+Визуальность

Пользователи отмечали положительные эмоции, сопряженные с выставлением галочки при достижении поставленной цели. Аналогично, положительному восприятию сопутствует наглядная графика, позволяющая соревноваться с самим собой, прослеживать динамику результатов во времени. В особенности это актуально для спортивных показателей, поскольку с ними могут быть сопряжены более глобальные цели: увеличить километраж, повысить скорость бега и т. д. В таком случае важно не только достижение ежедневной цели (например, 10 тысяч шагов), но и определенная положительная динамика результатов. Кроме того, визуальная составляющая в виде непрерывной линии медитации мотивировала пользователей неспортивных приложений. В данном случае важна регулярность осуществления вида деятельности, а не ее результативность. Помимо связи с целью, такой вид подачи также более прост для восприятия по сравнению с числовым выражением:

Вот еще функция мотивации — это не в виде каких-то мотивирующих цитат, это в виде просто наглядной графики, где я понимаю, на каком свете я, грубо говоря, нахожусь.
(P8, муж., 23 года)

+Подробность

С удобством и возможностью сравнивать себя с самим собой также соотносится следующая особенность — подробность биометрических данных. Более детальные данные могут открывать большие возможности для отслеживания своего состояния и сопоставления результатов с ним. Благодаря этому пользователи могут делать выводы, каким образом те или иные факторы влияют на измеряемые показатели и, соответственно, вносить корректизы в свою жизнь для достижения лучших результатов:

...В приложении более удобно, по часам, по неделям, по дням. А все, что записывают часы, там можно посмотреть просто на более большом экранчике телефона. (P10, жен., 31 год)

+Рефлексия

Следующий фактор, способствующий положительному восприятию дисциплинарных практик, связан с квантификацией телесности. Получая количественные данные о себе, человек делает выводы и совершает действия. Важно, что выводы не навязываются приложением, а делаются самостоятельно. Рефлексия тесно переплетается с осознанным видом отношений между пользователем и продуктом селф-трекинга. Решения принимает пользователь, он же способен скорректировать нормы, исходя из персонализированных целей и своих возможностей. В таком случае человек сохраняет определенную независимость от устройства, которая не исключает следование рекомендациям трекера и характеризуется способностью пользователя к рефлексии. Такое отношение, как правило, положительно коррелирует с восприятием трекера, поскольку пользователь ощущает меньший контроль со стороны приложения или устройства:

...Должен был заниматься 1000 часов, получилось 200 часов, окей. Соответственно, я понимаю, либо нужно планку понижать, либо пересматривать свое отношение... (Р8, муж., 23 года)

–Завышенные требования

Завышенные требования со стороны приложения вызывают у пользователя разочарование от несоответствия им и, как следствие, могут привести к избеганию контроля и негативному восприятию вмешательства трекера. Человеку важно соответствовать нормам или уровню других пользователей приложений, чтобы не подвергаться осуждению и чувствовать себя субъективно близким к некоторому идеалу. Сам по себе это мотивирующий фактор, положительно сказывающийся на результативности и вовлеченности индивида. В случае же невозможности соответствия пользователь испытывает разочарование и демотивацию:

А когда с тебя, не знаю, уже десятый литр пота стекает и тебе машинный голос говорит: «Давай, еще чуть-чуть, пожалуйста». (Р2, жен., 19 лет)

–Частота

Частота получения уведомлений напрямую связана с эмоциональными аспектами переживания опыта использования. Если приложение не напоминает о себе слишком часто, то оно не становится объектом негативных эмоций и переживаний пользователя. И, напротив, назойливость трекера, не легитимированная самим пользователем, как правило, воспринимается как вмешательство и контроль над повседневностью.

3. Динамика отношений между пользователем и трекером

Рассмотренные ранее способы избегания и легитимации контроля со стороны инструментов цифрового селф-трекинга играют также значительную роль в определении отношений, которые складываются между пользователем и трекером. Эти отношения, в свою очередь, способны опосредовать пользовательский опыт, влияя на его динамику. В рамках раздела мы рассмотрим три вида отношений, возникающих между пользователями и трекерами, и то, как они связаны с динамикой пользовательского опыта.

Первый тип отношений между пользователем и инструментами трекинга — зависимость. Эти отношения формируются во многом на основании непрерывного, ежедневного использования трекеров, и, как следствие, инструменты отслеживания показателей начинают восприниматься как неотъемлемая часть жизни:

...В начале использования было что-то вроде **зависимости**. Я точно проходила каждый день... (Р1, жен., 32 года)

Пользователи при этом становятся зависимыми как от самого трекера, о чем свидетельствует непрерывность их использования, так и от его влияния на их жизнь. Речь идет о том, что люди зацикливаются на цифрах, предоставляемых устройствами, и стремятся к ежедневному достижению установленных трекерами норм. Пользователи, находящиеся в отношениях зависимости, обычно готовы

прикладывать дополнительные усилия для достижения ежедневных целей, порой корректируя для этого свои ежедневные планы.

Другая важная характеристика отношений зависимости — усиление эмоциональных переживаний пользователя в случае несоответствия требованиям трекеров. В этой ситуации люди наделяют особым значением достижение ежедневных целей и несоответствие становится для них травматическим событием, которое усиливается переживаниями. Пользователи начинают винить себя, буквально испытывать страдание:

...Когда я не добираю, я начинаю страдать... (Р16, жен., 54 года)

В рамках отношений зависимости пользователи в наибольшей степени склонны легитимировать деятельность продуктов селф-трекинга, описывая ее в терминах пользы, уделяя внимание объективности предоставляемых данных и заботе устройств об их благополучии (два ключевых основания легитимации). Таким образом, чтобы человек легитимировал деятельность трекера, необходим вариант зависимости, который наделяют продукты цифрового селф-трекинга особым значением и особой позицией во взаимоотношениях.

Важно отметить, что такие отношения между пользователями и продуктами трекинга могут присутствовать только в начале использования, а могут и перерастать в постоянные. В ситуации, когда зависимость возникает только в начале использования, после ее исчезновения снижается и частота обращений к трекеру, ведь со временем пропадает эффект новизны, который также может способствовать интенсивному и постоянному отслеживанию показателей:

Вначале, наверное, я чаще смотрела на метрики. (Р7, жен., 29 лет)

Оказавшись в отношениях зависимости с трекером, пользователь, как уже отмечалось ранее, может ощущать настолько сильные переживания, связанные с завышенными требованиями устройств, что неспособность легитимировать контроль может привести к полному прекращению использования трекера или, как минимум, к сильной фruстрации:

А он мне потом пишет, что вы сегодня двигаетесь больше, чем обычно. Кто его спрашивал? Все, на этом все наши отношения закончились, разумеется. (Р12, жен., 55 лет)

С другой стороны, отношения зависимости со временем могут перерasti в отношения сопровождения (которые, конечно, могут присутствовать отдельно от отношений зависимости). В рамках отношений сопровождения взаимодействие между пользователем и трекером практически отсутствуют: трекинг уходит на задний план, перестает «существовать» в действительности человека. Продукты селф-трекинга лишь изредка напоминают о себе посредством уведомлений:

То есть для меня это, как я сказал, не более чем красивая цифра, которая как бы периодически под вечер мигает, что цель достигнута. (Р8, муж., 23 года)

Возможны несколько ситуаций, при которых взаимодействие между трекером и пользователем могут перерасти в отношения сопровождения. Во-первых, такой вариант отношений может сформироваться, когда у пользователей после некоторого периода активного использования продуктов трекинга возникает понимание того, какими физически ощущаются те или иные значения отслеживаемых показателей. Благодаря самоощущению пользователи способны самостоятельно, без вмешательства трекера, регулировать свою деятельность:

Сейчас-то я, в общем-то, наверное, даже без трекера примерно знаю по ощущениям, сколько я шагов примерно делаю в день. (Р9, муж., 54 года)

Во-вторых, отношения сопровождения могут возникать в ситуации, когда пользователь достигает поставленной цели, и полученный результат остается лишь поддерживать — для этого также не нужен плотный контроль со стороны инструментов трекинга.

Поскольку в рамках отношений сопровождения трекер выступает в качестве «безмолвного наблюдателя», контроль и вмешательство с его стороны практически сходят на нет, также исчезают и эмоциональные переживания пользователя в ситуации несоответствия требованиям устройства. Трекер напоминает о себе только тогда, когда пользователю удается достичь дневной нормы, а до этого напоминания в реальности пользователя устройства не существует. Таким образом, в отношениях сопровождения гаджет приносит человеку исключительно положительные эмоции: пользователь вспоминает о трекере при получении уведомления, оно носит приятный характер, доставляя индивиду истинное удовольствие от осознания соответствия заданным требованиям.

Наконец, отношения между пользователем и инструментами трекинга могут быть описаны в терминах осознанности. Такие отношения характеризуются способностью пользователей к рефлексивному осмысливанию деятельности устройства и оценке его влияния на свою жизнь. Люди не слепо следуют указаниям трекера, а оценивают и корректируют их, исходя из своих собственных представлений и целей. В рамках таких отношений пользователи самостоятельны и независимы от продуктов селф-трекинга:

...Сама лучше знаю, как мне, походить мне или не походить, поспать или не поспать. (Р10, жен., 31 год)

В осознанном стиле отношений происходит некоторое дистанцирование между пользователем и трекером. Индивид создает себе пространство свободы, вынося влияние и деятельность продуктов селф-трекинга за его пределы. К рекомендациям пользователи относятся скептически и соотносят их как со своим состоянием, так и со своими целями. Тем самым они оставляют себе право принятия решения.

Поскольку пользователи создают себе пространство независимости, то и переживания от несоответствия минимизируются:

Если вдруг я пропускаю тренировку, то я просто пропускаю тренировку без всяких угрызений совести. (Р7, жен., 29 лет)

Такая минимизация эмоциональных переживаний связана, прежде всего, именно с осознанностью. Пользователи понимают, почему они не смогли выполнить норму, и связывают это с объективными причинами (например, из-за плохой погоды или загруженного дня). Вероятно, именно поэтому возникновение отношений осознанности может приводить к снижению частоты использования продуктов цифрового трекинга: внешние факторы — сезонность использования или занятость — часто становятся причиной таковых изменений, а отношения осознанности позволяют пользователям не прикладывать дополнительных усилий для преодоления обстоятельств. Не только объективные причины оправдывают неактивность, но и сами пользователи позволяют себе расслабиться и отдохнуть:

Во время ЕГЭ я его практически не использовала, потому что как бы человек не спал, а еще надо отжиматься. Было немножко лениво, сейчас вот стараюсь использовать.
(Р2, жен., 19 лет)

Важно отметить, как именно пользователи приходят к осознанному отношений. Осознанный формат отношений может устанавливаться с самого начала применения продуктов селф-трекинга, а может быть реакцией на контроль со стороны устройства. Когда люди сталкиваются с невозможностью соответствовать высоким требованиям трекера, они ищут возможные способы выхода из этой ситуации, что может приводить к изменению их подхода к реализации «техник себя», о чем уже упоминалось ранее. Именно в рамках таких отношений пользователи провозглашают свою «независимость», призванную минимизировать эмоциональный дискомфорт.

Таким образом, в зависимости от роли контроля со стороны продуктов трекинга, выделяются три формы отношений, возникающих между пользователем и трекером: отношения зависимости, осознанности и сопровождения. Эти формы могут возникать на разных этапах трекинга, формируясь изначально или перетекая друг в друга по определенным причинам, одной из которых можно считать ощущение внешнего контроля, которое заставляет пользователя дистанцироваться от продуктов селф-трекинга. Более того, отношения между пользователями и продуктами трекинга могут опосредовать динамику пользовательского опыта. Например, отношения осознанности могут привести к постепенному снижению частоты использования, а отношения зависимости — к полному отказу от приложения или гаджета, поскольку пользователи могут не справиться с серьезными эмоциональными переживаниями от несоответствия требованиям устройств.

Заключение

Лейтмотивом данной работы стал теоретический и эмпирический анализ того, каким образом пользователи инструментов цифрового селф-трекинга определяют границу между контролем со стороны приложения и освобождением, которые такие устройства могут предоставить.

В процессе исследования были выделены три формы отношений между трекером и пользователем: зависимость, осознанность и сопровождение. Зависимый

трекинг проявляется в восприятии трекера как неотъемлемой части своей жизни, что приводит к его непрерывному и ежедневному использованию. При этом усиливается значение эмоциональных переживаний, которые пользователь испытывает при невыполнении рекомендаций устройства. Осознанный трекинг предоставляет пользователю большую степень свободы и независимости. В рамках сопровождения отношения между пользователями и устройствами для отслеживания минимизируются, а процесс трекинга проходит в фоновом режиме.

Восприятие трекера как заботливого наставника и внешнего надзирателя позволяет сделать вывод о том, как пользователи обозначают границу между контролем со стороны инструментов трекинга и «освобождением от ответственности», предоставляемым устройствами. Когда человек действительно ощущает контроль со стороны приложений, взаимодействие становится эмоциональным, болезненным, и пользователь стремится его избегать. В ситуации сверхконтроля со стороны трекера это часто приводит к отказу от гаджета или приложения. Пользователи также могут избегать контроля со стороны устройств трекинга, иначе реализовывая «техники себя». В этом случае люди прекращают попытки соответствовать ожиданиям трекеров и прибегают к оправданиям того, почему им не удается выполнять установленные нормы.

Однако чаще всего пользователи легитимизируют контроль со стороны приложения, описывая его как некоторое оправданное, необходимое давление. Именно подобное вмешательство в их жизнь со стороны устройств гарантирует пользователям получение достоверного представления о своих показателях.

Когда контроль легитимизируется, устройства для трекинга воспринимаются как заботливые наставники и вежливые надзиратели. Такая форма легитимизации контроля характерна для отношений зависимости: индивид начинает придавать очень большое значение цифрам и болезненно реагирует на случаи, когда не удается выполнить рекомендации устройства, поскольку верит, что трекеры действительно заботятся о его самочувствии.

Надзор со стороны инструментов цифрового селф-трекинга, а также то «давление», которое ощущают на себе пользователи, оправдывающие этот контроль, проявляется в дисциплинарных практиках. В качестве таковых нами были выделены уведомления, сопровождение, полоса непрерывного выполнения задач, оценивание пользователя приложением, система уровней или наград, различные способы визуализации результатов и данных, а также цель, которая может отличаться по формату и способу установления, в зависимости от ситуации использования. В ходе исследования был выделен ряд социальных механик, предполагающих наличие другого актора. Детальное изучение роли «другого» не было целью настоящей работы, однако представляет интерес для дальнейших исследований. Так, остается неизученным, каким образом пользователи выбирают социальные механизмы и свойственны ли этим механизмам специфичные факторы, воздействующие на восприятие практик и ощущение контроля.

Что же касается восприятия дисциплинарных практик, то на него влияют такие факторы, как персонализация, сенсорность, назойливость, наглядность, подробность предоставляемой информации, а также интенция. Последняя отражает наимерение пользователя позволять устройству мотивировать его на совершение

определенных действий, что еще раз подчеркивает значимость вовлеченности лица, использующего приложение.

Таким образом, по мнению авторов, именно интенция — ключевой элемент в проведении границы между контролем со стороны трекера и добровольной помощью в саморазвитии, которую предоставляют эти устройства. На контроль со стороны устройств отслеживания пользователи реагируют крайне болезненно, но также склонны легитимизировать его. Очень часто устройствам отслеживания отводится определенная узкая сфера воздействия, но даже в рамках этой сферы пользователи либо избегают значительных вмешательств в собственную деятельность со стороны устройств, либо легитимизируют их, оправдывая давление заботой со стороны приложений и устройств.

Список литературы (References)

- Голенков С.И. Понятие субъективации Мишеля Фуко // Вестник Самарской гуманитарной академии. Серия «Философия. Филология». 2007. № 1. С. 54—66. URL: <http://www.phil63.ru/ponyatie-subektivatsii-mishelya-fuko> (дата обращения: 25.01.2021).
- Golenkov S.I. (2007) Michel Foucault's Concept of Subjectivity. *Bulletin of Samara Academy for the Humanities. Series «Philosophy. Philology»*. No. 1. P. 54—66. URL: <http://www.phil63.ru/ponyatie-subektivatsii-mishelya-fuko> (accessed: 25.01.2021). (In Russ.)
- Ним Е. Г. Селф-трекинг как практика квантификации телесности: концептуальные контуры // Антропологический форум. 2018. № 38. С. 172—192. <https://doi.org/10.31250/1815-8870-2018-14-38-172-192>.
- Nim E. G. (2018). Self-Tracking as a Practice of Quantifying the Body: Conceptual Outlines. *Forum for Anthropology and Culture*. No. 38. P. 172—192. <https://doi.org/10.31250/1815-8870-2018-14-38-172-192>.
- Савельева Н. В. Между свободой и подчинением: «техники себя» в работах Мишеля Фуко // Интеракция. Интервью. Интерпретация. 2015. Т. 7. № 9. С. 7—15. URL: <https://www.inter-fnisc.ru/index.php/inter/article/view/4357> (дата обращения: 25.01.2021).
- Savelieva N. V. (2015) Between Freedom and Submission: “Techniques of the Self” in the Works of Michel Foucault. *Inter*. Vol. 7. No. 9. P. 7—15. URL: <https://www.inter-fnisc.ru/index.php/inter/article/view/4357> (accessed: 25.01.2021). (In Russ.)
- Фуко М. История сексуальности — III. Забота о себе. Киев: Дух и літера: Грунт; М.: Рефл-бук, 1998.
- Foucault M. Histoire de la sexualité — III. Le souci de soi. Kyiv: Dukh i litera, Grunt; Moscow: Refl-buk, 1998
- Ajana B. (2017) Digital Health and the Biopolitics of the Quantified Self. *Digital Health*. Vol. 3. P. 1—18. <https://doi.org/10.1177/2055207616689509>.
- Berson J. (2015) Computable Bodies: Instrumented Life and the Human Somatic Niche. London: Bloomsbury Academic.

Crawford K., Lingel J., Karppi T. (2015) Our Metrics, Ourselves: A Hundred Years of Self-Tracking From the Weight Scale to the Wrist Wearable Device. *European Journal of Cultural Studies*. Vol. 18. No. 4—5. P. 479—496. <https://doi.org/10.1177/1367549415584857>.

Hassan L., Dias A., Hamari J. (2019) How Motivational Feedback Increases User's Benefits and Continued Use: A Study on Gamification, Quantified-Self and Social Networking. *International Journal of Information Management*. Vol. 46. P. 151—162. <https://doi.org/10.1016/j.ijinfomgt.2018.12.004>.

Kristensen D. B., Ruckenstein M. (2018) Co-Evolving With Self-Tracking Technologies. *New Media & Society*. Vol. 20. No. 10. P. 3624—3640. <https://doi.org/10.1177/1461444818755650>.

Lyall B., Robards B. (2018) Tool, Toy and Tutor: Subjective Experiences of Digital Self-Tracking. *Journal of Sociology*. Vol. 54. No. 1. P. 108—124. <https://doi.org/10.1177/1440783317722854>.

Martin E. (2000) Mind-Body Problems. *American Ethnologist*. Vol. 27. No. 3. P. 569—590. <https://doi.org/10.1525/ae.2000.27.3.569>.

Ramirez E. (2014) An Introduction to the Quantified Self. *Quantified Self*. 26 February. URL: <https://quantifiedself.com/blog/introduction-to-the-quantified-self> (accessed: 27.01.2021).

Rapp A. (2018) Gamification for Self-Tracking: From World of Warcraft to the Design of Personal Informatics Systems. *CHI' 18: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. April 2018. Paper No. 80. P. 1—15. <https://doi.org/10.1145/3173574.3173654>.

Ruckenstein M. (2014) Visualized and Interacted Life: Personal Analytics and Engagements With Data Doubles. *Societies*. Vol. 4. No. 1. P. 68—84. <https://doi.org/10.3390/soc4010068>.

Swan M. (2015) The Quantified Self: Fundamental Disruption in Big Data Science and Biological Discovery. *Big Data*. Vol. 1. No. 2. P. 85—99 <https://doi.org/10.1089/big.2012.0002>.

Whitson J. (2013) Gaming the Quantified Self. *Surveillance and Society*. Vol. 11. No. 1/2. P. 163—176. <https://doi.org/10.24908/ss.v11i1/2.4454>.

Приложение

Список информантов

Номер респондента	Пол	Возраст	Приложение/трекер
1	Женский	32	Фитнес-браслет Fitbit; весы (синхронизация с телефоном); Nike Training Club
2	Женский	19	Nike Training Club, Quizlet, ABBYY Lingvo
3	Мужской	27	Apple watch, Runtastic
4	Мужской	36	Фитнес-браслет Xiaomi Mi Band 4
5	Мужской	29	Tabata timer, Duolingo
6	Мужской	56	Фитнес-браслет Xiaomi Mi Band 4
7	Женский	29	Фитнес-браслет Xiaomi Miband 2; Headspace, Elevate, Pick, Habits
8	Мужской	23	Фитнес-браслет Honor band 5, HabitHub, Focus To-Do
9	Мужской	54	Fitbit Alta Hr Bluetooth
10	Женский	31	Фитнес-браслет Suunto; Strava
11	Женский	30	eMoods, Insight Timer, ранее фитнес-браслет Xiaomi Miband
12	Женский	55	Фитнес-браслет Xiaomi
13	Женский	29	Часы Samsung Galaxy Active 2; VR meditation
14	Женский	45	Фитнес-браслет Xiaomi Mi Band, Polar, пульсометр, Strava
15	Женский	26	Nike Training Club, FatSecret, «Murphy. English grammar in use»
16	Женский	54	Omron Walking Style
17	Мужской	50	Samsung Galaxy Watch Active 2
18	Женский	32	Apple Watch, Pacer, Strava, Insight Timer

DOI: [10.14515/monitoring.2021.1.1810](https://doi.org/10.14515/monitoring.2021.1.1810)



А. И. Егорова, Н. Кловайт

КАК ПОПРОЩАТЬСЯ С РОБОТОМ? ИССЛЕДОВАНИЕ ДОСТИЖЕНИЯ РАЗГОВОРНОГО ЗАКРЫТИЯ

Правильная ссылка на статью:

Егорова А.И., Кловайт Н. Как попрощаться с роботом? Исследование достижения разговорного закрытия // Мониторинг общественного мнения: экономические и социальные перемены. 2021. № 1. С. 241—270. <https://doi.org/10.14515/monitoring.2021.1.1810>.

For citation:

Egorova A.I., Klowait N. (2021) How to Say Good-Bye to a Robot? The Matter of Conversational Closing. *Monitoring of Public Opinion: Economic and Social Changes*. No. 1. P. 241–270. <https://doi.org/10.14515/monitoring.2021.1.1810>. (In Russ.)

КАК ПОПРОЩАТЬСЯ С РОБОТОМ? ИССЛЕДОВАНИЕ ДОСТИЖЕНИЯ РАЗГОВОРНОГО ЗАКРЫТИЯ

ЕГОРОВА Анастасия Игоревна — младший специалист департамента сбора данных, группа «Выполнение опросов в Москве и Московской области», Фонд «Общественное мнение», Москва, Россия

E-MAIL: emeraldansa@gmail.com
<https://orcid.org/0000-0002-8889-6272>

КЛОВАЙТ Нильс — научный сотрудник Центра перспективных социальных исследований, Российской академия народного хозяйства и государственной службы при Президенте РФ, Москва, Россия; старший научный сотрудник Международного центра современной социологической теории, Московская высшая школа социальных и экономических наук, Москва, Россия; эксперт, Лаборатория геймификации Сбера, Москва, Россия

E-MAIL: nils.klowait@gmail.com
<https://orcid.org/0000-0002-7347-099X>

Аннотация. Разговорные компьютеризированные агенты совершенствуются, развиваются и получают все более широкое распространение в различных сферах жизни. Теперь искусственный интеллект может вести беседу с человеком, используя сложные разговорные приемы и человекоподобные речевые паттерны. Например, Google Duplex, современный голосовой интерфейс, способен автономно выполнять сервисные телефонные звонки таким образом, что собеседнику не удается опознать в нем робота.

С увеличением числа интерактивных коммуникаций человека и компью-

HOW TO SAY GOOD-BYE TO A ROBOT? THE MATTER OF CONVERSATIONAL CLOSING

Anastasia I. EGOROVA¹ — Junior Specialist at the Department of Data Collection, Surveys in Moscow and Moscow oblast Division

E-MAIL: emeraldansa@gmail.com
<https://orcid.org/0000-0002-8889-6272>

Nils KLOWAIT^{2,3,4} — Research Fellow at the Center for Innovative Social Research; Senior Research Fellow at the International Center for Contemporary Sociological Theory; Researcher

E-MAIL: nils.klowait@gmail.com
<https://orcid.org/0000-0002-7347-099X>

¹ Public Opinion Foundation, Moscow, Russia

² Russian Presidential Academy of National Economy and Public Administration, Moscow, Russia

³ Moscow School of Social and Economic Sciences, Moscow, Russia

⁴ Sber Gamification Lab, Moscow, Russia

Abstract. Artificial conversational agents are becoming increasingly popular in various spheres of life. Contemporary AI is able to talk to humans using sophisticated conversational techniques and human-like conversational patterns. For instance, Google Duplex, a cutting-edge voice interface, is capable of autonomously making customer service calls that barely register as robotic. With the frequency of human-computer interactions on the rise, there is a growing need to study their features: how misunderstandings are resolved, how conversational aims are achieved. This article is an empirical investigation of the interaction between an advanced conver-

тера становится очевидной необходимость их изучения: каким образом они происходят, как разрешаются недоразумения и достигаются цели разговора. Статья посвящена эмпирическому изучению взаимодействия между продвинутым разговорным агентом и человеком: в рамках этнографии и конверсационного анализа рассмотрены последовательно разворачивающиеся, естественно возникающие речевые практики собеседников, а также порождаемые ими нормативные структуры.

Материалом для анализа послужила аудиозапись телефонной беседы между Google Duplex и человеком-оператором. Исследовано, в частности, как достигается разговорное закрытие — микроструктура, требующая кооперации собеседников в организации чередования реплик. Показано, что завершение беседы происходит за счет постепенного достижения эпистемического равновесия между говорящими, несмотря на интеракционные проблемы, возникающие из-за неуместного просодического контура робота. Подчеркивается связь между институциональным контекстом взаимодействия и роботизированной природой не-человеческого собеседника.

Полученные результаты закладывают основу для дальнейших исследований специфического порядка во взаимодействии человека и робота и способствуют более глубокому пониманию того, как завершается разговор в пограничных случаях.

Ключевые слова: этнография, конверсационный анализ, разговор-

sational agent and human interactant. Using the framework of ethnography and conversation analysis, the authors analyze the procedurally unfolding and naturally emerging conversational practices, as well as the normative structures that emerge as a result. The research is based on a recording of a call between Google Duplex and a human operator. We explore how to reach conversational closing — a microstructure that requires cooperation between the speakers. Despite interactional tensions caused by the robot's incongruous prosody, conversational closing is produced by the gradual achievement of epistemic balance. The authors emphasize the relationship between the institutional context of the interaction and the non-human agent's robotic nature. The results serve as a foundation for future studies in human-robot interaction and provide a deeper understanding of how conversational closings are achieved in liminal cases.

Keywords: ethnography, conversation analysis, conversation closing,

ное закрытие, взаимодействие человека и компьютера, просодический контур, эпистемика

Благодарность. Статья подготовлена в сотрудничестве с Лабораторией геймификации Сбера.

Статья подготовлена в рамках гранта, предоставленного Министерством науки и высшего образования Российской Федерации (№ соглашения о предоставлении гранта: 075-15-2020-908).

human-computer interaction, prosodic contour, epistemics

Acknowledgments. The article is made in collaboration with the Sber (ex.—Sberbank's) Gamification Laboratory, under a grant from the Ministry of Science and Higher Education of the Russian Federation (contract No. 075-15-2020-908).

Как завершается разговор между собеседниками? Как участники понимают, что один из них сообщил другому достаточно информации для того, чтобы оба повесили трубки? Эти вопросы — предмет интереса этнографии и конверсационного анализа (EMCA), где разговор рассматривается как «средство, с помощью которого организуется социальный порядок» [Drew, Heritage, 1992: 12]. Данная теоретическая рамка изучает, как участники разговора, используя различные микроструктуры (например, организация чередования говорящих, последовательностей в разговоре, просодия), локально определяют ситуацию.

А что будет меняться в завершении разговора, если одним из собеседников будет не человек, а робот? Искусственный интеллект активно развивается, роботы становятся интерактивными и человеку даже не всегда удается распознать, что он общается с умной машиной. Однако действия не-человека все еще ограничены по сравнению с действиями человека [Button, 1990: 81], поэтому в данной работе будет рассмотрено, каким образом завершается разговор между человеком и продвинутым разговорным интерфейсом.

Перспектива конверсационного анализа

Согласно допущению конверсационного анализа, каждый отдельный разговор последовательно разворачивается во времени. Организация чередов¹ ("turns") — определяющий элемент последовательного структурирования обычного разговора. Сама последовательность конструируется с помощью различных единиц ("turn-constructional units") [Sacks et al., 1974: 720—721] (например, фраза "Hi, how are you?") через синтаксическое выделение и «производство звуков» [Sacks et.al., 1974: 180], или просодический контур [Selting, 1996], который включает в себя «замиинки, растягивание звука, осечки, вдохи, смех, микропаузы, паузы, ударения и интонации», а также тон, громкость [Couper-Kuhlen, Selting, 2009: 13].

¹ Здесь и далее использован термин «череда»/«череды», появившийся в переводе статьи «Простейшая систематика организации очередности в разговоре» [Сакс, Щеглофф, Джейферсон, 2015], сделанном кандидатом социологических наук Андреем Корбутом.

Смена чередов в разговоре тщательно координируется: следующий говорящий, как правило, вступает в момент возможного завершения типовой единицы, который называется «релевантным местом перехода» (“transition-relevance place”).

Поскольку в этой статье мы фокусируемся исключительно на закрывающей части беседы, нас будет интересовать, как чередование прерывается и разговор заканчивается. Вместо того, чтобы просто перестать говорить (что может восприниматься как молчание), для завершения взаимодействия участники закрывают саму возможность передачи череда, что обеспечивается терминальной парой смежности (“adjacency pairs”) [Sacks et al., 1974: 728]. Пары смежности представляют собой два высказывания, занимающие соседнее расположение и производящиеся различными говорящими. Пример подобных парных типов — обмен прощаниями «пока — пока» [Schegloff, Sacks, 1973: 296]. Это терминальная пара, то есть фактический обмен в заключительной части разговора.

Перед терминальным обменом последовательно инициируется заключительная часть разговора. Когда один из собеседников не претендует на инициацию новой темы, после закрытия предыдущей он вводит первую парную часть в виде возможного предзакрытия (“pre-closing”), например, “We-ell...”, “О.К...”, “So-oo” [Schegloff, Sacks, 1973: 303]. Затем черед переходит ко второму говорящему, который либо начинает новую тему и тем самым продолжает разговор, либо отвечает на предзакрытие (Пример 1) — тогда возможное предзакрытие становится фактическим, и разговор завершается [Schegloff, Sacks, 1973: 309].

Пример 1 1 А: О.К.
 2 В: О.К.

Однако реакцией на предзакрытие может быть не только окончание разговора, но и минимальное или радикальное уклонение от закрытия [Button, Lee, 1987: 109—111]. В обоих случаях разговор переоткрывается, но в первом случае последовательность расширяется ненадолго, а во втором закрытие резко отодвигается во времени.

Таким образом, завершение взаимодействия представляет собой сложный процесс, кооперацию в организации чередования собеседников, а не простое следование желанию прекратить разговор.

Проблематичность завершения разговорного взаимодействия между людьми была изучена в телефонных звонках [Schegloff, Sacks, 1973: 696—735; Button, Lee, 1987: 101—151], устном разговоре лицом к лицу [Broth, Mondada, 2013; Haddington, 2019; Mondada, 2014], компьютерно-опосредованном взаимодействии [Markman, 2006: 138—172], виртуальных играх [Rojaparipua, Jaroenkitboworn, 2011]. Также предпринимались попытки изучать закрытия в симулированном разговоре между человеком и компьютерной системой [Wooffitt, 1997]: при этом человек-участник был убежден, что разговаривает с голосовым неинтерактивным роботом, но роль компьютера играл другой человек. Было обнаружено, что в отличие от разговора между людьми, в котором терминальный обмен значим для собеседников, в симулированном разговоре терминальный компонент, как правило, систематически отсутствует. Маркером того, что говоря-

щему больше нечего сказать, выступает в таком случае прекращение разговора (сброс звонка) без терминальной последовательности.

Однако настоящие разговоры робота с человеком представляют новый объект для изучения закрытий². В таком случае форма социальной ситуации меняется, и роботичность начинает влиять на завершение разговора, что будет показано в текущей работе. Таким образом, эта статья вносит вклад в изучение выстраивания специфического порядка взаимодействия в контексте разговора между человеком и не-человеком.

В данной работе изучается, как происходит закрытие разговора во взаимодействии человека и искусственного интеллекта Google Duplex, сложной компьютерной системы, в основе которой лежит рекуррентная нейронная сеть, способная использовать свою внутреннюю память для распознавания речи. Ответ на исследовательский вопрос основан на анализе конкретных эмпирических данных, где Google Duplex эксплицитно демонстрирует способность участвовать в разговоре.

Данные

Рассматривается аудиозапись естественно протекающего телефонного разговора между оператором-человеком и голосовой системой Google Duplex, продолжительность которого составляет 56 секунд. Запись предоставлена компанией Google и находится в открытом доступе³. Система выступает в качестве клиента, способного делать телефонные звонки исключительно для оформления заказов (например, бронирование столика в ресторане), а человек является оператором, который выполняет эти поручения.

Изучать аудиозапись мы будем на основании методологии конверсационного анализа с помощью детального анализа транскриптов естественно протекающих разговоров [Sacks et.al., 1974: 696—734]. Транскрипт исследуемого разговора, выполненный по системе Гейл Джейферсон [Jefferson, 2004: 14—31] (см. Приложение 1), представлен в Приложении 2.

Для человека беседа протекает в обычном режиме, но в реальности разговорное взаимодействие происходит не между людьми, а между человеком и роботической системой. Благодаря тому, что в рассматриваемом разговоре один из участников не-человек, можно характеризовать некоторые наблюдаемые особенности его речи как роботические. Иными словами, роботичность разговора превносит определенные отличия во взаимодействие. Если изучать разговор с точки зрения конверсационного анализа [Schegloff, 2009: 370], длинные паузы (более 1,2 сек.), видимое отсутствие реакции и малая вовлеченность (робот только отвечает, не задает вопросы, ничего не уточняет), повышенная громкость, ровный тон, растянутость речи можно отнести к характерным чертам этой компьютерной системы.

Исследуемое взаимодействие отличается от «обычных разговоров» [Drew, Heritage, 1992: 3—6]. Изучаемый нами кейс представляет собой институциональ-

² Каким образом конверсационный анализ относится к проблеме человеко-компьютерного взаимодействия в контексте развития искусственного интеллекта см. в статье Н. Кловайта и М. Ерофеевой в этом выпуске [Klowait, Erofeeva, 2021].

³ Leviathan Y., Matias Y. Google AI Blog: Google Duplex: An AI System for Accomplishing Real-World Tasks Over the Phone. 2018. 8 May. URL: <https://ai.googleblog.com/2018/05/duplex-ai-system-for-natural-conversation.html> (accessed: 18.02.2021).

ный разговор [Drew, Heritage, 1992: 21—25], накладывающий определенные ограничения на действия его участников и их интерактивные компетенции [Klowait, 2017: 21—24; Klowait, 2018: 25—26]. Разговор Google Duplex и человека осуществляется в формате сервисного звонка, в котором роли фиксированы [Moore, Arar, 2019: 40]: клиент Duplex запрашивает услугу, а человек предоставляет ее от имени организации. Отметим, что форма смены череда в таком разговоре, а также диапазон возможностей для действий отличаются от повседневного разговора [Drew, Heritage, 1992: 25—27]. Разговор между человеком и системой, как и многие другие институциональные разговорные взаимодействия [Drew, Heritage, 1992: 54], характеризуется последовательностью вопросов-ответов, в которой клерк задает вопросы, а клиент на них отвечает, хотя это ограничение не исключает свободу действий говорящих.

Проблематичное завершение

Разговорное закрытие в текущем кейсе затягивается после того, как оператор (R) получил всю релевантную информацию от робота Google Duplex (D) (35—48):

- 34 R: Okay↑ (0.2) I got it↓
35 (2.9)
36 R: S:::o. It will be two people at seven o'clock'↑ ⁴
37 (1.6)
38 D: YES mmhm=
39→ R: =Mmhmm (.) okay↑ (.) got it↓↓yeh (0.2)
40→ Thank you↑
41 (1.9)
42→ D: OK:AY A::WESOME (.) THANKS A LOT↓
43 (0.5)
44→ R: >Thank you have good night.<
45→ (0.2) Bye
46→ (1.2)
47→ D: OKAY (.) BYE-BYE
48 ... End call ...

Несмотря на то, что говорящие ориентируются на закрытие разговора, оно продолжается заметно долго. Сначала инициируются возможные предзакрытия (okay got it yeh) и благодарности (Thank you) (строки 39—40), которые принимаются вторым говорящим (42), а затем происходит терминальный обмен (44—47).

Подобное продолжительное закрытие нехарактерно для институционального разговора, в котором после выполнения запроса разговор последовательно, но быстро завершается. Многие институциональные разговоры имеют некую стандартную форму, или этапный порядок [Drew, Heritage, 1992: 43]: открытие — запрос — серия вопросов — ответ — закрытие. Хотя уклонение от закрытий для институциональных разговоров также естественно (например, когда у звонящего появляется новый запрос), но в данном случае Duplex не делает новых запросов, а закрытие все равно является длительным.

⁴ По методологическим причинам в конверсационно-аналитическом транскрипте обычные правила орфографии или пунктуации намеренным образом исключаются.

В работе мы рассмотрим, каким образом специфика роботического разговора между оператором и Google Duplex влияет на его завершение. Для этого в статье будут рассмотрены два эмпирических вопроса. 1) Каким образом роботические черты связаны с институциональными ролями говорящих? То есть как связаны длинные паузы и видимое отсутствие робота с тем, что он делает запрос, а оператор его выполняет? 2) С помощью каких средств происходит закрытие данного разговора?

Эпистемика как ресурс организации последовательностей в разговоре

Для ответа на вопрос, каким образом осуществляется закрытие разговора в исследуемом кейсе, рассмотрим разговорную последовательность через концепцию эпистемики, которая исследует релевантность информации во взаимодействии. Ресурс эпистемики поможет выявить, каким образом закрываются и открываются последовательности в разговоре, а также в какой момент разговора говорящие ориентируются на его закрытие.

Введение в эпистемику

Согласно данной концепции, говорящие имеют доступ к областям знания, располагающимся на эпистемическом градиенте: говорящий с абсолютным эпистемическим преимуществом, то есть осведомленный о каком-либо предмете, обладает «знанием», К+ (от Knowledge), а получатель, который совершенно не знает предмета, о котором идет речь, не обладает «знанием», К– (менее осведомленный) [Heritage, 2012а: 4—6]. Отметим, что для определения того, обладает говорящий знанием или нет, эпистемические градиенты прослеживаются через соотношение чередов друг с другом, а не по отдельности [Heritage, 2013: 553].

Ключевые компоненты эпистемики: эпистемический статус («epistemic status»), эпистемическая позиция («epistemic stance»), а также взаимодействие между ними. Эпистемический статус включает в себя «не только реальное обладание информацией, но и права на ее обладание и формулирование» [Heritage, 2013: 558]. С одной стороны, говорящие могут занимать разные позиции на эпистемическом градиенте, и даже если они занимают одну позицию, степень «глубины» их знаний может различаться. С другой стороны, участники разговора могут обладать равным объемом информации [Heritage, 2012а: 4—6]. Эпистемический статус определяется во взаимодействии самими участниками, то есть может быть признан или нет.

Эпистемическая позиция «касается последовательного выражения социальных отношений, управляемого посредством конструирования чередов в разговоре» [Heritage, 2013: 558]. Это то, как говорящие позиционируют себя с точки зрения эпистемического статуса в процессе разговора.

Эпистемика необходима для того, чтобы определить одно из фундаментальных различий в конверс-анализе — запрос информации и информирование, и, соответственно, понять, является ли действие запросом информации или ее предоставлением, а также как формируются данные действия [Heritage, 2013: 557].

Запрос информации происходит с помощью трех компонентов: вопросительный морфосинтаксис (грамматическая форма вопроса), просодия (интонационный контур), эпистемика.

Согласно традиционному представлению, утвердительные высказывания, декларативы, (3) в обычном употреблении утверждают информацию (“assert information”), а вопросительные высказывания, интерrogативы, (1,2)— запрашивают информацию (“request information”) [Heritage, 2012a: 6].

- (1) Are you married?
- (2) You’re married, aren’t you?
- (3) You’re married.

Однако Дж. Херитидж показывает, что именно эпистемика, а не морфосинтаксис или просодический контур играет решающую роль в определении действия как запроса информации или информирования [Heritage, 2012a: 7—23]. Например, в (2) [KR:2] является ли вопрос матери в строке 11 («Знаешь ли ты, кто идет на эту встречу?») запросом информации? Для того чтобы ответить на этот вопрос, необходимо посмотреть на последующие череды. Расс в строке 12 спрашивает: «Кто.», тем самым осуществляя запрос информации, после чего мать отвечает, что не знает (строка 13), что указывает на ее статус К-. Тогда Расс дает ответ (строки 15—16) на запрашиваемую информацию. Следовательно, в данном случае вопрос в строке 11 действительно был запросом информации, хотя он и был сначала проинтерпретирован Рассом как информирование, на что указывает его вопрос «Кто.».

- (2) [KR:2][Heritage, 2013: 553]
- 11 Mom: -> Do you know who’s going to that meeting?
12 Rus: Who.
13 Mom: I don’t kno:w.
14 (0.2)
15 Rus: .hh Oh::: Prob’ly .h Missiz Mc Owen (‘n Dad said)
16 prob’ly Missiz Cadry and some of the teachers.

Информационная асимметрия: Duplex (K+) и оператор (K-)

В нашем кейсе Duplex (D) звонит и декларативно формулирует высказывание, инициируя начало разговора: “Hey, I’m calling to make reservation(.) I’m looking for a table on Friday the fourth”. Это высказывание явно попадает в его эпистемическую область (K+), так как последующий черед от оператора (R) представляет собой интерrogатив “This Friday.”, после чего D подтверждает информацию (строка 09):

- 04 **D:** Hey
05 I:’m calling to make reservation? (.) I’m looking for a table
on Friday the fourth↓
06 (1.2)
07 **R:** This Friday,
08 (1.0)
09 **D:** Yeah

При этом вопросительная формулировка R (строка 07) говорит о том, что необходимая информация находится вне его эпистемической области, то есть R запрашивает информацию, которая в полном объеме ему не доступна. Этот вывод можно сделать на основании того, что, во-первых, после предположения R в седьмой строке D подтверждает его. Во-вторых, из контекста данного разговора

мы знаем, что D делает запрос (бронирует столик), сообщая информацию о себе, доступ к которой есть только у него.

Затем последовательность продолжается в виде цепочки вопросов и ответов, где каждая инициируется через вопрос K– и завершается предоставлением от K+ ответа (1), подтверждающим повторением слова от K– (2), либо подтверждающим токеном⁵ (3).

В (1), (2), (3) R в вопросительной форме, которая соответствует его эпистемическому статусу (K–), запрашивает релевантные детали (“So how many people,”, “For two,”, “In what time you will come,”, “Can you say name,”). В свою очередь, D сообщает необходимую информацию, используя декларативный синтаксис, соответствующий его области знания (“It’s for two people”, “At seven pm”, “Ehm (0.3) first name is Yaniv (0.5) YA (0.2) NI (.) V”). Таким образом, эпистемический статус участников остается неизменным на протяжении рассмотренных последовательностей разговора.

(1)

- 14 (0.9)
15→ **R:** S:o how many people,
16 (1.1)
17 **R:** Fi[ve,
18→ **D:** [It's fo::r two people↓
19 (0.4)
20→ **R:** For two,
21 (1.2)
22→ **D:** Yes

(2)

- 23 (0.6)
24→ **R:** E:n what time you will come,
25 (1.2)
26→ **D:** At seven pm
27 (0.9)
28→ **R:** °Seven°
29 (0.9)

(3)

- 30→ Can you say name,
31 (1.3)
32→ **D:** E:::HM (0.3) FIRST NAME IS YANIV (0.5) Y A (0.2) N I (.) V
33 (2.7)
34→ **R:** Okay↑ (0.2) I got it↓
35→ (2.9)
36→ **R:** S:::o. It will be two people at seven o'clock' ↑

⁵ Токен подтверждения — маркер, подтверждающий принятие/понимание предыдущего череда говорящего, после чего открывается следующая последовательность. См. [Jefferson, 1984: 199].

Теперь рассмотрим подведение итога разговора (строки 35—36) с точки зрения эпистемики. R декларативно формулирует “So it will be two people at seven o’clock”, получив к этому моменту уже всю информацию от D, однако он произносит это с восходящим просодическим контуром, что все еще является запросом информации. Восходящий просодический контур характерен для запроса информации, даже если морфосинтаксис декларативный [Heritage, 2013: 564].

Таким образом, до строки 34 спикеры занимают разные позиции на эпистемическом градиенте, то есть один более осведомленный, а другой — менее. Статусом K– обладает человек, и в соответствие с этим он запрашивает информацию, а статусом K+ обладает Duplex, в соответствие с чем он делает запрос (бронирование столика) и предоставляет информацию.

Эпистемическое равновесие как инициирование заключительной части разговора

Когда сокращается эпистемический градиент, то есть информационный дисбаланс признается «уравновешенным для всех практических целей» [Heritage, 2012c: 32], происходит изменение состояния (“change of state”) говорящего, который запрашивал информацию: его эпистемический статус меняется с K– на K+. Тогда оба говорящих приобретают эпистемический статус K+ и достигают эпистемического равновесия, которое «подразумевается при закрытии последовательностей, тем и даже целых бесед» [Heritage, 2012b: 390]. В институциональных разговорах при изменении состояния последовательности завершаются с использованием токенов подтверждения “okay”, “alright”, которые регистрируют и принимают ответное действие [Drew, Heritage, 1992: 126], а также признательности (“thank you”), смеха (“haha”), положительной (“great”, “excellent”) и негативной (“too bad”, “oh well”) оценок [Moore, Arar, 2019: 142]. Таким образом, при устранении информационного дисбаланса последовательность может быть закрыта, однако это не означает, что новые темы не возникнут.

Токены подтверждения как возможные ресурсы предзакрытия разговора

Токены подтверждения закрывают текущую последовательность и открывают новую. В заключительной же части разговора они представляют собой возможные предзакрытия разговора в целом. Рассмотрим, как работают данные токены в разговоре клерка и Google Duplex и какой из них способствует завершению данной беседы.

В разговоре Duplex и человека R предоставляет токен подтверждения (“Mmhm”) в строке 11, который закрывает последовательность, относящуюся к выбору дня недели (“This Friday”), однако еще не может закрыть весь разговор, так как не все детали известны:

- 07→ R: This Friday,
- 08 (1.0)
- 09→ D: Yeah
- 10 (0.5)
- 11→ R: °Mmhm°
- 12 (0.9) °em Let me see°=

Следующий токен, повторяющий ответ D (“seven”) в строке 28, также закрывает только одну последовательность, относящуюся к вопросу о времени, и после паузы R инициирует новый черед (“Can you say name”):

- 24 **R:** E:n what time you will come,
25 (1.2)
26 **D:** At seven pm
27 (0.9)
28→ **R:** °Seven°
29→ (0.9)
30→ Can you say name,

Не очевидно, к чему относятся два подтверждающих токена в строке 34 (первый — “Okay”, второй — “I got it”): только ли к предыдущему череду или ко всему предшествующему разговору. Можно заметить, что до и после этого череда возникают длинная заминка [Sacks et al., 1974: 715]⁶ (строка 33) и пауза [Sacks et al., 1974: 715]⁷ (строка 35) со стороны клерка. Заминку можно объяснить тем, что клерк записывал информацию, которая была ему сообщена в предыдущем череде (32 строка). Пауза же завершается растянутым “so” в начальной позиции череда (turn-initial position), произнесенным слегка понижающимся тоном. Ее можно объяснить тем, что клерк систематизирует (строка 36) ранее полученную от клиента информацию (строки 15—29) — токен “so” указывает на то, что дальше будет высказан итог предшествующего разговора [Raymond, 2004: 186—191]. Во-первых, клерк проверяет, правильно ли он услышал и записал информацию клиента, и во-вторых, в случае правильности — токен “so” после подтверждения закрывает последовательность основной части разговора, после чего говорящий может приступить к его закрытию [Schegloff, Sacks, 1973: 303]:

- 32→ **D:** E:::HM (0.3) FIRST NAME IS YANIV (0.5) YA (0.2) NI (.) V
33→ (2.7)
34→ **R:** Okay↑ (0.2) I got it↓
35→ (2.9)
36→ **R:** S:::o. It will be two people at seven o'cloc'↑
37 (1.6)

Следовательно, токены “Okay I got it” подтверждают лишь предыдущий черед, так как после подтверждения следует пауза (строка 35) и подводится итог (строка 36) предшествующих последовательностей.

Однако в строке 39 эти же токены (“okay got it”) выступают в качестве предзакрытия, так как информационный дисбаланс был уравновешен, то есть клерк получил достаточную информацию для выполнения запроса (дата, количество человек, время, имя) и изменил свой эпистемический статус с K- на K+:

- 36 **R:** S:::o. It will be two people at seven o'cloc'↑
37 (1.6)
38 **D:** YES Mmhmm=
39→ **R:** =Mmhmm (.) okay↑ (.) got it ↓↓ yeh (0.2)

⁶ Заминки (“lapses”) — продолжительное молчание между чередами говорящих.

⁷ Пауза — молчание внутри чередов, то есть когда нет смены говорящих.

Прослеживается связь между токенами подтверждения, закрывающими последовательность, и эпистемикой: именно в момент достижения эпистемического баланса возможное предзакрытие становится фактическим.

Мы продемонстрировали, что ресурс эпистемики активно задействован говорящими и становится решающим для понимания, в какой момент разговор должен закончиться. Клерк, запрашивая информацию, инициирует последовательности, в которых он обладает статусом К–, до тех пор, пока говорящие не достигнут эпистемического равновесия. В то же время робот, предоставляя информацию, закрывает некоторые последовательности, после чего человеком инициируются новые. Кроме того, в нашем кейсе последовательности закрываются с помощью токенов подтверждения, занимающих третью позицию (вопрос — ответ — подтверждение), которые исходят от человека и после которых также инициируются новые последовательности. При этом единственны подобные токены (строка 39) являются предзакрытием всего разговора в момент достижения эпистемического баланса, когда все детали запроса стали известны клерку.

Теперь для того чтобы посмотреть, каким образом говорящие подходят к завершению, обратимся к другому теоретическому ресурсу — организации последовательностей в разговоре. Этот ресурс позволит обнаружить особенности перехода к заключительной части по отношению к другим частям разговора.

Организация последовательностей в разговоре

В рамках изучения организации последовательностей (“sequence organisation”) рассматривалось, каким образом каждый черед в разговоре либо принимает новое направление, либо сохраняет тот же самый курс действий (приветствия, запрос, серия уточняющих вопросов, закрытие) [Heritage, 1989: 32—34; Couper-Kuhlen, 2004: 335—360; Zellers, 2012: 233—236; Hoey, 2015: 443—448]. Помимо последовательной организации, в разговоре существует тематическая организация (“topic organisation”). Смена или продолжение темы и последовательностей в разговоре взаимосвязаны и могут пересекаться: по концу темы можно определить конец последовательности, но тема может продолжаться, хотя последовательность уже завершена.

Как новые начинания, так и продолжения происходят в местах возможного закрытия последовательности (“at points of possible sequence closure” [Couper-Kuhlen, 2004: 346]). Рассмотрим особенности последовательной организации для выявления средств закрытия разговора в нашем кейсе.

Инициирование новых последовательностей в разговоре

Курс действий может меняться в трех разных частях разговора. Во-первых, когда открывющая часть (“opening section of conversation”) подходит к месту возможного закрытия, говорящий может ввести тему своего звонка [Couper-Kuhlen, 2004: 337].

Во-вторых, новый курс действий может быть инициирован в основной части разговора в местах возможного закрытия темы (“at points of possible topic closure in the body of the conversation”), то есть как внутри череды одного говорящего, так и между чередами разных говорящих. Последовательность может закрываться

в третьей позиции (“sequence-closing third”), что предшествует смене темы и инициации новой последовательности [Couper-Kuhlen, 2004: 344—345].

В-третьих, новое начинание может вводиться не как новая тема, а как полный переход от темы разговора к заключительной части (“closing section of conversation”) в месте возможного открытия закрывающей секции [Couper-Kuhlen, 2004: 347].

Таким образом, последовательность принимает новое направление либо путем ввода новой темы в качестве причины звонка, либо через инициирование нового курса действий внутри разговора, либо когда происходит открытие заключительной части разговора.

Структура разговора робота и человека: инициирование новых подтем

Из-за того, что в нашем исследовании анализируется институциональный разговор, его общая и единственная тема — исходящий от робота запрос бронирования столика в ресторане. При этом в соответствии с организацией последовательностей можно выделить восемь блоков, подтем, которые на протяжении всего разговора открываются и закрываются.

Первый блок (строки 00—04) представляет собой открытие звонка, в котором говорящие приветствуют друг друга. После этого начинается второй блок, когда звонящий D вводит тему своего запроса, причину звонка, и происходит обсуждение даты (строки 05—11). Таким образом, открывается основная часть разговора. В строке 11 последовательность, относящаяся к обсуждению даты бронирования, подходит к месту возможного закрытия с помощью подтверждающего токена D “Mmhm”:

00 ((Ring))
01 (0.3)
02 R: ()
03 (0.9)
04 D: Hey
05 I:’m calling to make reservation? (.) I’m looking for a table
on Friday the fourth↓
06 (1.2)
07 R: This Friday,
08 (1.0)
09 D: Yeah
10 (0.5)
11 R: °Mmhm.

Затем курс действий меняется, начиная с фразы “Let me see” производится обработка полученной информации, а “Mmhm” D сигнализирует R, что первый ожидает продолжения разговора с его стороны (строки 12-14):

12 (0.9) °em Let me see.=
13 D: =Mmhm
14 (0.9)

После этого идет четвертый блок, в котором R инициирует новую последовательность и тему — на какое количество человек необходимо забронировать сто-

лик (строки 15—23). Данная подтема подходит к завершению в строках 22—23, когда D подтверждает количество человек, которое обсуждалось в предшествующих чередах:

- 15 **R:** S:o how many people,
16 (1.1)
17 **R:** Fi[ve,
18 **D:** [It's fo::r two people↓
19 (0.4)
20 **R:** For two,
21 (1.2)
22 **D:** Yes
23 (0.6)

В пятом блоке обсуждается новая подтема (строки 24—28), которая инициируется R в месте возможного завершения последовательности внутри основной части разговора (строка 23). Новая подтема касается определения времени; эта последовательность завершается в строке 28 с помощью повтора юнита ("Seven"). Шестой блок (строки 29—34) инициируется R после небольшой паузы (строка 29) и касается выяснения имени звонящего:

- 24 **R:** E:n what time you will come,
25 (1.2)
26 **D:** At seven pm
27 (0.9)
28 **R:** °Seven°
29 (0.9)
30 Can you say name,
31 (1.3)
32 **D:** E:::HM (0.3) FIRST NAME IS YANIV (0.5) Y A (0.2) N I (.) V
33 (2.7)
34 **R:** Okay↑ (0.2) I got it↓

После закрытия предыдущей последовательности (строка 34) происходит переход к седьмому блоку, когда после значительной паузы (строка 35) вводится резюме (строка 36) того, что обсуждалось ранее:

- 35 (2.9)
36 **R:** S::::o. It will be two people at seven o'clock↑
37 (1.6)
38 **D:** YES Mmhmm=

Наконец, переход в заключительную часть разговора осуществляется в восьмом блоке, в котором говорящие отходят от общей темы разговора с помощью предзакрытия (строка 39). Этот блок включает в себя обмен предзакрытиями—подтверждениями и благодарностями (39—42)—и заключительную последовательность (44-47), после чего звонок заканчивается:

- 39 **R:** =Mmhmm (.) okay↑ (.) got it ↓↓ yeh (0.2)
40 Thank you↑
41 (1.9)
42 **D:** OK:AY A::WESOME (.) THANKS A LOT↓

43 (0.5)
44 R: >Thank you have good night.<
45 (0.2) Bye
46 (1.2)
47 D: OKAY (.) BYE-BYE
48 ... End call ...

Таким образом, в этом разговоре последовательности поочередно сменяют друг друга, пока тема продолжается. Она заканчивается в момент, когда открывается завершающая последовательность разговора, говорящие обмениваются предзакрытиями, а также вводится заключительная последовательность.

Способы изменения тем и последовательностей

Согласно Э. Купер-Куллен, тема в разговоре может продолжаться, либо принимать новое направление двумя способами: дизъюнктивно или пошагово [Couper-Kuhlen, 2004: 335—352]. В то время как последовательности либо меняются дизъюнктивно, либо продолжаются. Поскольку изменение последовательности часто пересекается со сменой темы [ibid.: 347], то можно рассматривать изменение темы как потенциальный маркер смены последовательностей.

Дизъюнктивный вид смены темы представляет собой доведение темы до конца и инициирование другой. В то время как «пошаговый» способ предполагает постепенные изменения, которые достигаются за счет того, что говорящий увязывает упомянутое раньше с новой темой, поэтому такая смена остается практически незаметной [Zellers, 2012: 235].

Способ изменения/продолжения последовательности или темы маркируется с помощью лексических, грамматических и просодических сигналов в начале нового череда [Couper-Kuhlen, 2004: 371]. Эти ресурсы работают в совокупности, однако в случаях, когда в разговоре не используются первые два ресурса, его участники могут полагаться только на просодию [ibid.: 366—372].

Купер-Куллен определяет следующие характерные черты дизъюнкции и продолжения. Лексические маркеры дизъюнкции — “hey”, “oh” и подобные слова, а также обращения [ibid.: 360—371]. В роли грамматических маркеров продолжения последовательности или темы выступают «анафорические формы, указывающие на преемственность» [ibid.: 372] (например, «это», «оно»). Дизъюнктивными маркерами просодического контура могут быть паузы и вдохи, заметное повышение громкости речи и повышение тона в начале единицы конструирования череда, то есть на первом слоге [ibid.: 337—352].

Предзакрытия обычно вводятся дизъюнктивно в местах возможного завершения последовательности и характеризуются паузами и/или вдохами и повышением тона. При этом не обязательно вводится новая тема, потому что, как мы отметили ранее, курс действий может меняться, например, через переход к заключительной части разговора [ibid.: 350].

Таким образом, для того чтобы определить направление курса действий (изменение или продолжение последовательности) и тип смены темы (дизъюнктивный или пошаговый), необходимо анализировать как синтаксические, так и просодические ресурсы.

Пошаговый и дизъюнктивный способы смены темы в разговоре робота и человека

В нашем кейсе первая смена подтемы происходит из начальной части вызова постепенно и вводится с помощью растянутого “I’m”, произнесенного ровным тоном (строка 05):

- 05→ **D:** I’m calling to make reservation? (.) I’m looking for a table on Friday the fourth↓
06 (1.2)
07 **R:** This Friday,
08 (1.0)
09 **D:** Yeah
10 (0.5)
11 **R:** °Mmhmm.

Затем после паузы (строка 12) подтема пошагово сменяется с помощью тихо введенного “em” с ровной интонацией, когда R обрабатывает информацию:

- 12→ **R:** (0.9) °em Let me see°=
- 13 **D:** =Mmhmm
- 14 (0.9)

Третья смена происходит также пошагово с помощью растянутого “s:o” с ровной интонацией (строка 15):

- 15→ **R:** S:o how many people,
16 (1.1)
17 **R:** Fi[ve,
18 **D:** [It's fo::r two people↓
19 (0.4)
20 **R:** For two,
21 (1.2)
22 **D:** Yes
23→ (0.6)

Обсуждение времени начинается после небольшой задержки (строка 23) и вводится с помощью вопроса, который инициируется с растянутой “E::n” (строка 24) без повышения тона на первом слоге:

- 24→ **R:** E:n what time you will come,
25 (1.2)
26 **D:** At seven pm
27 (0.9)
28→ **R:** °Seven°.

Пятая смена подтемы осуществляется после того, как предыдущая последовательность завершается в закрывающей третьей позиции, в которой R повторяет слово “seven” в качестве подтверждения (строка 28). Новая последовательность вводится пошагово после небольшой паузы (строка 29) и инициируется с помощью вопроса, который начинается с “can” и произносится ровным тоном:

- 29→ (0.9)
30 Can you say name,
31 (1.3)

32 **D:** E:::HM (0.3) FIRST NAME IS YANIV (0.5) Y A (0.2) N I (.) V
33 (2.7)

34→ **R:** Okay↑ (0.2) I got it↓

Затем резюме вводится постепенно после значительной паузы (строка 35) с помощью растянутого токена “So” со слегка понижающимся тоном (строка 36):

35→ (2.9)

36→ **R:** S:::o. It will be two people at seven o'clock↑

37 (1.6)

38 **D:** YES Mmhmm=

Данное изменение последовательности отличается от всех предыдущих — пауза длиннее, тон приглушается (35—36). Это может говорить о том, что последовательность будет меняться, то есть вместо уточняющих вопросов будет введен новый курс действий (систематизация информации). Таким образом, инициирование череда R в 36 строке усиливает закрытие разговора, так как после подтверждения информации, возможно, начнется завершение разговора.

Наконец, седьмая смена, переход от темы разговора к заключительной части, осуществляется дизъюнктивно с помощью предзакрытия, которое произносится с повышением тона в месте возможного завершения последовательности (строка 39):

39→ **R:** =Mmhmm (.) okay↑ (.) got it↓↓yeh (0.2)

40 Thank you↑

41 (1.9)

42 **D:** OK:AY A:::WESOME (.) THANKS A LOT↓

43 (0.5)

44 **R:** >Thank you have good night.<

45 (0.2) Bye

46 (1.2)

47 **D:** OKAY (.) BYE-BYE

48 ... End call ...

Токены в третьей закрывающей позиции (строка 39) произносятся быстрее по сравнению с предыдущими токенами клерка (строка 34), а также увеличивается их количество. Это может говорить о том, что в заключительной части разговора клерк начинает прилагать больше усилий к его закрытию.

Отметим, что, во-первых, рассмотренные выше смены последовательностей инициируются только человеком в местах их возможного завершения. Занимая статус К-, клерк инициирует новые последовательности до момента достижения эпистемического равновесия, что характерно для подобного институционального разговора [Drew, Heritage, 1992: 54]. Однако робот не меняет тему, не инициирует новую последовательность и не продолжает ее в момент достижения эпистемического баланса, что уже не относится к вопросам эпистемики. Несмотря на то, что ему доступны места потенциального завершения последовательности, в которых может быть открыта новая, робот остается в статусе реципиента, побуждая клерка быть модератором [Müller, 2009: 138]. Например, в строках 34—35 в момент, когда подтверждение от клерка получено и наступает длинная пауза, где было бы уместно роботу вступить в последовательность и инициировать предзакрытие.

Однако он этого не делает, что, возможно, побуждает R мгновенно переходить к предзакрытию в строке 39.

Во-вторых, анализ ожидаемо показал, что первые шесть смен последовательности происходят пошагово, так как разговор на общую тему состоит из связанных между собой подтем. Иными словами, каждая следующая смена подтемы осуществляется постепенно, потому что она отсылает к теме общего запроса.

В-третьих, человек постепенно инициировал новые темы до тех пор, пока не получил все детали, необходимые для выполнения запроса, то есть до момента достижения эпистемического баланса между говорящими. Когда эпистемическое равновесие достигнуто, переход в заключительную часть разговора происходит не последовательно, а дизъюнктивно. Это указывает на доведение темы до конца, а также на готовность клерка закончить разговор.

Фактически разговор не заканчивается в тот момент, когда должно произойти закрытие. Для анализа этой проблематичной ситуации рассмотрим просодический контур говорящих и его влияние на исход завершения разговора.

Заключительная часть разговора: препятствия к завершению

Несмотря на то, что подтверждающие токены (“okay got it yeh”) являются предзакрытием разговора при эпистемическом балансе и вводятся дизъюнктивно, что указывает на готовность клерка заканчивать разговор, его завершению предшествует следующая последовательность (строки 41—42):

- 36 **R:** S::::o. It will be two people at seven o’clock↑
37 (1.6)
38 **D:** YES Mmhmm=
- 39→ **R:** =Mmhmm (.) okay↑ (.) got it↓↓yeh (0.2)
40 Thank you↑
41→ (1.9)
42→ **D:** OK:AY A:::WESOME (.) THANKS A LOT↓

Робот вступает в разговор в строке 42, но его черед не продвигает завершение, а расширяет его: Duplex заметно задерживается с ответом (строка 41), растягивает первые два слова и говорит громче, чем в предыдущих чередах (строка 42).

Как мы выяснили ранее, робот во всех частях разговора выступает в роли реципиента. Подобная реципиентность легитимна для этого типа сервисного звонка, где инициирует последовательности говорящий с К—статусом. Однако в заключительной части разговора реципиентность робота, которая растягивает закрытие разговора, остается непонятной.

Реципиентность в ситуациях аффилирования и дизаффилирования

Для анализа закрывающей секции разговора рассмотрим две ситуации использования токенов подтверждения реципиентом (“hm”, “uh huh”, “yeah”, “yes”, “right” и другие): «аффилирования» (“affiliation”) и «дизаффилирования» (“disaffiliation”) [Müller, 2009: 138—139].

Аффилирование представляет собой ситуацию, когда модератор подходит к заключительной части разговора — например, запрашивает подтверждение (“requests clarification”), и реципиент подтверждает запрос с помощью соответ-

ствующего токена, который своей синтаксической конструкцией и просодией согласуется с последовательностью говорящего-модератора. Просодически такие токены произносятся с пониженной громкостью, более низким или ровным тоном, при этом речь может быть ускорена. Подобный токен с описанными интонационными характеристиками может закрывать разговор [Müller, 2009: 140—149].

Однако, когда модератор инициирует заключительную часть разговора, а реципиент уклоняется (“moves out”) от закрытия и подобное расширение последовательности не поощряется, это характеризует ситуацию дизаффилирования. При этом токены подтверждения могут отсутствовать или звучать «не в тон» (“out of tune”), быть «не на месте» (“out of place”), «вне ритма» (“out of rhythm”). Тогда модератор может управлять сложившейся ситуацией, не позволяя реципиенту задерживать закрытие и сводя к минимуму дальнейшее расширение [Müller, 2009: 149—159].

Реципиентность Duplex: ситуация дизаффилирования

Клерк, который является модератором данного разговора, готов к его закрытию. Как мы выяснили ранее, на это указывает дизъюнктивный ввод предзакрытия (строка 39), после того как был достигнут эпистемический баланс. Кроме того, на завершенность также указывает просодический контур этого предзакрытия, который характеризуется интонационным спадом. Следующий черед клерка (строка 40)—токен благодарности, произнесенный со слегка повышающейся интонацией,—запрашивает подтверждение закрытия:

- 36 R: S::::o. It will be two people at seven o'clock↑
37 (1.6)
38 D: YES Mmhmm=
- 39→ R: =Mmhmm (.) okay↑ (.) got it ↓↓ yeah (0.2)
- 40→ Thank you↑
41→ (1.9)
42→ D: OK:AY A::WESOME (.) THANKS A LOT↓

В то же время Duplex, хотя синтаксически подтверждает предзакрытие (строка 42), то есть не собирается вводить новую тему, демонстрирует просодические несоответствия заключительной части разговора: делает заметно длинную паузу (строка 41), растягивает слова и говорит громче. Из этого следует, что токены подтверждения (“Okay awesome”) и благодарности (“Thanks a lot”) робота-реципиента находятся в ситуации дизаффилирования, то есть просодически «не на месте».

Клерк пресекает просодическое расширение заключительной части разговора с помощью ускоренной и тихой речи, а также вводит прощание (строка 45):

- 44 R:> Thank you have good night.<
45→ (0.2) Bye

Заключительный черед робота также сопровождается паузой, хоть и меньшей по длительности (строка 46), однако произносится громко и с ровной интонацией (строка 47). Это говорит о том, что робот продолжает совершать терминальный обмен, просодически не соответствующий действиям клерка:

- 46→ (1.2)
47→ D: OKAY (.) BYE-BYE
48 ... End call ...

Подводя итог: в заключительной части разговора просодия робота находится в ситуации дизаффилирования и затягивает закрытие, вступая в противоречие с намерением и готовностью клерка закончить разговор. Тем самым разговор продолжается, но ничего нового не происходит; прощание задерживается. Это противоречие разрешается быстро, благодаря ограничительным действиям клерка — вступлению в черед с минимальной заминкой, ускоренному темпу высказываний и вводу токена прощения. Подобные действия могут объясняться тем, что «для принимающего вызов это совершенно рутинный, уже десятый звонок, который он получил в этот день» [Drew, Heritage, 1992: 51]. В источнике идет речь о звонке в экстренную службу, но в нашем примере для клерка это столь же привычный, будничный звонок, поэтому он быстрее хочет его закончить.

Характерные черты Duplex как препятствие к завершению разговора

К рассмотренным речевым чертам робота (повышенная громкость, ровный тон, растянутость речи) добавим, что выявленные продолжительные задержки с ответом (строка 41, 46) в заключительной части разговора ничем не заполняются и никак не маркируются. Их нельзя легитимировать совмещением двух активностей, как в случае с клерком в нашем примере (строки 27 и 33), который одновременно запрашивает и записывает информацию. Таким образом, задержки клиента с ответом в нашем примере можно объяснить только тем, что в качестве говорящего выступает робот. Несмотря на то что разработчики данной системы запрограммировали использование “hmm”s, “uh”s и пр. для того, чтобы сигнализировать собеседнику обработку задачи и нивелировать ожидание⁸, наш анализ показывает, что у робота не всегда есть подобные маркеры, что создает проблемные моменты в разговоре. В отличие от человека «оправдать» задержки робота в этом кейсе можно только долгой обработкой данных.

Таким образом, препятствуют завершению беседы растянутость, громкость, ровный тон речи и заметно длинные паузы робота в заключительной части разговора, где эти роботические черты неуместны.

Выводы

В данном исследовании был проанализирован разговор между человеком-оператором и роботической системой Google Duplex с точки зрения закрытия данного разговора. Мы рассмотрели, каким образом организуются темы и последовательности в разговоре, с помощью эпистемики выяснили, как достижение эпистемического баланса способствует закрытию разговора, а также выявили присущие роботу черты, которые влияют на данное закрытие.

Мы искали ответы на два ключевых вопроса: 1) как демонстрируемые роботом черты связаны с институциональными ролями говорящих, и 2) с помощью каких средств происходит закрытие данного разговора. Наш анализ показал, что клерк, обладая статусом К-, в течение всего разговора выступает в роли модератора: задает вопросы, контролирует введение тем, инициирует фактическое

⁸ Leviathan Y., Matias Y. Google AI Blog: Google Duplex: An AI System for Accomplishing Real-World Tasks Over the Phone. 2018. 8 May. URL: <https://ai.googleblog.com/2018/05/duplex-ai-system-for-natural-conversation.html> (accessed: 18.02.2021).

предзакрытие разговора, а также первую парную часть терминального обмена (“bye”). В то же время клиент-робот, обладая статусом К+, является реципиентом — отвечает клерку. В данном случае институциональная роль говорящего и его эпистемический статус соответствуют друг другу: для говорящего, выступающего в роли профессионала, характерно управлять последовательностями в разговоре и задавать вопросы. Этим профессионалом является оператор со статусом К-, который инициирует вопросы, а также вопросительные декларативы, для того чтобы достичь эпистемического равновесия с собеседником. Равным образом, для говорящего, выступающего в роли клиента, также характерно давать ответы, что совпадает с его эпистемическим статусом К+. Кроме того, в момент достижения эпистемического равновесия оператор инициирует предзакрытие, что тоже характерно для институционального разговора, в котором клерк, выполнив запрос, начинает завершать разговор. При этом инициирование фактического предзакрытия оператором осуществляется дизъюнктивно, что указывает на его готовность заканчивать разговор. Таким образом, до заключительной части роботические черты не влияют на происходящее взаимодействие.

В заключительной части робот по-прежнему ведет себя как реципиент, однако просодический контур его ответов начинает конфликтовать с намерением клерка завершить разговор и задерживает закрытие. Duplex, с одной стороны, продолжает разговор, а с другой — не говорит ничего нового, его действия не способствуют закрытию. Синтаксически робот поддерживает предзакрытие, а также само прощание, но его просодия противоречива. Именно специфические черты речи робота (заметные задержки с ответом, растянутая речь, высокая громкость, ровная интонация), а не его институциональная роль задерживают разговорное закрытие, так как эти просодические черты находятся «не на месте».

Наблюдаемое уклонение от закрытия решается благодаря институциональной роли оператора. Текущий звонок должен закончиться, когда будет выполнен запрос и если не будут инициированы другие, поэтому проблематичное завершение разговора управляет клерком с помощью речевых ограничений на затянутое прощание, что в конечном итоге приводит к закрытию разговора.

Дискуссия

В данной работе мы проанализировали, как функционирует разговорное закрытие в специфическом кейсе, когда один из говорящих — робот. Полученные результаты закладывают основу для детального исследования организации разговора человека с продвинутым искусственным интеллектом и способствуют более глубокому пониманию того, как завершается разговор в подобных пограничных случаях. Несмотря на то что в данном разговоре собеседники обмениваются предзакрытиями, а затем прощаниями, его успешному завершению препятствует не подходящий для закрывающей части просодический контур робота. При дальнейшем изучении разговоров между человеком и искусственным интеллектом необходимо учитывать просодические особенности робота, так как они влияют на протекание сервисных звонков. В целом просодический контур должен рассматриваться как значимый ресурс для успешного завершения разговора.

Кроме того, в настоящей статье мы предприняли попытку соединить вопросы эпистемики, анализ просодического контура и организацию последовательностей в разговоре. Это дает целостное представление о том, в какой момент разговора собеседники намереваются его закончить, каким образом они этого достигают, какие проблематичные ситуации возникают и как говорящие их решают.

Неисследованным остался вопрос, как роботизированное эпистемическое позиционирование влияет на разговорные результаты и динамику информационной асимметрии между взаимодействующими. В рамках этой статьи робот занимал позицию «знающего», не способного первым завершить разговор, в последующих исследованиях было бы продуктивно изучить влияние «незнания» робота на результаты разговора.

Наконец, наши результаты наглядно демонстрируют, как программирование искусственного интеллекта может привести к неожиданным последствиям для разговора, и задают курс на основанное на эмпирических данных проектирование интерфейсов. Обладая более натуральным звучанием по сравнению с другими роботами, Google Duplex очень похож на человека [O'Leary, 2019]. Несмотря на это, сделанные выводы показывают, что взаимодействие между оператором и роботом не протекает гладко, так как просодия робота не настроена на успешное прощание.

Список литературы

- Сакс Х., Щеглофф Э., Джейферсон Г. Простейшая систематика организации очередности в разговоре // Социологическое обозрение. 2015. Т. 14. № 1. С. 142—202. URL: <https://sociologica.hse.ru/2015-14-1/147275640.html> (дата обращения: 22.02.2021)
- Sacks H., Schegloff E., Jefferson G. (2015) A Simplest Systematics for the Organization of Turn-Taking for Conversation. *Sociological Review*. Vol. 14. No. 1. P. 142—202. URL: <https://sociologica.hse.ru/2015-14-1/147275640.html> (accessed: 22.02.2021). (in Russ.)
- Broth M., Mondada L. (2013) Walking Away: The Embodied Achievement of Activity Closings in Mobile Interaction. *Journal of Pragmatics*. Vol. 47. No. 1. P. 41—58. <https://doi.org/10.1016/j.pragma.2012.11.016>.
- Button G. (1990) Going Up a Blind Alley: Conflating Conversation Analysis and Computational Modelling. In: Luff P., Gilbert N., Fronlich D. *Computers and Conversation*. London: Academic Press. P. 67—90. <https://doi.org/10.1016/b978-0-08-050264-9.50009-9>.
- Button G., Lee J. R. E. (1987) Talk and Social Organisation. *Intercommunication*. Vol. 1. Clevedon: Multilingual Matters.
- Couper-Kuhlen E. (2004) Prosody and Sequence Organization in English Conversation. In: Couper-Kuhlen E., Ford C. E. (eds.) *Sound Patterns in Interaction. Typological Studies in Language*. Vol. 62. Amsterdam: John Benjamins Publishing Company. P. 335—376. <https://doi.org/10.1075/tsl.62.17cou>.

Couper-Kuhlen E., Selting M. (2009) Towards an Interactional Perspective on Prosody and a Prosodic Perspective on Interaction. In: Couper-Kuhlen E., Selting, M. (eds.) *Prosody in Conversation*. Cambridge: Cambridge University Press. P. 11—56. <https://doi.org/10.1017/cbo9780511597862.003>.

Drew P, Heritage J. (1992) Talk at Work: Interaction in Institutional Settings. *Studies in Interactional Sociolinguistics*. Vol. 8. Cambridge: Cambridge University Press.

Haddington P. (2019) Leave-Taking as Multiactivity: Coordinating Conversational Closings With Driving in Cars. *Language & Communication*. Vol. 65. P. 58—78. <https://doi.org/10.1016/j.langcom.2018.04.003>.

Heritage J. (1989) Current Developments in Conversation Analysis. In: Roger D., Bull P. (eds.) *Conversation: An Interdisciplinary Perspective*. Clevedon: Multilingual Matters. P. 21—47.

Heritage J. (2012a) Epistemics in Action: Action Formation and Territories of Knowledge. *Research on Language and Social Interaction*. Vol. 45. No. 1. P. 1—29. <https://doi.org/10.1080/08351813.2012.646684>.

Heritage J. (2012b) Epistemics in Conversation. In: Sidnell J., Stivers, T. (eds.) *The Handbook of Conversation Analysis*. Chichester: Wiley-Blackwell. P. 370—394. <https://doi.org/10.1002/9781118325001.ch18>.

Heritage, J. (2012c) The Epistemic Engine: Sequence Organization and Territories of Knowledge. *Research on Language and Social Interaction*. Vol. 45. No. 1. P. 30—52. <https://doi.org/10.1080/08351813.2012.646685>.

Heritage J. (2013) Action Formation and Its Epistemic (and Other) Backgrounds. *Discourse Studies*. Vol. 15. No. 5. P. 551—578. <https://doi.org/10.1177/1461445613501449>.

Hoey E. M. (2015) Lapses: How People Arrive at, and Deal With, Discontinuities in Talk, *Research on Language and Social Interaction*. Vol. 48 No. 4. P. 430—453. <https://doi.org/10.1080/08351813.2015.1090116>.

Jefferson G. (1984) Notes on a Systematic Deployment of the Acknowledgement Tokens «Yeah»; and «Mm Hm». *Paper in Linguistics*. Vol. 17. No. 2. P. 197—216. <https://doi.org/10.1080/08351818409389201>.

Jefferson G. H. (2004) Glossary of Transcript Symbols With an Introduction. In: Lerner G. H. (ed.), *Conversation Analysis: Studies From the First Generation*. *Pragmatics & Beyond New Series*. Vol. 125. Amsterdam: John Benjamins Publishing. P. 13—31. <https://doi.org/10.1075/pbns.125.02jef>.

Klowait N. (2017) A Conceptual Framework for Researching Emergent Social Orderings in Encounters With Automated Computer-Telephone Interviewing Agents. *The International Journal of Communication and Linguistic Studies*. Vol. 15. No. 1. P. 19—37. URL: https://cgscholar.com/bookstore/works/a-conceptual-framework-for-researching-emergent-social-orderings-in-encounters-with-automated-computertelephone-interviewing-ag?category_id=cgrn&path=cgrn%2F228%2F230 (accessed: 18.02.2021).

Klowait N. (2018) Reflexive Anthropomorphism: Ontological Ignorance, or Ignorant Ontology? *Sociological Journal*. Vol. 24. No. 1. P. 8—33. <https://doi.org/10.19181/socjour.2018.24.1.5711>.

Klowait N., Erofeeva M. A. (2021) The Rise of Interactional Multimodality in Human-Computer Interaction. *Monitoring of Public Opinion: Economic and Social Changes*. No. 1. P. 46—70.

Markman K. M. (2006) Computer-Mediated Conversation: The Organization of Talk in Chat-Based Virtual Team Meetings, Ph.D. diss... Department of Communication Studies. Austin, TX: The University of Texas at Austin. URL: <http://hdl.handle.net/2152/2582> (accessed: 18.02.2021).

Mondada L. (2014) Pointing, Talk, and the Bodies. In: Seyfeddinipur M., Gullberg M., Kendon A. (eds.) *From Gesture in Conversation to Visible Action as Utterance: Essays in Honor of Adam Kendon*. Amsterdam; Philadelphia, PA: John Benjamins Publishing Company. P. 95—124. <https://doi.org/10.1075/z.188.06mon>.

Moore R. J., Arar R. (2019) Conversational UX Design: A Practitioner's Guide to the Natural Conversation Framework. New York, NY: ACM Books. <https://doi.org/10.1145/3304087>.

Müller F. E. (2009) Affiliating and Disaffiliating With Continuers: Prosodic Aspects of Recipiency. In: Couper-Kuhlen E., Selting M. (eds.) *Prosody in Conversation: Interactional Studies*. Cambridge: Cambridge University Press. P. 131—176.

O'Leary D.E. (2019) GOOGLE'S Duplex: Pretending to be Human. *Intelligent Systems in Accounting, Finance and Management*. Vol. 26. No. 1. P. 46—53. <https://doi.org/10.1002/isaf.1443>.

Pojanapunya P., Jaroenkitboworn K. (2011) How to Say «Good-Bye» in Second Life. *Journal of Pragmatics*. Vol. 43. No. 14. P. 3591—3602. <https://doi.org/10.1016/j.pragma.2011.08.010>.

Raymond G. (2004) Prompting Action: The Stand-Alone «So» in Ordinary Conversation. *Research on Language & Social Interaction*. Vol. 37. No. 2. P. 185—218. https://doi.org/10.1207/s15327973rlsi3702_4.

Sacks H., Schegloff E.A., Jefferson, G. (1974) A Simplest Systematics for the Organization of Turn-Taking for Conversation. *Language*. Vol. 50. No. 4. P. 696—735. <https://doi.org/10.1353/lan.1974.0010>.

Schegloff E. A. (2009) A Practice for (Re-)Exiting a Sequence: And/But/So + Uh(m) + Silence. In: Turner, K. (ed.) *Language in Life, and a Life in Language: Jacob Mey, a Festschrift. Studies in Pragmatics*. Bingley: Emerald. P. 365—374. https://doi.org/10.1163/9789004253209_047.

Schegloff E. A., Sacks H. (1973) Opening up Closings. *Semiotica*. Vol. 8. No. 4. P. 289—327. <https://doi.org/10.1515/semi.1973.8.4.289>.

Selting M. (1996) On the Interplay of Syntax and Prosody in the Constitution of Turn-Constructional Units and Turns in Conversation. *Pragmatics*. Vol. 6. No. 3. P. 371—388. <https://doi.org/10.1075/prag.6.3.06sel>.

Wooffitt R. (1997) Humans, Computers and Wizards: Analysing Human (Simulated) Computer Interactions. London: Routledge.

Zellers M. (2012) Prosodic Variation for Topic Shift and Other Functions in Local Contrasts in Conversation. *Phonetica*. Vol. 69. No. 4. P. 231—253. <https://doi.org/10.1159/000353445>.

Приложение 1.

Список символов транскрипта:

[Левая квадратная скобка означает, что на начало речи текущего говорящего накладывается речь другого.

R: Fi[ve,

D: [It's fo::r two people↓

= Знак равно означает, что нет промежутка между словами.

R: °Mmhmm° (0.9) °em Let me see°=

D: =Mmhmm

(0.0) Число в скобках означает истекшее время в десятых долях секунды.

R: E:n what time you will come,

(1.2)

D: At seven pm

(.) Точка в скобках означает короткий интервал (\pm одна десятая секунды) в пределах или между высказываниями.

D: I: m calling to make reservation? (.) I'm looking for a table on Friday the fourth↓

: Двоеточие означает продление непосредственно предшествующего звука. Чем больше двоеточий, тем больше удлинение.

R: S:o how many people,

(1.1)

R: Fi[ve,

D: [It's fo::r two people↓

— Подчеркивание обозначает форму ударения.

R: E:n what time you will come,

(1.2)

D: At seven pm

↑↓ Стрелки означают повышение или понижение тона.

D: YES Mmhmm↑=

R: =Mmhmm↑ (.) okay↑ (.) got it↓↓yeh (0.2)

, Запятая означает направление интонации вверх, что обычно соответствует вопросительной форме.

R: S:o how many people,

(1.1)

R: Fi[ve,

. Точка означает слегка поникающийся тон.

R: S:::o. It will be two people at seven o'clock↑

? Вопросительный знак связан с вопросительной интонацией.

D: Hey, I:’m calling to make reservation? (.) I’m looking for a table on Friday the fourth↓

WORD Прописные буквы означают высокую громкость по отношению к окружающему разговору.

D: OK:AY A::WESOME (.) THANKS A LOT↓

(0.5)

R:>◦Thank you have good night◦<

(0.2) Bye

(1.2)

D: OKAY (.) BYE-BYE

◦word◦ Знаки градуса, заключающие высказывание или его часть, означают, что звуки тише, чем окружающий разговор.

D: At seven pm

(0.9)

R: ◦Seven◦ (0.9) Can you say name,

><word Знаки больше и меньше означают повышенную скорость высказывания по отношению к общему темпу разговора.

R:>◦Thank you have good night◦<

() Пустые круглые скобки означают невозможность понять, что говорится.

R: ()

Приложение 2

Транскрипт

00 ((Ring))
01 (0.3)
02 **R:** ()
03 (0.9)
04 **D:** Hey
05 I: 'm calling to make reservation? (.) I'm looking for a table
on Friday the fourth↓
06 (1.2)
07 **R:** This Friday,
08 (1.0)
09 **D:** Yeah
10 (0.5)
11 **R:** °Mmhmm°
12 (0.9) °em Let me see°=
13 **D:** =Mmhmm
14 (0.9)
15 **R:** S:o how many people,
16 (1.1)
17 **R:** Fi[ve,
18 **D:** [It's fo::r two people↓
19 (0.4)
20 **R:** For two,
21 (1.2)
22 **D:** Yes
23 (0.6)
24 **R:** E:n what time you will come,
25 (1.2)
26 **D:** At seven pm
27 (0.9)
28 **R:** °Seven°
29 (0.9)
30 Can you say name,
31 (1.3)
32 **D:** E:::HM (0.3) FIRST NAME IS YANIV (0.5) Y A (0.2) N I (.) V
33 (2.7)
34 **R:** Okay↑ (0.2) I got it↓
35 (2.9)
36 **R:** S:::o. It will be two people at seven o'clock↑
37 (1.6)
38 **D:** YES Mmhmm=
39 **R:** =Mmhmm (.) okay↑ (.) got it↓yeh (0.2)
40 Thank you↑
41 (1.9)

- 42 **D:** OK:AY A::WESOME (.) THANKS A LOT↓
43 (0.5)
44 **R:** >Thank you have good night.<
45 (0.2) Bye
46 (1.2)
47 **D:** OKAY (.) BYE-BYE
48 ... End call ...

DOI: [10.14515/monitoring.2021.1.1736](https://doi.org/10.14515/monitoring.2021.1.1736)



**И. А. Блохин, С. П. Морозов, В. Ю. Чернина, А. Е. Андрейченко,
И. В. Шахабов, А. В. Смышляев, В. А. Гомболевский**

ИСПОЛЬЗОВАНИЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В ЗДРАВООХРАНЕНИИ: ОПЫТ ВАЛИДАЦИИ АЛГОРИТМА ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В МЕДИЦИНСКИХ ОРГАНИЗАЦИЯХ В УСЛОВИЯХ ПАНДЕМИИ COVID-19

Правильная ссылка на статью:

Блохин И. А., Морозов С. П., Чернина В. Ю., Андрейченко А. Е., Шахабов И. В., Смышляев А. В., Гомболевский В. А. Использование искусственного интеллекта в здравоохранении: опыт валидации алгоритма искусственного интеллекта в медицинских организациях в условиях пандемии COVID-19 // Мониторинг общественного мнения: экономические и социальные перемены. 2021. № 1. С. 271—282. <https://doi.org/10.14515/monitoring.2021.1.1736>.

For citation:

Blokhin I. A., Morozov S. P., Chernina V. Y., Andreychenko A. E., Shakhabov I. V., Smyshlyaev A. V., Gombolevskiy V. A. (2021) Artificial Intelligence in Healthcare: Validating an AI Algorithm in Health Institutions in the COVID-19 Pandemic (a Use Case). *Monitoring of Public Opinion: Economic and Social Changes*. No. 1. P. 271–282. <https://doi.org/10.14515/monitoring.2021.1.1736>. (In Russ.)

ИСПОЛЬЗОВАНИЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В ЗДРАВООХРАНЕНИИ: ОПЫТ ВАЛИДАЦИИ АЛГОРИТМА ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В МЕДИЦИНСКИХ ОРГАНИЗАЦИЯХ В УСЛОВИЯХ ПАНДЕМИИ COVID-19

БЛОХИН Иван Андреевич — младший научный сотрудник сектора исследований в лучевой диагностике, Научно-практический клинический центр диагностики и телемедицинских технологий Департамента здравоохранения города Москвы

E-MAIL: i.blokhin@npcmr.ru

<https://orcid.org/0000-0002-2681-9378>

МОРОЗОВ Сергей Павлович — доктор медицинских наук, профессор, директор, Научно-практический клинический центр диагностики и телемедицинских технологий Департамента здравоохранения города Москвы, Москва, Россия

E-MAIL: morozov@npcmr.ru

<https://orcid.org/0000-0001-6545-6170>

ЧЕРНИНА Валерия Юрьевна — руководитель сектора исследований в лучевой диагностике, Научно-практический клинический центр диагностики и телемедицинских технологий Департамента здравоохранения города Москвы, Москва, Россия

E-MAIL: chernina909@gmail.com

<https://orcid.org/0000-0002-0302-293X>

АНДРЕЙЧЕНКО Анна Евгеньевна — руководитель сектора медицинской информатики, радиомикси и радиогеномики, Научно-практический клинический центр диагностики и телемедицинских технологий Департамента здравоохранения города Москвы, Москва, Россия

E-MAIL: a.andreychenko@npcmr.ru

<https://orcid.org/0000-0001-6359-0763>

ARTIFICIAL INTELLIGENCE IN HEALTH-CARE: VALIDATING AN AI ALGORITHM IN HEALTH INSTITUTIONS IN THE COVID-19 PANDEMIC (A USE CASE)

Ivan A. BLOKHIN¹—Junior Researcher at the Radiology Research Sector

E-MAIL: i.blokhin@npcmr.ru

<https://orcid.org/0000-0002-2681-9378>

Sergei P. MOROZOV¹—Dr. Sci. (Med.), Professor, Director of the Research and Practical Clinical Center for Diagnostics and Telemedicine Technologies of the Moscow Health Care Department

E-MAIL: morozov@npcmr.ru

<https://orcid.org/0000-0001-6545-6170>

Valeria Yu. CHERNINA¹—Head of Radiology Research Sector

E-MAIL: chernina909@gmail.com

<https://orcid.org/0000-0002-0302-293X>

Anna E. ANDREYCHENKO¹—Head of Medical Informatics, Radiomics and Radiogenomics Division

E-MAIL: a.andreychenko@npcmr.ru

<https://orcid.org/0000-0001-6359-0763>

¹ Research and Practical Clinical Center for Diagnostics and Telemedicine Technologies of the Moscow Health Care Department, Moscow, Russia

ШАХАБОВ Ислам Висханович — кандидат медицинских наук, научный сотрудник отдела научных медицинских исследований, Научно-практический клинический центр диагностики и телемедицинских технологий Департамента здравоохранения города Москвы, Москва, Россия

E-MAIL: islam75@mail.ru

<https://orcid.org/0000-0001-6409-1384>

СМЫШЛЯЕВ Алексей Викторович — кандидат медицинских наук, ведущий научный сотрудник отделения научных основ организации амбулаторной помощи, ФГБУ «Центральный научно-исследовательский институт организации и информатизации здравоохранения Министерства здравоохранения Российской Федерации», Москва, Россия

E-MAIL: alexeysmishlyaev@yandex.ru

<https://orcid.org/0000-0003-3099-2517>

ГОМБОЛЕВСКИЙ Виктор Александрович — кандидат медицинских наук, руководитель отдела научных медицинских исследований, Научно-практический клинический центр диагностики и телемедицинских технологий Департамента здравоохранения города Москвы, Москва, Россия

E-MAIL: g_victor@mail.ru

<https://orcid.org/0000-0003-1816-1315>

Аннотация. В статье рассматриваются актуальные проблемы современности в сфере охраны здоровья граждан. Многочисленные вызовы вынуждают предпринимать действия, повышающие скорость оказания медицинской помощи и ее доступность при сохранении качества. Одним из инновационных решений этого вопроса может быть внедрение систем автоматического анализа данных, основанных на искусственном интеллекте.

Islam V. SHAKHABOV¹— Cand. Sci.

(Med.),

Researcher at the Medical Research Department

E-MAIL: islam75@mail.ru

<https://orcid.org/0000-0001-6409-1384>

Aleksey V. SMYSHLYAEV²— Cand. Sci. (Med.), Leading Researcher of the Department of Scientific Foundations of Organization of Outpatient Care

E-MAIL: alexeysmishlyaev@yandex.ru

<https://orcid.org/0000-0003-3099-2517>

Viktor A. GOMBOLEVSKIY¹— Cand. Sci. (Med.), Head of Medical Research Department

E-MAIL: g_victor@mail.ru

<https://orcid.org/0000-0003-1816-1315>

¹ Research and Practical Clinical Center for Diagnostics and Telemedicine Technologies of the Moscow Health Care Department, Moscow, Russia

² Federal Research Institute for Health Organization and Informatics of Ministry of Health of the Russian Federation, Moscow, Russia

Abstract. The paper considers new challenges related to public health. Action is needed to improve access to healthcare while maintaining its quality. The introduction of AI-based automated data analysis systems can be a solution to that. The present study seeks to assess the use of AI in outpatient care to detect pathological changes in the lungs typical of a coronavirus amidst the pandemic. The sample size was 600 patients. The

Проведено исследование, цель которого — оценить эффективность применения в амбулаторных условиях в период пандемии искусственного интеллекта, направленного на выявление патологических изменений в легких, характерных для COVID-19. Объем выборки составил 600 пациентов, проведены статистическая и аналитическая обработка результатов. Установлено, что чувствительность метода составила 94 %, специфичность — 77 %, точность — 83 %, площадь под характеристической кривой — 87 %. Зафиксирован уровень прогностической ценности отрицательного результата — 97 %, положительного результата — 66 %. Полученные данные свидетельствуют о надежном отсеивании алгоритмом результатов компьютерной томографии без признаков патологических изменений в легких. Сделан вывод о возможности применения искусственного интеллекта в здравоохранении на современном этапе — в условиях пандемии COVID-19 это крайне актуально и позволяет увеличить эффективность диагностики заболеваний. Алгоритмы ИИ также могут быть использованы для медицинской сортировки пациентов вне условий пандемии, что повысит доступность медицинской помощи.

Ключевые слова: искусственный интеллект, медицинская помощь, лучевая диагностика, здравоохранение, пандемия, COVID-19

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Финансирование. Авторы заявляют об отсутствии финансирования всех этапов данного исследования.

results were statistically and analytically processed. The sensitivity attained 94%; the specificity, accuracy and the area under the ROC curve were 77%, 83%, and 87%, respectively. The negative predictive value was 97%; the positive predictive value was 66%. The data obtained show that the algorithm separates the CT scan results having no abnormalities in the lungs. The authors conclude that the usage of AI technologies helped to improve diagnostic accuracy during the COVID-19 pandemic. Artificial intelligence algorithms can also work with patients in non-pandemic times, thus improving healthcare access.

Keywords: artificial intelligence, medical aid, radiology, healthcare, pandemic, COVID-19

Благодарность. Авторы выражают глубокую признательность коллективам отделений лучевой диагностики медицинских организаций, оказывающих первичную медико-санитарную помощь взрослому населению Москвы. Авторы благодарят за сплоченную работу коллектив Департамента информационных технологий города Москвы и ООО «Лаваль», обеспечивающих внедрение искусственного интеллекта в практическое здравоохранение на уровне крупного мегаполиса. Кроме этого, авторы отдельно благодарят разработчиков компании Binomics ray. Каждый из участников процесса приложил важные усилия к исследованию в непростой период эпидемии.

Введение

Проблемы старения населения, увеличения потока пациентов, значительный рост баз данных и научно-методических материалов, подлежащих обработке — все это вызовы для здравоохранения. На фоне этих изменений повышается нагрузка на врачебный персонал, нехватка которого ощущима в настоящее время. Одновременно с этим система здравоохранения быстро меняется, трансформируясь из аналоговой в цифровую информационную среду и набирая огромные массивы данных. Вычислительные мощности развиваются по закону Мура, удваиваясь каждые два года. Эти компоненты обуславливают необходимость и возможность анализа больших данных. В свете вышеперечисленного возможным решением представляется внедрение систем автоматического анализа данных, основанных на алгоритмах машинного обучения — «искусственного интеллекта» (далее — ИИ). Валидированный алгоритм ИИ сопоставим с работой врача в отношении качества анализа материала, но превосходит его по скорости обработки и охвату данных, что способствует совершенствованию лечебно-диагностических процессов. Другими словами, возможности ИИ позволяют осуществлять дополнительное обучение на клинических данных, оптимизировать лечение, прогнозирование и тактику ведения пациента, улучшать процесс отбора пациентов для участия в клинических испытаниях, снижать количество ошибок. Благоприятная основа для развития алгоритмов ИИ — структурированные и стандартизованные данные, именно поэтому в нашем случае они представляют наибольшую ценность.

Повышение нагрузки на медицинский персонал ведет к увеличению вероятности врачебных ошибок из-за влияния человеческого фактора. Это требует поиска новых инструментов, направленных на повышение скорости и качества работы медицинского персонала. ИИ в такой ситуации может оказаться чрезвычайно

эффективным. В перспективе возможна передача ИИ части врачебных компетенций первого (элементарного) уровня.

Для национальных систем здравоохранения наиболее остро встает вопрос использования ИИ в чрезвычайных ситуациях. Так, в январе 2020 г. Всемирная организация здравоохранения объявила чрезвычайную ситуацию из-за распространения новой коронавирусной инфекции COVID-19, и национальные системы здравоохранения были вынуждены перейти к работе в «нештатном» режиме. В период пандемии COVID-19 актуальность использования ИИ в здравоохранении резко возросла: это обусловлено повышением уровня заболеваемости и смертности населения. По официальным данным, уже к середине 2020 г. в мире было зарегистрировано более 6 млн подтвержденных случаев заболевания и более 375 тыс. летальных случаев [Dong et al., 2020b].

ИИ необходим главным образом для повышения доступности медицинской помощи через оптимизацию сортировки пациентов, то есть для определения вероятности заболевания и формирования «первого мнения» до интерпретации данных врачом. Так, многие вендоры¹ представляют свои решения ИИ для анализа рентгенографии органов грудной клетки [Murphy et al., 2020], однако компьютерная томография (далее — КТ) имеет более высокую диагностическую ценность в отношении тактики ведения пациента [Lee et al., 2020]. Помимо этого, ИИ или другие методы количественного анализа изображений необходимы для ускорения анализа медицинских изображений при эпидемии COVID-19 [Dong et al., 2020a]. В настоящее время исследования в области ИИ привели к использованию таких инструментов, как объективная оценка объема неповрежденной ткани легкого по сравнению с пораженной [Belfiore et al., 2020]. Однако до сих пор отсутствуют публикации, посвященные использованию ИИ как эффективного инструмента повышения доступности лечебно-диагностических мероприятий посредством увеличения пропускной способности диагностических каналов через сортировку пациентов на ранних этапах диагностики. Стоит отметить, что в условиях эпидемии появление автоматизированного инструмента, способного повлиять на сортировку первичных пациентов, представляет собой важный организационно-управленческий аспект в борьбе с COVID-19 [Dong et al., 2020b]. До настоящего времени не проводилось сравнение возможностей ИИ на подготовленном тестовом датасете и в реальной практике для триажа (сортировки) пациентов в условиях пандемии [Fang et al., 2020]. Все это становится более актуальным в контексте текущих эпидемиологических вызовов, которые испытывают на прочность национальные системы здравоохранения.

Цель исследования — оценить эффективность ИИ в диагностике заболеваний в условиях медицинских организаций, оказывающих амбулаторную помощь во время пандемии COVID-19.

Материалы и методы

Для оценки эффективности ИИ в медицинских организациях, оказывающих помощь в амбулаторных условиях, проведено ретроспективное многоцентровое

¹ Вендор — компания, которая разрабатывает и/или поставляет заказчикам собственные продукты, технологии, услуги под своим товарным знаком (брендом).

диагностическое исследование. Организационно-методическое обеспечение, статистическая и аналитическая обработка результатов выполнена сотрудниками ГБУЗ «Научно-практический клинический центр диагностики и телемедицинских технологий ДЗМ» (далее — НПКЦ ДИТ ДЗМ). Осуществлялась валидация алгоритма ИИ по выявлению патологических изменений в легких, характерных для COVID-19. В исследование были включены анонимизированные результаты компьютерной томографии органов грудной клетки из 53 амбулаторных медицинских организаций города Москвы.

Оценка степени поражения COVID-19 по данным КТ органов грудной клетки на территории Российской Федерации проводится на основании визуальной полуколичественной шкалы [Лучевая диагностика..., 2020]. Визуальный количественный анализ изменений на КТ органов грудной клетки помогает оценивать клиническую тяжесть COVID-19 [Sverzellati et al., 2020]. По КТ-данным оценивался объем нормально аэрированного легкого при первоначальной КТ органов грудной клетки, коррелирующий с прогнозом у пациентов, поступивших в приемное отделение неотложной помощи с пневмонией.

Объем выборки составил 600 пациентов европеоидной расы: 265 мужчин (средний возраст $54 \pm 17,3$ лет) и 335 женщин (средний возраст $52,6 \pm 16,8$ лет). При анализе результаты исследований были разделены авторами на три группы: группа № 1 «Вирусные пневмонии» (ранние признаки COVID-19) ($n=172$); группа № 2 «Неизмененные легкие» ($n=195$), группа № 3 «Другие патологии в легких» ($n=174$). Критерии для включения в группу 1: инфильтрация легочной паренхимы по типу матовых стекол с обеих сторон, преимущественно периферической локализации; инфильтрация легочной паренхимы по типу «бульжной мостовой» (утолщение междоллькового интерстиция на фоне «матового стекла») с обеих сторон, преимущественно периферической локализации; инфильтрация легочной паренхимы по типу консолидации с положительным признаком воздушной бронхограммы при наличии пункта 1 или пункта 2. Критерий для включения в группу 2 — отсутствие признаков вирусной пневмонии по результатам проведения КТ органов грудной клетки, в группу 3 — наличие патологического процесса в легких, отличного от пневмонии.

Каждое исследование интерпретировано в два этапа. Первично — врачом-рентгенологом, непосредственно проводившим исследование; второе чтение одномоментно проводилось двумя независимыми экспертами до достижения консенсуса. При отсутствии расхождений во мнении исследование относили к одной из трех групп. Алгоритм ИИ состоял из многоэтапной обработки исследования нейронной сетью и был интегрирован в Единый радиологический информационный сервис (далее — ЕРИС), охватывающий 53 отделения лучевой диагностики амбулаторно-поликлинического звена г. Москвы. Результаты проводимых исследований подвергались автоматизированному анализу. Врач, непосредственно проводивший исследования, оценивал решения алгоритма посредством формы обратной связи, интегрированной в ЕРИС. Далее исходное исследование передавалось экспертам для ретроспективной оценки. По завершении сбора данных произведено сопоставление решений алгоритма с результатами референс-теста, построена четырехпольная таблица, осуществлены расчет показателей (чувствительность,

специфичность, точность (общая валидность), отношение правдоподобия и прогностическая ценность положительного/отрицательного результата), построена и проанализирована характеристическая кривая (ROC-кривая). Для каждого из указанных диагностических показателей рассчитан 95 % доверительный интервал (ДИ).

Проводилась статистическая и аналитическая обработка результатов.

Результаты

Результаты исследований были разделены на три группы: Группа 1 — «Вирусные пневмонии» (ранние признаки COVID-19) ($n=172$); Группа 2 — «Неизмененные легкие» ($n=195$); Группа 3 — «Другие патологии в легких» ($n=174$). Все результаты обследования проанализированы с применением оригинального алгоритма ИИ. ИИ-сервис обработал 541 обследование (89,3 %). Решения алгоритма ИИ сопоставлены с референс-тестом, значения которого приняты за истинный статус обследуемых лиц. В результате получено 162 истинно-положительных результата, 285 — истинно-отрицательных, 84 — ложноположительных, 10 — ложноотрицательных. Установлены следующие показатели: чувствительность — 94 %, специфичность — 77 %, точность — 83 %, площадь под характеристической кривой — 87 %. В условиях пандемии COVID-19 и связанного с нею непрерывного потока пациентов, которым проводилось обследование, протестированный алгоритм ИИ показал высокую прогностическую ценность отрицательного результата (97 %). Это означает крайне высокую гарантию отсутствия признаков заболевания (вирусной пневмонии) в случаях, которые были отнесены алгоритмом ИИ к группам 2 и 3. Данный результат представляется наиболее значимым, доказывающим возможность применения технологий ИИ для сортировки пациентов в амбулаторных условиях. Вместе с тем оцениваемый алгоритм ИИ имеет относительно низкое значение прогностической ценности положительного результата (66 %). Это означает, что в условиях пандемии COVID-19 и высокой априорной вероятности наличия признаков соответствующего заболевания (вирусной пневмонии) ценность алгоритма ИИ, указывающего на наличие у пациента пневмонии, вызванной именно COVID-19, представляется низкой.

Полученные нами значения свидетельствуют об удовлетворительном качестве алгоритма ИИ. Чувствительность индекс-теста составляет 94 %, специфичность — 77 %, точность — 83 %. Эксперты, проводившие анализ результатов работы алгоритма, отметили следующие особенности взаимодействия: качественная оценка локализации неудовлетворительная, присутствие ложноположительных (нормальная ткань легкого, фиброзные изменения, рак легкого) и ложноотрицательных (не отмечено целиком «матовое стекло») отметок на изображении; представление результатов работы ИИ-сервиса ограничено выборочными срезами, что вызывает неудобство при практическом использовании данных врачом-рентгенологом; DICOM SR (структурированный отчет) содержит много лишней информации, замедляющей работу врача-рентгенолога с результатами ИИ-сервиса.

Обсуждение

Увеличение потока исследований в условиях пандемии инфекционного заболевания требует быстрого принятия решений для сортировки пациентов на тех, кому тре-

буется медицинская помощь прямо сейчас, и тех, кому понадобится плановый уход. В связи с этим актуальна задача автоматической медицинской сортировки (триажа) КТ-исследований. С помощью такой сортировки можно существенно сократить время от момента завершения процедуры до получения врачебного заключения путем повышения приоритета в рабочем списке КТ-исследований с подозреваемой патологией для более быстрой интерпретации рентгенологом. Примером может служить полученная с помощью специального программного обеспечения количественная оценка поражения легочной паренхимы: была установлена высокая корреляция между оставшейся хорошо вентилируемой легочной тканью и исходами (перевод в палату интенсивной терапии или смерть) [Colombi et al., 2020]. Помимо этого, поражение легких при пневмонии COVID-19 можно оценить количественно автоматически, используя алгоритмы машинного обучения [Li et al., 2020].

Однако медицинскую сортировку (триаж) исследований с целью облегчения взаимодействия профессионально-коммуникативной модели «врач—врач» и «врач—пациент» следует проводить в рамках принятых территориально или профессионально систем классификаций «находок» (обнаруженных патологий). Изучаемый алгоритм ИИ оценивает вероятность наличия изменений, характерных для коронавирусной инфекции, при КТ органов грудной клетки, но не классифицирует их в привычном для врача-рентгенолога виде (по степеням от КТ-1 до КТ-4) [Gozes et al., 2020].

С управлеченческо-организационной точки зрения актуальным представляется разделение медицинской сортировки исследований на две подзадачи: отделение пациентов с COVID-19 от всех остальных (включая норму); сортировка пациентов с COVID-19 по степени тяжести. В зависимости от потока пациентов в конкретном медицинском учреждении значимость задач меняется. Для амбулаторно-поликлинического звена больший приоритет имеет обнаружение пациентов с высоким риском наличия новой коронавирусной инфекции для их дальнейшей изоляции. Для стационарного звена, в частности для перепрофилированных учреждений, важна оценка степени тяжести вирусной пневмонии, в том числе динамики контрольных исследований.

Заключение

В условиях непрерывного и зачастую всевозрастающего потока исследований в амбулаторно-поликлиническом звене здравоохранения протестированный алгоритм ИИ отличается высокой прогностической ценностью отрицательного результата (97 %). Это означает крайне высокую гарантию отсутствия признаков вирусной пневмонии в случаях, которые были отнесены алгоритмом ИИ к группам 2 и 3. Этот результат представляется наиболее значимым, доказывающим возможность применения технологий ИИ для сортировки пациентов по данным КТ-исследований в условиях первичного звена. Однако оцениваемый алгоритм ИИ имеет относительно низкое значение прогностической ценности положительного результата (66 %). Это означает, что в условиях пандемии COVID-19 и высокой пред-тестовой вероятности наличия признаков вирусной пневмонии практическая ценность алгоритма ИИ, указывающего на наличие у пациента вирусной пневмонии COVID-19, представляется низкой.

В целом алгоритм ИИ обеспечивает медицинскую сортировку КТ-исследований в условиях первичного (амбулаторно-поликлинического) звена здравоохранения с выделением исследований без патологии, что имеет низкую ценность при высокой априорной вероятности наличия изменений, характерных для вирусной пневмонии. По мнению экспертов, применение алгоритмов машинного обучения способно повысить доступность и качество медицинской помощи [Thrall et al., 2017], однако широкое внедрение ИИ в медицину должно производиться в соответствии с этическими и правовыми нормами [Pesapane et al., 2018]. Следует также отметить тревожную тенденцию в виде возникновения огромного количества алгоритмов, не прошедших валидацию на внешних данных. Так, в анализе 516 публикаций из баз данных PubMed MEDLINE и Embase за 2018 год только в 6 % исследований (31 публикация) проводилась валидация на внешних данных [Kim et al., 2019]. При этом ни одно из проанализированных исследований не имело пропективный дизайн или многоцентровой характер. Таким образом, проведенное исследование показало эффективность использования ИИ в условиях медицинских организаций, оказывающих помощь в амбулаторных условиях. Валидация алгоритма искусственного интеллекта по данным компьютерной томографии проведена без предварительной калибровки. Это свидетельствует о надежном «отсеивании» алгоритмом результатов исследований без признаков заболевания.

Выводы

Использование ИИ в диагностике заболеваний в условиях медицинских организаций, оказывающих амбулаторную помощь в период пандемии COVID-19, представляется, по мнению авторов, эффективным как с лечебно-диагностической, так и с организационно-управленческой точек зрения. Применение ИИ в здравоохранении позволяет повысить эффективность диагностики заболеваний. Применение алгоритмов ИИ в условиях пандемии COVID-19 обоснованно в части проведения медицинской сортировки результатов КТ органов грудной клетки. Использование ИИ в работе амбулаторно-поликлинического звена здравоохранения позволяет повысить количественные показатели доступности оказания медицинской помощи населению через увеличение пропускной способности (нагрузки) диагностических подразделений при сохранении надлежащего качества оценки результатов КТ-исследований на этапе медицинской сортировки.

Список литературы (References)

Лучевая диагностика коронавирусной болезни (COVID-19): организация, методология, интерпретация результатов. Препринт № ЦДТ. 2020. II. Версия 2 от 17.04.2020/ сост. Морозов С.П., Проценко Д.Н., Сметанина С.В. и др. М.:ГБУЗ «НПКЦ ДИТ ДЗМ», 2020. (Серия «Лучшие практики лучевой и инструментальной диагностики». Вып. 65). URL: https://tele-med.ai/documents/346/luchevaya_diagnostika_koronavirusnoj_infekcii_covid-19_v2_2.pdf (дата обращения: 03.02.2021).

Radiation Diagnosis of Coronavirus Disease (COVID-19): Organization, Methodology, Interpretation of the Results. (2020) Preprint No. CDT-2020-II. Version 2 dated 04/17/2020. Comp. Morozov S.P., Protsenko D.N., Smetanina S.V. et al. Series “Best

Practices of Radiation and Instrumental Diagnostics". Vol. 65. M.: GBUZ "NPKTs DiT DZM". URL: https://tele-med.ai/documents/346/luchevaya_diagnostika_koronavirusnoj_infekcii_covid-19_v2_2.pdf (accessed: 03.02.2021). (In Russ.)

Belfiore M.P., Urraro F., Grassi R., Giacobbe G., Patelli G., Cappabianca S., Reginelli A. (2020) Artificial Intelligence to Codify Lung CT in COVID-19 Patients. *La Radiologia Medica*. Vol. 125. No. 5. P. 500—504. <https://doi.org/10.1007/s11547-020-01195-x>.

Colombi D., Bodini F.C., Petrini M., Maffi G., Morelli N., Milanese G., Silva M., Sverzella-ti N., Michieletti E. (2020) Well-Aerated Lung on Admitting Chest CT to Predict Adverse Outcome in COVID-19 Pneumonia. *Radiology*. Vol. 296. No. 2. P. E86—E96. <https://doi.org/10.1148/radiol.2020201433>.

Dong D., Tang Zh. Wang Sh., Hui H., Gong L., Lu Y., Xue Zh., Liao H., Chen F., Yang F., Jin R., Wang K., Liu Zh., Wei J., Mu W., Zhang H., Jiang J., Tian J., Li H. (2020a) The Role of Imaging in the Detection and Management of COVID-19: A Review. *IEEE Reviews in Biomedical Engineering*. Vol. 14. P. 16—29. <https://doi.org/10.1109/RBME.2020.2990959>.

Dong E., Du H., Gardner L. (2020b) An Interactive Web-Based Dashboard to Track COVID-19 in Real Time. *The Lancet Infectious Diseases*. Vol. 20. No. 5. P. 533—534. [https://doi.org/10.1016/s1473-3099\(20\)30120-1](https://doi.org/10.1016/s1473-3099(20)30120-1).

Fang Y., Zhang H., Xie J., Lin M., Ying L., Pang P., Ji W. (2020) Sensitivity of Chest CT for COVID-19: Comparison to RT-PCR. *Radiology*. Vol. 296. No. 2. P. E115—E117. <https://doi.org/10.1148/radiol.2020200432>.

Gozes O., Frid-Adar M., Greenspan H., Browning P.D., Zhang H., Ji W., Bernheim A., Siegel E. (2020) Rapid AI Development Cycle for the Coronavirus (COVID-19) Pandemic: Initial Results for Automated Detection & Patient Monitoring using Deep Learning CT Image Analysis. *arXiv:2003.05037 [eess IV]*. URL: <https://arxiv.org/abs/2003.05037> (accessed: 3.02.2021).

Kim D.W., Jang H.Y., Kim K.W., Shin Y., Park S.H. (2019) Design Characteristics of Studies Reporting the Performance of Artificial Intelligence Algorithms for Diagnostic Analysis of Medical Images: Results from Recently Published Papers. *Korean Journal of Radiology*. Vol. 20. No. 3. P. 405—410. <https://doi.org/10.3348/kjr.2019.0025>.

Lee E.Y.P., Ng M.-Y., Khong, P.-L. (2020) COVID-19 Pneumonia: What Has CT Taught Us? *The Lancet Infectious Diseases*. Vol. 20. No. 4. P. 384—385. [https://doi.org/10.1016/S1473-3099\(20\)30134-1](https://doi.org/10.1016/S1473-3099(20)30134-1).

Li K., Fang Y., Li W., Pan C., Qin P., Zhong Y., Liu X., Huang M., Liao Y., Li Sh. (2020) CT Image Visual Quantitative Evaluation and Clinical Classification of Coronavirus Disease (COVID-19). *European Radiology*. Vol. 30. No. 8. P. 4407—4416. <https://doi.org/10.1007/s00330-020-06817-6>.

Murphy K., Smits H., Knoops A.J.G., Korst M.B.J.M., Samson T., Scholten E.T., Schalekamp S., Schaefer-Prokop C.M., Philipsen R.H.H.M., Meijers A., Melendez J., van Ginneken B., Rutten M. (2020) COVID-19 on Chest Radiographs: A Multireader

Evaluation of an Artificial Intelligence System. *Radiology*. Vol. 296. No. 3. P. E166—E172. <https://doi.org/10.1148/radiol.2020201874>.

Pesapane F., Volonté C., Codari M., Sardanelli F. (2018) Artificial Intelligence as a Medical Device in Radiology: Ethical and Regulatory Issues in Europe and the United States. *Insights Into Imaging*. Vol. 9. P. 745—753. <https://doi.org/10.1007/s13244-018-0645-y>.

Sverzellati N., Milanese G., Milone F., Balbi M., Ledda R. E., Silva M. (2020) Integrated Radiologic Algorithm for COVID-19 Pandemic. *Journal of Thoracic Imaging*. Vol. 35. No. 4. P. 28—233. <https://doi.org/10.1097/RTI.0000000000000516>.

Thrall J. H., Li X., Li Q., Cruz C., Do S., Dreyer K., Brink J. (2018) Artificial Intelligence and Machine Learning in Radiology: Opportunities, Challenges, Pitfalls, and Criteria for Success. *Journal of the American College of Radiology*. Vol. 15. No. 3. P. 504—508. <https://doi.org/10.1016/j.jacr.2017.12.026>.

DOI: [10.14515/monitoring.2021.1.1821](https://doi.org/10.14515/monitoring.2021.1.1821)



Н. Д. Трегубова, М. Л. Ни, А. А. Китаева

ТРАНСНАЦИОНАЛИЗМ В РЕЖИМЕ ОНЛАЙН: МИГРАЦИОННЫЕ ИССЛЕДОВАНИЯ В УСЛОВИЯХ ИСКУССТВЕННОЙ СОЦИАЛЬНОСТИ

Правильная ссылка на статью:

Трегубова Н.Д., Ни М.Л., Китаева А.А. Транснационализм в режиме онлайн: миграционные исследования в условиях искусственной социальности // Мониторинг общественного мнения: экономические и социальные перемены. 2021. № 1. С. 283—303. <https://doi.org/10.14515/monitoring.2021.1.1821>.

For citation:

Tregubova N. D., Nee M. L., Kitaeva A. A. (2021) Transnationalism Online: Migration Research in the Age of Artificial Sociality. *Monitoring of Public Opinion: Economic and Social Changes*. No. 1. P. 283–303. <https://doi.org/10.14515/monitoring.2021.1.1821>. (In Russ.)

ТРАНСНАЦИОНАЛИЗМ В РЕЖИМЕ ОНЛАЙН: МИГРАЦИОННЫЕ ИССЛЕДОВАНИЯ В УСЛОВИЯХ ИСКУССТВЕННОЙ СОЦИАЛЬНОСТИ

ТРЕГУБОВА Наталья Дамировна — кандидат социологических наук, ассистент кафедры сравнительной социологии, Санкт-Петербургский государственный университет, Санкт-Петербург, Россия
E-MAIL: n.tregubova@spbu.ru
<https://orcid.org/0000-0003-3259-5566>

НИ Максим Леонидович — сотрудник Международной исследовательской лаборатории «Транснационализм и миграционные процессы: сравнительный и институциональный анализ», Санкт-Петербургский государственный университет, Санкт-Петербург, Россия
E-MAIL: jintae@mail.ru
<https://orcid.org/0000-0002-7285-0237>

КИТАЕВА Анастасия Александровна — студент, Совместная магистерская программа «Наука и технологии в обществе», Университет ИТМО и Европейский университет в Санкт-Петербурге, Санкт-Петербург, Россия
E-MAIL: kitnsty@yandex.ru
<https://orcid.org/0000-0002-8883-8013>

Аннотация. Миграция в Россию из стран СНГ — явление, которое ставит перед социальными учеными многочисленные исследовательские проблемы. Разнообразие миграционных потоков и трудности в доступе к «полю» приводят к тому, что различные фрагменты общей картины лишь в той или иной степени согласуются друг с другом. Статья посвящена осмыслению и сравнению нескольких таких фрагментов, полученных в ходе изучения «следов» активности транснациональных мигрантов

TRANSNATIONALISM ONLINE: MIGRATION RESEARCH IN THE AGE OF ARTIFICIAL SOCIALITY

Natalia D. TREGUBOVA¹ — Cand. Sci. (Soc.), Assistant Professor at the Chair of Comparative Sociology
E-MAIL: n.tregubova@spbu.ru
<https://orcid.org/0000-0003-3259-5566>

Maxim L. NEE¹ — Researcher at the International Research Laboratory “Transnationalism and Migration Processes: Comparative and Institutional Analysis”
E-MAIL: jintae@mail.ru
<https://orcid.org/0000-0002-7285-0237>

Anastasia A. KITAЕVA² — Student
E-MAIL: kitnsty@yandex.ru
<https://orcid.org/0000-0002-8883-8013>

¹ Saint Petersburg State University, Saint Petersburg, Russia

² ITMO University and EUSP Joint Master's Program 'Science and Technologies in Society', St. Petersburg, Russia

Abstract. Migration to Russia from the CIS countries poses various research problems for social scientists. Diversity of migration flows and difficulties in access to the field lead to various researchers seeing different fragments of the overall picture, which do or do not match with each other. This paper aims to interpret and compare several research findings obtained in the study of digital ‘traces’ of transnational migrants’ online activities. The paper seeks to answer two questions: Which aspects of migration

в интернете. Авторы стремятся ответить на два вопроса: какие стороны миграционных процессов проявляются в собранных данных, а какие — ускользают от исследователей; как полученные результаты и их ограничения соотносятся со свойствами онлайн-среды, организованной деятельностью поисковых, классифицирующих и сортирующих алгоритмов? Эмпирический материал представлен двумя типами онлайн-данных, характеризующих миграцию в Россию из пяти стран (Азербайджана, Казахстана, Кыргызстана, Таджикистана и Узбекистана). Это группы «ВКонтакте» и сайты, найденные с помощью поисковика Google.

Данные были собраны на шести языках: русском и стран исхода. В качестве методов выступают сетевой анализ, контент-анализ, а также онлайн-наблюдение. Охарактеризованы процессы сбора и анализа данных и основные результаты исследования для каждого из методов. Результаты соотнесены друг с другом, сделаны выводы о содержательных характеристиках миграционных процессов и о возможностях и ограничениях исследования миграции в онлайн-среде. Авторы описывают «разделение труда» между людьми и алгоритмами на каждом из этапов исследования, выделяют три типа алгоритмов согласно их роли в исследовательском процессе, а также размышляют над тем, как принятые методологические решения повлияли на полученные результаты.

Ключевые слова: миграция в Россию, транснациональная миграция, цифровая этнография, онлайн-наблюдение, сетевой анализ, взаимодействие «человек — алгоритм»

processes manifest in the collected data, and which ones elude from the researchers' view? How do the research findings and limitations correspond to the properties of the online milieu that are organised by search and classification algorithms? The research is based on two online data types: groups in VKontakte, the Russian largest social media, and sites found using the Google search engine. This data characterises migration to Russia from five countries: Azerbaijan, Kazakhstan, Kyrgyzstan, Tajikistan, and Uzbekistan, and was collected in six languages (Russian and the languages of the countries of origin). The methods of analysis are network analysis, online observation, and content analysis. The paper characterises data collection and analysis steps and presents the main research results for each of the methods. In conclusion, the authors discuss how the research findings are related to each other, what conclusions could be drawn about transnationalism in Russia, and how to organize a study in the online milieu considering various algorithms' active role.

Keywords: migration to Russia, transnational migration, digital ethnography, online observation, network analysis, human-algorithm interaction

Благодарность. Исследование выполнено при поддержке Российского научного фонда, грант № 18-78-10049.

Acknowledgments. The study was supported by the Russian Science Foundation, grant No. 18-78-10049.

Постановка проблемы

Миграция в Россию из стран СНГ — явление, которое ставит перед социальными учеными многочисленные исследовательские проблемы. Разнообразие миграционных потоков и трудности в доступе к «поля» приводят к тому, что различные фрагменты общей картины лишь в той или иной степени согласуются друг с другом. Статья посвящена осмыслинию и сравнению нескольких таких фрагментов, полученных в ходе анализа «следов» активности транснациональных мигрантов в интернете.

Транснациональный подход — один из наиболее популярных и перспективных в современных миграционных исследованиях. В самом общем виде транснационализм предполагает параллельное пребывание «здесь и там» — «одновременное включение мигранта в социальные сети как страны исхода, так и принимающего сообщества, характеризующееся постоянным движением мигранта между различными национальными пространствами, повседневными мирами, финансовым участием в экономике страны исхода и экономической активностью в принимающем государстве» [Степанов, 2018: 119].

Проект, результаты которого представлены в настоящей статье, направлен на фиксацию и анализ различных проявлений транснационализма онлайн для мигрантов из стран СНГ в Россию.

Концепция «транснационализма онлайн» предполагает, что онлайн-технологии расширяют возможности установления транснациональных связей и, в предельном выражении, изменяют саму природу транснационализма: «здесь и там» преображается в «здесь, там и где-то еще» [Starikov et al., 2018]. Проблематика транснациональной миграции на постсоветском пространстве представляет особый интерес в соотнесении с процессами, происходящими в онлайн-среде. С одной стороны, исследования последних лет показывают, что миграция в Россию, особенно миграция из стран Средней Азии, часто транснациональна [Абашин, 2017; Бредникова, 2017] и мигранты активно используют интернет в повседневной жизни [Тимошкин, 2019; Ruget, Usmanalieva, 2019; Urinboyev, 2017]. С другой стороны, существует важная концептуальная проблема, связанная с осмыслинием влияния на миграционные процессы технологий, которые делают возможными различные виды опыта пребывания онлайн [Трегубова, 2020].

Авторы настоящей статьи ставят перед собой две цели.

Первая — представить результаты конкретных эмпирических исследований, проведенных в рамках вышеупомянутого исследовательского проекта, посвященного проблематике транснационализма онлайн.

Вторая цель заключается в осмыслинии опыта исследования, в котором в качестве «поля» выступает онлайн-среда. Данные, генерируемые онлайн, становятся все более востребованы социальными учеными, появляются новые методы сбора и анализа этих данных, возникают имена для новых областей: Internet studies

[Рыков, Нагорный, 2017], digital humanities [Digital Humanities: гуманитарные науки в цифровую эпоху, 2016], digital ethnography [Hjorth et al., 2017], digital sociology [Сивков, 2018], и др. Настоящая статья представляет рефлексию об опыте исследования онлайн, а именно — о том, как в исследовательском процессе воплощаются особенности развития «онлайн-культуры» и «искусственной социальности» [Резаев, Трегубова, 2019]. Иными словами, нас интересуют а) тот «сплав» технологических возможностей, культурно обусловленных практик и повсеместного проникновения интернета в повседневную жизнь общества, который характеризует жизнь современного человека; б) влияние алгоритмов, активных посредников и участников онлайн-взаимодействий, в том числе — влияние алгоритмов искусственного интеллекта (ИИ). Как следствие этих двух факторов, в любом онлайн-исследовании возникает своеобразное «разделение труда» между человеком и машиной (алгоритмом), которое мы и стремимся осмыслить в данной статье.

В рамках настоящей статьи эти цели взаимосвязаны, причем первая подчинена второй. Результаты исследований транснационализма онлайн изложены и соотнесены между собой таким образом, чтобы они могли стать материалом для методологической рефлексии о проведении онлайн-исследований в условиях искусственной социальности.

В качестве эмпирического материала выступают два типа данных: группы в социальной сети «ВКонтакте» (далее — ВК) и сайты, найденные с помощью поисковика Google. Решение о сборе именно таких данных связано с тем, что социальные сети и поисковики сегодня используются практически повсеместно и с различными целями. Выбор поисковика Google обусловлен тем, что это крупнейший международный поисковик, используемый в различных языковых сегментах интернета и имеющий национальные домены. ВК был выбран потому, что это самая популярная социальная сеть в России и одна из наиболее популярных — в странах СНГ [Тимошкин, 2019]. Полученные сведения характеризуют миграцию и транснациональные связи для мигрантов в Россию из пяти стран: Азербайджана, Казахстана, Киргизстана, Таджикистана и Узбекистана¹. Сбор данных на шести языках (русском и стран исхода) проходил с октября 2019 г. по май 2020 г.

Исследования транснациональной миграции зачастую ограничиваются фиксацией связи между Родиной и страной исхода. Однако в рамках настоящего проекта была поставлена более широкая задача: мы стремились зафиксировать существование транснационального пространства, которое объединяло бы все шесть стран, а не только Россию с отдельными странами-донорами. Данная постановка вопроса связана с интересом к тому, как общее советское наследие влияет на характер транснационализма мигрантов в России.

В соответствии с поставленными целями авторы статьи стремятся дать ответы на два вопроса.

¹ Выбор именно этих стран связан с тремя обстоятельствами. Во-первых, величина миграционных потоков: по данным МВД РФ (URL: <https://xn--b1aew.xn--p1ai/Deljatelnost/statistics/migracionnaya/item/19365693/> (дата обращения: 10.02.2021.)), в 2019 г. Азербайджан, Казахстан, Киргизстан, Таджикистан и Узбекистан входили в число семи государств, из которых в Россию приезжает наибольшее число мигрантов (также в эту семерку вошли Китай и Украина). Во-вторых, переселенцы из этих стран в отличие, например, от украинцев, часто идентифицируют местным населением именно как мигранты. В-третьих, страны объединяют то, что в каждой из них преобладает мусульманское население.

Первый: какие стороны миграционных процессов проявляются в собранных данных, а какие — ускользают от исследователей? В частности, фиксируются ли «следы» общего транснационального пространства, связывающего страны исхода с Россией?

Второй: как полученные результаты и их ограничения соотносятся со свойствами онлайн-среды, организованной деятельностью поисковых, классифицирующих и сортирующих алгоритмов?

Отвечая на первый вопрос, мы характеризуем результаты применения к собранным данным различных методов анализа. Для ответа на второй вопрос мы рассмотрим «разделение труда» между людьми и алгоритмами в ходе исследования и обсудим, как могла бы повлиять на полученные результаты иная организация исследования.

Сайты, найденные через Google: контент-анализ

Поиск сайтов по миграционной тематике через Google осуществлялся на шести языках по ключевым словам: «мигрант русский», «местный русский», «мигрант азербайджанец/казах/киргиз/узбек/таджик», «местный азербайджанец/казах/киргиз/узбек/таджик» (на русском языке — 12 словосочетаний, на азербайджанском, казахском, киргизском, таджикском и узбекском языках — по четыре словосочетания с указанием соответствующей национальности). Поиск производился в национальных доменах для соответствующих языков; для всех языков, кроме русского, — в режиме инкогнито.

Для каждого из словосочетаний просматривались сайты, выдаваемые поисковиком на первых десяти страницах, и из них вручную отбирались те, что имеют отношение к миграционным процессам на постсоветском пространстве. При необходимости содержание сайтов переводилось с помощью переводчика Google. Всего было отобрано более 700 сайтов (от 100 до 200 для каждого языкового сегмента).

Для анализа текстов, представленных на сайтах, применялся метод качественного контент-анализа. Внимание исследовательского коллектива было направлено на следующие характеристики: 1) содержание: тип сайта, характер информации, целевая аудитория; 2) язык (языки) сайтов; 3) совпадения: мы фиксировали случаи, когда на разных языках поисковик выдавал одни и те же сайты². Основные типы сайтов, найденные нами через Google, — это информационные (новостные) порталы, коллекции видео, сайты организаций (правительственных, международных, правозащитных, предоставляющих услуги мигрантам и др.), персональные сайты, сообщества пользователей (включая социальные сети), сайты с нормативными документами, онлайн-словари, а также специальные сайты по миграционной тематике. Наиболее распространены новостные и информационные порталы — они составляют около половины сайтов, попавших в выборку.

Ключевой вопрос на данном этапе реализации проекта был сформулирован следующим образом: *Какие источники информации по миграционной тематике оказываются общими для национальных сегментов интернета?*

² Детальное изложение результатов и выводов данного этапа реализации исследовательского проекта см. в [Лицицын, Трегубова, Орлова, 2021].

В основном поиск на разных языках обнаруживал разные сайты. Если формализовать сравнение по названиям сайтов до домена, то лишь 1/9 из них совпадет для двух и более языков. Таким образом, информация о миграции оказывается в значительной степени замкнутой в границах шести национальных сегментов интернета. Тем не менее проведенный анализ обнаружил некоторые элементы, связывающие отдельные языки, страны и сегменты интернета.

В отношении новостей страны оказываются связаны или через общие дискурсы, или через региональную принадлежность. Основные новостные источники, общие для шести языковых сегментов,— это «Радио Свобода», для которого характерен правозащитный дискурс, и «Спутник»— российский проект информационного портала, представляющего «многополярный мир». Два новостных канала транслируют разные дискурсы: направленный на защиту прав человека и «западные» образцы и центрированный на Россию и ее влияние в мире.

В отношении национальных интернет-доменов связи между странами свидетельствуют скорее об отсутствии единого транснационального пространства новостей о миграции, так как эти связи представляют вовсе не самые мощные миграционные потоки. Российское информационное пространство новостей включает новости азербайджанского и казахстанского доменов (но не наоборот), а кыргызстанское — новости российских СМИ. Узбекистан и Таджикистан оказываются вне этих информационных обменов. Кроме того, для поиска по-киргизски, по-узбекски и по-таджикски важный источник информации — коллекции видео, которые отсутствуют при поиске на других языках. Таким образом, сама форма донесения информации оказывается разной для разных языков.

Основная региональная принадлежность в новостных порталах— это Средняя/Центральная Азия. Она объединяет Казахстан, Кыргызстан, Россию, Таджикистан и Узбекистан, но не Азербайджан. Некоторые ее конкретизации — Туркестан, Фергана, Таджикистан и Центральная Азия. В сравнении с ней принадлежность к Кавказу представлена слабо. Принадлежность к исламу, которая могла бы связывать шесть стран, также представлена слабо и при этом практически не увязывается с региональной принадлежностью. В целом же в информационном пространстве Россия и страны Средней Азии оказываются умеренно связанными между собой, а Азербайджан находится в стороне от них.

Возвращаясь к общей выборке сайтов, найденных при помощи Google, мы фиксируем следующие особенности распределения по языкам и странам.

Для поиска на русском языке более распространены сайты о миграции, сайты организаций, предлагающих услуги мигрантам, сайты землячеств и иных ассоциаций, а также правозащитных, и международных организаций, персональные сайты и информационные порталы о межнациональных отношениях.

Для поиска на киргизском, узбекском и таджикском языках, по сравнению с другими языками, более важными источниками информации являются видео, а также сайты, посвященные непосредственно миграционному процессу. Кроме того, на киргизском языке были найдены специальные форумы, созданные пользователями для взаимопомощи в адаптации в России. Сравнительно менее важны для этих трех языков оказываются сайты правозащитных и международных организаций, для узбекского и таджикского языка — сайты, предлагающие услуги мигрантам.

Для поиска на казахском и азербайджанском языках сравнительно более важными источниками информации являются сайты государственных, международных и правозащитных организаций. Сайты о миграции при поиске на этих языках обнаружены не были.

Таким образом, для мигрантов из Кыргызстана, Узбекистана и Таджикистана, с одной стороны, и мигрантов из Казахстана и Азербайджана, с другой стороны, оказываются важными разные источники информации. Первые ориентированы на сайты, предлагающие помочь в адаптации к действительным проблемам, вторые — на сайты различных регулирующих организаций, от государственных до правозащитных. Следует особо отметить, что сайты по миграционной тематике связывают Россию с основными странами-донорами — Кыргызстаном, Таджикистаном и Узбекистаном; связь эта определяется релевантной для мигрантов информацией об адаптации в России и знанием языков. Кроме того, для Кыргызстана характерна большая связь с русскоязычным информационным пространством, а также большая степень самоорганизации мигрантов. Мы связываем это с особым статусом мигрантов из Кыргызстана в связи с членством страны в ЕАЭС, предполагающим упрощенные процедуры по трудоустройству в России.

Наконец, к описанию содержательных характеристик онлайн-среды следует добавить некоторые важные формальные ее особенности.

Первая особенность заключается в том, что социальные сети и приложения оказываются потенциальными «точками роста» для транснациональных пространств, связанных с миграцией. Социальные сети важны, поскольку они популярны и представляют общую площадку для активностей и «сверху» (СМИ и представительства различных организаций), и «снизу» (сообщества пользователей и личные страницы). Приложения важны, так как они ориентированы на решения проблем мигрантов и часто имеют версии на разных языках.

Вторая особенность состоит в том, что домен «ru» иногда обозначает принадлежность к России, а иногда выступает как домен общего характера. В этом отношении он схож с тем, как употребляется русский язык на постсоветском пространстве. В целом отнесение сайта к национальному домену не гарантирует того, что содержание сайта будет относиться к искомой стране.

Третья особенность состоит в том, что для сайтов СМИ и различных организаций характерно включение отдельных версий на разных языках. Сайты с версиями на разных языках ориентированы на модель переключения между языками и национальными государствами, даже если апеллируют к иным идентичностям (например, региональной среднеазиатской). Переключение между языками — часто между русским и другим/другими языками — играет объединяющую роль в постсоветском транснационализме онлайн.

Группы «ВКонтакте»: сетевой анализ

Поиск по группам ВК осуществлялся по следующим группам морфем, характеризующим страны исхода с их этнотERRиториальным разнообразием: мигрант*, миграц*, тиhoj*, муҳоҷ*, мӯҳоҷ*, тімqr*, азәr*, azər*, казах*, казаш*, қазақ*, киргиз*, кыргыз*, кыргызч*, таджик*, таджич*, тоҷик*, узбек*, узбеч*, ўзбек* o'zbek*, ферган*, farg'on, каралакалпак*, карақалпак*, памир*, помир*, бадахшан*, ба-

дахшон*, дунган*, хуэйцз*, талыш*, толыш* *tolış**, тольш*, уйгур*, *uyghur**, уйфур*. Из групп, появившихся в выдаче по соответствующей морфеме, вручную отбирались те, что имеют отношение к тематике исследования. Кроме собственно миграционных, в выборку включались группы о стране/регионе (20 самых крупных) и группы об исламе. Всего было отобрано более 2000 групп. При необходимости содержание сайтов переводилось с помощью переводчика Google.

Полученная выборка составила основу для сетевого анализа: группы представляют собой узлы, а их общие участники — ребра. Ключевой вопрос, на который был ориентирован сетевой анализ: *Какие группы оказываются связанными между собой через общих участников?*

Информация о сообществах была получена при помощи API VK и далее структурирована при помощи функционала библиотеки NetworkX языка программирования python. Полученные данные далее визуализировались в форме графа, что позволило проанализировать структуру взаимосвязей между группами через их участников. Визуализация и анализ параметров графа производились с применением программного пакета Gephi. Визуализация укладки графа была построена с использованием алгоритма Force Atlas 2, представляющего граф как совокупность «пружин» (ребер) и «металлических колец» (узлов), для которого производится симуляция физических сил (притяжения/отталкивания), действующих в системе. Это позволяет визуально выделить вершины, имеющие высокую степень взаимодействия.

Для количественного анализа свойств графа рассматривался набор стандартных параметров каждого ребра и вершины:

- степень посредничества (Betweenness) дает число кратчайших путей в графе через эту вершину, показывает важность узла для связности;
- центральность по собственному значению (Eigenvector centrality) характеризует важность вершины, учитывая сумму важности степеней ее соседей;
- степень вершины (Degree) показывает количество ее связей;
- взвешенная мощность (Weighted Degree) показывает количество связей узла, учитывая их вес;
- метрика модулярности (Modularity) с учетом весов ребер использовалась для кластеризации — она выделяет сообщества, основываясь на плотности: плотность внутри сообществ предполагается выше, чем плотность в остальном графе.

Сетевой анализ был проведен для всей выборки, а также отдельно — для каждой подвыборки групп, найденных по соответствующим группам морфем (мигр* и морфемы, связанные со страной, регионом или национальностью на разных языках).

В целом центральность групп и кластеризация «по морфемам» сохранялась, когда группы объединялись в общую сеть. Это свидетельствует об устойчивой структуре сетевой активности, проходящей по национальным и региональным линиям.

В полном графе, включающем все группы, были выделены семь кластеров: 1) большинство — группы по морфеме мигр* — в основном группы юридической помощи мигрантам; 2) большинство — таджикистанские группы, а также памирские

и бадахшанские группы; 3) большинство — узбекистанские, а также ферганские и каракалпакские группы; 4) большинство — кыргызстанские группы; 5) большинство — казахстанские группы; 6) большинство — азербайджанские, а также талышские группы; 7) большинство — уйгурские и дунганские группы. Кластеры 1 и 3 оказываются более «размытыми», кластеры 5, 6 и 7 — напротив, более «чистыми», то есть они содержат практически только группы с преобладающими морфемами в названии. Таким образом, разделение кластеров проходит в целом в соответствии с морфемами. Причем регионы/этнические категории группируются по «своим» странам. Исключение составляют уйгуры и дунгане, выделяющиеся в одну группу (возможно, на основании их исторических связей с Китаем).

Интересные эффекты наблюдаются, когда из полного графа удаляются группы общенациональной и исламской тематики. Общая структура кластеров сохраняется, однако с некоторыми изменениями. Во-первых, выделяется кластер групп «казахстанцев» — русскоязычных переселенцев из Казахстана (в полном графе самые центральные группы этого кластера попадают в кластер 1). Это, по-видимому, свидетельствует о существовании отдельной группы переселенцев из Казахстана, объединенных прагматическим интересом, связанным с гражданско-правовой интеграцией в российское общество. Во-вторых, группы с морфемой дунган* объединяются в кластер с узбекистанскими группами, а группы с морфемой уйгур* остаются в отдельном кластере. Данное изменение позволяет выдвинуть предположение, что общность между дунганами и уйгурами основана на культурных и национальных символах, а не на миграционных практиках. В-третьих, от кластера 1 отделяется кластер с разного рода услугами и объявлениями для мигрантов, а от кластера 3 — кластер знакомств для мигрантов. Это измерение свидетельствует о дифференцированности мигрантской экономики/сферах услуг, которая не сводится лишь к юридическим и околоюридическим вопросам.

Внутри подвыборок групп также были выделены отдельные кластеры, однако разница между ними, оцениваемая с помощью выбранных параметров графа, была гораздо меньшей, чем для кластеров в полном графе. Однако она использовалась на следующем этапе при выборе групп для онлайн-наблюдения.

Группы «ВКонтакте»: онлайн-наблюдение

На третьем этапе исследования для более детального анализа из исходной выборки ВК были отобраны 23 группы. В качестве критериев отбора выступали: а) центральность (по результатам сетевого анализа); б) разнообразие групп (по принадлежности к кластерам, выделенным в ходе сетевого анализа); в) отчетливая тематизация миграции в Россию; г) интенсивность взаимодействий в группе.

Основной метод анализа здесь — онлайн-наблюдение, его выбор связан с тем, что исследовательский коллектив интересовали категории и смыслы, возникающие, поддерживающиеся и трансформирующиеся в ходе взаимодействия участников онлайн-групп. Таким образом, нас интересовали только те посты, которые становились предметом обсуждения. Онлайн-наблюдение предполагает, что анализ содержания постов и комментариев «вписан» в анализ динамики взаимодействия в группах. Это, в свою очередь, требует фиксации временной последовательности постов и комментариев, оценки интенсивности взаимодей-

ствий и эмоциональных реакций участников, а также учета таких специфических для онлайн-групп данных, как число лайков и комментариев. В качестве базовых единиц наблюдения рассматривается пост в группе и тред обсуждения (набор комментариев под постом). Пост принимается за стартовую точку взаимодействия, тред фиксирует его развитие.

Вопросы, ключевые для онлайн-наблюдения, могут быть сформулированы следующим образом: *в каких категориях участники групп осмысляют опыт пребывания в России? Как они идентифицируют себя и как их идентифицируют другие? Какие категории выступают в качестве объединяющих для мигрантов из разных стран?*

Для исследования взаимодействий в социальных сетях коллектив использовал теорию цепочек ритуалов взаимодействия Рэндалла Коллинза [Collins, 2004]. Для адаптации концептуальных положений данной теории под специфику онлайн-сообществ и выработки методики исследования мы обратились к работе Пола ДиМаджо и его коллег [DiMaggio et al., 2018]³.

Мы анализировали активность в группе начиная с января 2019 г. (для двух групп — с января 2017 г., чтобы охватить больше активности) по апрель 2020 г. В анализ включались посты и комментарии, имеющие отношение к миграционной тематике: посты, имеющие 10 и более комментариев, для небольших и средних групп и посты с 20 и более комментариями для больших групп. Посты просматривались вручную, их выдача упорядочивалась с помощью инструментария для анализа социальных сетей Popsteps.

Группы, которые были включены в анализ, подразделяются на четыре типа:

1) Сообщества мигрантов из определенной страны в определенном городе, созданные для общения, взаимопомощи и обмена информацией. К ним относятся одна таджикистанская, три кыргызстанские и две казахстанские группы (одна из них — группа русскоязычных переселенцев в самом Казахстане). Данный тип был для нас приоритетным: мы включали в анализ подобные сообщества, если обнаруживали их в нашей выборке.

2) «Информационные группы», включающие новости и различные сюжеты о стране исхода, в том числе по миграционной тематике. К данному типу относятся одна кыргызстанская, одна казахстанская, две азербайджанских, две таджикистанских группы (одна из них посвящена истории страны) и четыре узбекистанских группы с разной целевой аудиторией (узбекистанская молодежь за пределами страны, эмигранты 1990-х годов, смешанный состав).

3) Группы национально-культурных организаций в отдельном городе, которые представлены одной азербайджанской и двумя казахстанскими группами.

4) Группы юридической помощи или взаимопомощи для мигрантов из разных стран. Данный тип представляют четыре группы, отобранные по морфеме мигр*, разного размера и с разной ориентацией — от прагматической до правозащитной.

Общение в группах первого, второго и третьего типа обычно происходит на двух языках — русском и языке страны исхода, в то время как четвертый тип групп полностью русскоязычный.

³ Более подробно теоретико-методологические основания, а также методика и результаты онлайн-наблюдения изложены в работе «Beyond Nations and Nationalities: Discussing the Variety of Migrants' Identifications in Russian Social Media» [Tregubova, Nee, 2020].

По результатам исследования, динамика взаимодействий и основные темы в группах зависят не только от страны и от конкретного информационного повода. Важными также оказываются характеристики конкретной группы: ее размер, отбор и способ публикации материалов, особенности использования языка, количество участников и др.

В ходе анализа были обнаружены следующие устойчивые идентификации мигрантов⁴, характерные для всех типов групп. Первая — «трудовой мигрант — экономическая необходимость». Данная идентификация связывает трудовую миграцию с экономическим или политическим положением в стране исхода. Вторая — «мигрант — подлежащий правилам миграции как гражданин страны». Здесь часто обсуждается облегченный режим миграции для стран ЕАЭС и возможности получения российского гражданства. Еще три устойчивые идентификации связаны с тем, как сами мигранты видят себя внутри принимающего общества. Во-первых, «мигрант — тот, кто представляет нацию своим поведением». Здесь имеет место критика (или, реже, похвала) поведения «своих» мигрантов, призыв «вести себя хорошо», ибо по поведению одного будут судить обо всех. С этой самоидентификацией тесно связана другая: «мигрант — носитель национального языка». Здесь появляются призывы говорить на родном языке, которые часто связаны с идеей, что мигрант теряет свою культуру и язык, русифицируется. Наконец, еще одна идентификация — «мигрант — тот, кто воспринимается хуже, чем местные». Она проявляется при обсуждении трудностей в поиске работы, расового профайлинга, а также дискуссий о том, воспринимают ли мигрантов в России как людей «второго сорта».

Устойчивые идентификации, характерные для отдельных стран, включают две категории. Во-первых, это идентификация «низкоквалифицированные трудовые мигранты из Средней Азии», которая оказывается важна и для россиян (в представлении участников групп), и для самих мигрантов. При этом в Среднюю Азию обычно включаются Кыргызстан, Таджикистан и Узбекистан. Данная идентификация может быть поводом и для объединения, и для отмежевания. Так, мигранты из Казахстана и Азербайджана в своих взаимодействиях подчеркивают, что они «не гастарбайтеры». Во-вторых, для групп Казахстана и Азербайджана характерна самоидентификация «член диаспоры», причем в казахстанских группах оказывается важным разделение на казахов из России и из Казахстана.

Таким образом, для взаимодействий в группах ВК массовая трудовая миграция из Средней Азии (Узбекистан, Таджикистан, Кыргызстан) в Россию — это факт, с которым соотносят себя представители этих стран, и Казахстана, и Азербайджана. Следует также отметить, что дискrimинация мигрантов («гастарбайтеров») не всегда срабатывает как символ, вокруг которого строится взаимодействие. Значительная часть наблюдавшихся нами взаимодействий имеют прагматический характер или эмоции возникают вовсе не по тому поводу, который предлагают исходные посты.

Ислам присутствует в дискурсе онлайн-групп, но в разной степени и в разных вариациях. Он важен скорее для характеристики нации, чем для ситуации миграции, причем снацией он в разных случаях соотносится по-разному. Ислам

⁴ Термины «идентификация» и «самоидентификация» здесь и далее используются в соответствии с теорией Роджерса Брубейкера [Brubaker, 2004].

может выступать как объединяющая идентификация, но не как основная в случае миграции. Для характеристики миграционных процессов оказываются важны два сюжета: 1) представление о том, что мигрантов-мусульман не любят, 2) представление о том, что ислам легче практиковать в России, чем на родине (этот сюжет характерен для таджикистанских и узбекистанских групп).

Наследие советской идеологии — «дружба народов» и классификация на национальности — регулярно актуализируется во взаимодействиях. Зачастую — в усеченном виде, в качестве распространенного суждения «нет плохих наций, есть плохие люди», однако иногда пользователи прямо ссылаются на «дружбу народов» и ее реализацию в советской истории.

Наконец, транснационализм в рассматриваемых группах существует в трех основных формах. Во-первых, это сообщества людей из конкретной страны в конкретном городе; в основном — кыргызстанские сообщества. Во-вторых, это диаспоральные связи, которые характерны для групп, связанных с Казахстаном и Азербайджаном. В-третьих, это ностальгические взаимодействия русскоязычных мигрантов, преимущественно в узбекистанских группах. Кроме того, в группах юридической помощи возникают временные транснациональные связи между мигрантами из СНГ, которые угасают с получением желаемого статуса (ВНЖ, гражданство).

Связь между двумя типами данных

Для характеристики связи между двумя типами данных (группы ВК и сайты, найденные через поиск Google) были проанализированы ссылки в 23 группах ВК на два новостных портала — «Спутник» и «Радио Свобода». Мы рассматривали все посты, содержащие упоминание этих новостных порталов за 2019 и 2020 гг. (Если за указанный временной отрезок ссылок не было, мы фиксировали ссылки за более ранний период.)

Для всех типов групп ссылки на новостные порталы «Спутник» преобладают над «Радио Свобода»: более 300 ссылок было обнаружено для первого и только 34 — для второго. При этом ссылки на «Спутник» и «Радио Свобода» в сравнении с общим объемом постов оказываются немногочисленными; их больше в крупных группах с большим количеством постов. Ссылки на новостные порталы обычно используются для «наполнения» содержания групп новостями, и обсуждение новостей проходит в связи с динамикой, характерной для самих групп. Таким образом, связь между двумя типами собранных данных оказывается довольно слабой.

Для кыргызстанских, узбекистанских и таджикистанских групп характерно использование новостного контента без предоставления гиперссылок, а лишь с упоминанием источника (обычно с сайта «Спутник» о своей стране). В группах разного типа выкладываются новости о происшествиях в стране, о действиях своего правительства, о национальной культуре, о миграции в Россию. В таджикистанских группах ссылок на «Радио Свобода» сравнительно больше, причем часть из них связана с обсуждением ограничения его деятельности в Таджикистане.

В казахстанских и азербайджанских группах ссылки на новостные порталы появляются в основном в крупных «информационных группах», причем упоминание источников без гиперссылок также встречается чаще. В казахстанской группе

обсуждают новости о своей стране, в азербайджанской — новости Азербайджана и соседей по СНГ; в связи с этим там присутствуют отсылки к другим доменам: российским, казахстанским, туркменистанским, кыргызстанским.

Наконец, в группах юридической помощи мигрантам большинство ссылок приходится на одну, самую крупную группу: это новости о миграционной документации и о миграции вообще. В отличие от страновых групп, здесь большинство новостей сопровождается гиперссылками, причем на различные национальные домены: таджикистанский, азербайджанский, узбекистанский, кыргызстанский, абхазский, осетинский, эстонский.

Мы видим, что для страновых (национальных) групп характерно использование содержания новостей без гиперссылки и с полным или частичным «замыканием» на своем домене. Для стран с массовой трудовой миграцией в Россию (Кыргызстан, Узбекистан, Таджикистан) это новости о происходящем в своей стране и о миграции; для Казахстана и Азербайджана — новости о своей стране и, во втором случае, о других странах постсоветского региона. В группах юридической помощи мигрантам преобладают новости с гиперссылками на различные национальные домены, характеризующие миграционные процессы из стран СНГ в Россию. При этом численный перевес «Спутника» над «Радио Свобода», по-видимому, свидетельствует о концентрации на процессах в своей стране и в СНГ в сравнении с интересом к глобально-правозащитной тематике.

Выводы

Начнем обсуждение с того, что суммируем полученные результаты, а также процедуры сбора и анализа данных для трех этапов реализации исследовательского проекта (см. табл. 1).

Таблица 1. Три этапа реализации проекта: методика и результаты

Тип данных	Метод анализа	Исследовательские вопросы	Основные результаты	Сбор и анализ данных: разделение труда между человеком и алгоритмом
Google-сайты	Контент-анализ	Какие источники информации по миграционной тематике оказываются общими для национальных сегментов интернета?	В целом на разных языках находятся разные сайты. Общие новостные порталы — «Спутник» и «Радио Свобода», также распространена отсылка к Средней Азии. Сайты по миграционной тематике связывают Россию с Кыргызстаном, Таджикистаном и Узбекистаном. Сайты часто представляют версии на разных языках.	Сбор данных. Алгоритм-поисковик осуществляет поиск сайтов, персонализированный под рядовых пользователей и их девайсы. Алгоритм-переводчик переводит содержание сайтов. Человек (исследовательский коллектив) 1) определяет ключевые слова для поиска, 2) ограничивает число просматриваемых сайтов, 3) решает, какие сайты включать в выборку. Анализ данных. Контент-анализ проводится исследовательским коллективом.

Тип данных	Метод анализа	Исследовательские вопросы	Основные результаты	Сбор и анализ данных: разделение труда между человеком и алгоритмом
Группы ВК	Сетевой анализ	Какие группы оказываются связанными между собой через общих участников?	Группы кластеризуются по морфемам, выбранным для поиска, причем регионы/этнические категории, в основном группируются по «своим» странам (исключение составляют уйгуры и дунганде). Кластеризация внутри подвыборок по выбранным морфемам оказывается гораздо меньшей, чем для кластеров в полном графе.	Сбор данных. Алгоритм выдачи групп ВК выдает фиксированное число групп по ключевым словам. Человек (исследовательский коллектив) 1) определяет морфемы для поиска, 2) решает, какие группы включать в выборку. Алгоритм-переводчик переводит содержание. Человек создает (адаптирует) алгоритм, чтобы собирать информацию об участниках групп. Анализ данных: человек создает (адаптирует) алгоритмы для построения и визуализации графов. Человек определяет, для каких выборок строить графы, и интерпретирует полученные данные.
Группы ВК	Онлайн-наблюдение	В каких категориях участники групп осмысляют опыт пребывания в России? Как они идентифицируют себя и как их идентифицируют другие? Какие категории объединяют мигрантов из разных стран?	Важные категории: «низкоквалифицированные трудовые мигранты из Средней Азии» (обычно применяется к мигрантам из Кыргызстана, Таджикистана и Узбекистана) и «члены диаспоры» (для казахстанских и азербайджанских групп). Группы часто представляют собой смешанную языковую среду.	Сбор данных. Человек (исследовательский коллектив) определяет 1) выборку групп (на основании сетевого анализа), 2) категории наблюдения, 3) границы для отбора наблюдений: временной период и количество сообщений. Алгоритм (специально созданный для подобных исследований, используемый исследовательским коллективом) отбирает сообщения по заданным критериям. Человек проводит наблюдение. Алгоритм-переводчик переводит содержание. Анализ данных: исследовательский коллектив интерпретирует результаты онлайн-наблюдения.

Мы видим, что три этапа реализации проекта представляют, по сути, отдельные исследования с разными исследовательскими вопросами и, что не менее важно,— с особым разделением труда между людьми — членами исследовательского коллектива и разного рода алгоритмами, задействованными при сборе или анализе данных.

Транснационализм онлайн

Исследовательские вопросы на всех этапах реализации проекта были связаны с поиском общего транснационального пространства, которое бы объединяло Россию и страны исхода. По результатам исследования не было выделено единое транснациональное пространство. Скорее, различные элементы транснациональных практик мигрантов находят свое отражение в различных языковых и содержательных сегментах интернета. Отобранные через Google сайты не создают единого пространства: информация о миграции в основном оказывается замкнутой для шести языков. Группы ВК, по результатам сетевого анализа, также оказываются замкнуты внутри страновых кластеров: общих участников обычно есть у групп, имеющих ссылку к одной стране. Результаты онлайн-наблюдения, однако, свидетельствуют о том, что транснациональные пространства иногда могут возникать, и почти всегда они связывают две страны — Россию и страну исхода, — не включая другие страны.

Несмотря на «замыкание» внутри национальных сегментов, для каждого типа сайтов могут быть выделены некоторые элементы транснациональных пространств, связывающие отдельные языки, страны и сегменты интернета. По результатам контент-анализа и онлайн-наблюдения, наиболее перспективной для возникновения транснационализма в анализируемых сетевых пространствах является идентичность «Средняя/Центральная Азия», а также концентрация на повседневных проблемах трудовых мигрантов в России⁵.

Кроме того, для обоих типов данных было зафиксировано различие между тематизацией миграции из Азербайджана и Казахстана, с одной стороны, и Киргизстана, Таджикистана и Узбекистана, с другой стороны. В группах ВК оно относилось к типам групп и идентификациям мигрантов, на сайтах, найденных через Google-поиск, — к обнаруженным типам сайтов. Это деление соответствует объему и характеру миграции: Киргизстан, Таджикистан и Узбекистан являются крупнейшими странами-донорами для трудовой миграции в Россию; миграция из Казахстана и из Азербайджана не столь значительна и более разнообразна по своим целям⁶. Что не очевидно из статистики, но следует из нашего анализа онлайн-наблюдений: для Азербайджана и Казахстана важна не столько миграция, сколько диаспоральность. При этом и казахи, и азербайджанцы отделяют себя от «гастарбайтеров» из Киргизстана, Таджикистана и Узбекистана.

Следует также отметить, что различия в миграционной политике РФ в отношении граждан разных стран постоянно актуализируются в онлайн-ресурсах. Особенно явно это проявляется в отношении мигрантов из Киргизстана. С одной стороны, Киргизстан — страна массовой трудовой миграции в Россию, с другой стороны, это член ЕАЭС, а значит для его граждан действует облегченный режим трудоустройства в России. По результатам Google-поиска, киргизстанское информационное пространство оказывается лучше интегрировано с российским и содержит

⁵ В пяти странах исхода, включенных в анализ, большинство населения исповедует ислам. Исламская тематика также присутствует на сайтах обоих типов, однако там она второстепенна. Особенно явно это проявляется в сравнении со значением, которое приобретает региональная идентичность «Средняя Азия».

⁶ См. данные МВД РФ за 2019 г. URL: <https://xn--b1aew.xn--p1ai/Deljatelnost/statistics/migracionnaya/item/19365693/> (дата обращения: 10.02.2021).

больше сайтов по миграционной тематике, а по результатам онлайн-наблюдения кыргызстанские группы представляют транснационализм в наиболее чистом виде. Таким образом, результаты анализа обоих типов данных указывают на то, что в России наиболее транснациональными оказываются мигранты из Кыргызстана.

Наконец, отличия в практиках транснационализма связаны не только с содержанием сайтов, но и с формальными характеристиками онлайн-среды. Наиболее очевидный пример — отличие в том, как проявляется транснационализм в языковых практиках. В группах ВК часто формируется смешанная языковая среда: в постах, в комментариях, даже в названии и описании группы одновременно используются два или более языков. В то же время для сайтов новостей, государственных организаций, правозащитных организаций и т. п. характерно включение отдельных версий на разных языках. Таким образом, такие сайты ориентированы на модель переключения между национальными языками и национальными государствами, даже если апеллируют к иным идентичностям (региональной — Средняя Азия, религиозной — ислам и др.). В отличие от них группы в социальных сетях предоставляют больше возможностей для спонтанной активности и взаимодействия между пользователями и в этом отношении в большей степени отражают реальные процессы транснациональной миграции.

Разделение труда в онлайн-исследовании

Из таблицы 1 видно, что три этапа исследования различаются по тому, в какой мере исследовательский процесс предполагает включение разного рода алгоритмов, и по тому, какие это алгоритмы.

Сетевой анализ оказывается связан с работой алгоритмов и при сборе, и при анализе данных, тогда как для контент-анализа и онлайн-наблюдения алгоритмы действуют только при сборе данных. Это различие соотносится с привычным для социальных наук разделением на качественные и количественные данные: для анализа последних необходимо использовать количественные показатели (и визуализацию), что и осуществляют алгоритмы.

Задействованные в исследовании алгоритмы могут быть разделены на три типа.

1) Алгоритмы, которые являются составной частью онлайн-среды, — в нашем случае это поисковик Google и поиск групп ВК. Подобные алгоритмы создаются не для исследователя, но для пользователей конкретных сайтов. Их свойства и практики их использования — предмет особого внимания со стороны исследователей: чтобы получить осмысленные результаты, нужно знать, как пользоваться этими алгоритмами (как формулировать запрос, как ограничивать поиск, и т. п.). Например, в нашем исследовании выдача Google была ограничена первыми десятью страницами (100 сайтами), потому что повседневное знание о поисковике свидетельствует о том, что дальше первых десяти страниц пользователи заходят крайне редко⁷.

⁷ Иное решение по использованию поисковиков в исследовании миграционных процессов на постсоветском пространстве предлагается в работе Д. О. Тимошкина [Тимошкин, 2020]. Автор сравнивает результаты для трех поисковиков в два временных периода, однако с меньшей глубиной поиска и используя только русский язык. Различия в выборе поисковых алгоритмов и в процедурах поиска в значительной степени объясняются разницей в постановке исследовательского вопроса: исследователя интересует не транснационализм мигрантов, а образ мигранта в интернете.

2) Алгоритмы, созданные для реализации исследовательского процесса онлайн. Если первый тип конституирует объект исследования, то второй — это составная часть метода. Иными словами, возникает ситуация, когда алгоритмы взаимодействуют с алгоритмами, извлекая некоторую информацию из сети. Например, специальная программа извлекает информацию об участниках групп ВК, и алгоритмы ВК позволяют ей это делать (в отличие, например, от Facebook, где та же задача вызывает гораздо больше трудностей). Алгоритмы данного типа существуют в континууме: от созданных под реализацию конкретного проекта до типовых инструментов анализа; между ними располагаются типовые алгоритмы, адаптированные под конкретные нужды. Соответственно, задачи исследователя здесь варьируются от создания принципиально новых программ или адаптации старых до выбора и детального изучения функций уже имеющегося программного обеспечения. Кроме того, исследователю необходимо знать ту онлайн-среду, с которой эти программы должны взаимодействовать (то есть алгоритмы первого типа).

3) Универсальный инструмент — переводчик Google. В отличие от программ второго типа, универсальные алгоритмы направлены на решение задач, которые возникают в самых разных, не обязательно исследовательских контекстах⁸. Здесь от исследователя требуется понимание возможностей и ограничений существующих алгоритмов и выбор подходящего (например, понимание, когда использовать переводчик Google, а когда — Яндекса).

Существование всех трех типов алгоритмов связано с развитием онлайн-среды. Более того, использование таких повсеместно распространенных, но культурно специфичных технологий, как онлайн-переводчики, свидетельствует о том, что мы имеем дело не просто с онлайн-средой, но с онлайн-культурой в разнообразии ее проявлений. Поэтому невозможно написать учебник (handbook) по онлайн-исследованиям, который будет актуален более чем три-пять лет: меняются не только технологии — меняются практики их использования и степень распространенности.

Как следствие, в социальных науках возникает необходимость коллективной исследовательской работы, в которой будут задействованы специалисты с разным образованием и/или профессиональными умениями. Вслед за развитием онлайн-культуры и дифференциацией методов ее анализа происходит дифференциация профессиональных ролей. *Разделение труда между алгоритмами требует разделения труда между людьми, взаимодействующими с алгоритмами.*

Наконец, необходимо сказать несколько слов собственно об алгоритмах искусственного интеллекта. В настоящем исследовании под эту рубрику подпадают поисковик Google и переводчик Google. Алгоритмы ИИ встречаются скорее среди составляющих онлайн-среды, чем среди методического инструментария социальных наук. Работа по адаптации методов ИИ для социальной аналитики ведется, однако она требует значительных ресурсов, активного сотрудничества с техническими специалистами, а также постановки нетривиальных исследовательских задач. Следует также отметить: то, что сегодня не является ИИ, завтра может им стать (так, можно себе представить «умную» выдачу групп ВК, которая

⁸ Другие примеры универсальных алгоритмов-помощников — навигатор, профессиональный калькулятор, голосовой помощник. Разумеется, универсальные алгоритмы сами могут стать предметом исследования, и тогда их следует классифицировать как алгоритмы первого типа.

сортирует группы персонализированно). По сравнению с обычными алгоритмами для использования алгоритмов ИИ требуется более детальное знание об их возможностях и ограничениях⁹.

Общий вывод из представленной методологической рефлексии состоит в том, что *нейтральной онлайн-среды как объекта анализа не существует*. Результаты анализа зависят от того, какой способ организации исследования выбран. Задача исследователя (чаще — исследовательского коллектива) состоит в том, чтобы осмысленно и грамотно конструировать процедуры сбора, обработки и анализа онлайн-данных, ориентируясь на то, какие именно практики важны для тех процессов, по поводу которых был поставлен искомый исследовательский вопрос.

В подкрепление данного тезиса представим несколько вопросов о том, что было бы, если бы наше исследование было организовано по-другому:

- Как бы поменялась кластеризация групп, если бы в выборку включались все группы ВК, найденные по соответствующим морфемам, без отбора релевантных групп вручную?
- Как бы изменилась выборка сайтов, если бы для поиска были использованы «чистые» Google-машины (устройства без истории поиска)?
- Как бы изменились результаты исследования, если бы было выбрано другое «поле»: поисковик Яндекс, а не Google, социальные сети «Одноклассники» или Facebook, а не ВК?
- Как бы изменились результаты исследования, если бы вместо переводчика Google мы привлекали носителей языка?
- Как могло бы выглядеть онлайн-исследование в мессенджерах, с их растущей популярностью среди трудовых мигрантов в России?

Каждый из вопросов определяет перспективы для одного или нескольких новых исследований, результаты которых, вероятно, частично бы совпадали, а частично — отличались бы от результатов настоящего проекта. Это означает, что исследования транснационализма онлайн в России только начинаются.

Список литературы (References)

Digital Humanities: гуманитарные науки в цифровую эпоху / под ред. Г. В. Можаевой. Томск: Издательство Томского университета, 2016.

Digital Humanities: Humanities in the Digital Age. (2016) Tomsk: Tomsk State University Press. (In Russ.)

Абашин С. Н. Интеграция vs транснационализм: миграционные стратегии жителей Центральной Азии // Пути России. Война и мир. Т. XXII. / под ред. Пугачевой М. Г., Жаркова В. П. М.; СПб.: Нестор-История, 2017. С. 203—220.

Abashin S. N. (2017). Integration vs Transnationalism: Migration Strategies of Residents of Central Asia. In: Pugacheva M. G., Zharkov V. P. (eds.) *Ways of Russia. War and Peace*, Vol. XXII. Moscow; St. Petersburg: Nestor-Istoriia. P. 203—220. (In Russ.)

⁹ Например, знание о том, как формулировать и изменять запросы, — это тонкое, трудно формализуемое знание, которым мы обладаем как опытные пользователи (о чем часто не задумываемся). Предлагаем уважаемому читателю осмыслить собственный опыт использования онлайн-поисковиков и подумать о том, как передать кому-то этот опыт.

Бредникова О. Е. (Не)возвращение: могут ли мигранты стать бывшими? // Этнографическое обозрение. 2017. № 3. С. 32—47. URL: <https://journal.iea.ras.ru/archive/2010s/2017/no3/032.htm> (дата обращения: 11.02.2021).

Brednikova, O. E. (2017) (Non)Return: Is It Possible to Become an Ex-Migrant? *Etnograficheskoe obozrenie*. No. 3. P. 32—47. URL: <https://journal.iea.ras.ru/archive/2010s/2017/no3/032.htm> (accessed: 11.02.2021). (In Russ.)

Лисицын П.П., Трегубова Н.Д., Орлова Н.А. В поисках транснационализма онлайн: обзор сайтов по миграционной тематике на шести языках // Телескоп: журнал социологических и маркетинговых исследований. № 1. 2021. С. 126—133. https://doi.org/10.51692/1994-3776_2021_1_126.

Lisitsyn P.P., Tregubova N.D., Orlova N.A. (2021) Searching for Transnationalism Online: Review of Web Sites on Migration Issues in Six Languages. *Telescope*. No. 1. P. 126-133. https://doi.org/10.51692/1994-3776_2021_1_126. (In Russ.)

Резаев А.В., Трегубова Н.Д. «Искусственный интеллект», «онлайн-культура», «искусственная социальность»: определение понятий // Мониторинг общественного мнения: экономические и социальные перемены. 2019. № 6. С. 35—47. <https://doi.org/10.14515/monitoring.2019.6.03>.

Rezaev A.V., Tregubova N.D. (2019) Artificial Intelligence, On-line Culture, Artificial Sociality: Definition of the Terms. *Monitoring of Public Opinion: Economic and Social Changes*. No. 6. P. 35—47. <https://doi.org/10.14515/monitoring.2019.6.03>. (In Russ.)

Рыков Ю., Нагорный О. Область интернет-исследований в социальных науках // Социологическое обозрение. 2017. Т. 16. № 3. С. 366—394. <https://doi.org/10.17323/1728-192X-2017-3-366-394>.

Rykov Y., Nagornyy O. (2017) Internet Studies in Social Sciences. *The Russian Sociological Review*. Vol. 16. No. 3. P. 366—394. <https://doi.org/10.17323/1728-192X-2017-3-366-394>. (in Russ.)

Сивков Д.Ю. Дигитальная социология: напряжения, неопределенности, нерешенные проблемы // Социология власти. 2018. Т. 30. № 3. С. 8—13. <https://doi.org/10.22394/2074-0492-2018-3-8-13>.

Sivkov D.Y. (2018) Digital Sociology: Tensions, Ambiguities, Unsolved Problems. *Sociology of Power*. Vol. 30. No. 3. P. 8—13. <https://doi.org/10.22394/2074-0492-2018-3-8-13>. (In Russ.)

Степанов А.М. Транснациональный подход в современных миграционных исследованиях // Вестник Санкт-Петербургского университета. Социология. 2018. Т. 11. № 1. С. 116—127. <https://doi.org/10.21638/11701/spbu12.2018.110>.

Stepanov A.M. (2018) Transnational Approach in Contemporary Migration Studies. *Vestnik of Saint Petersburg University. Sociology*. Vol. 11. No. 1. P. 116—127. <https://doi.org/10.21638/11701/spbu12.2018.110>. (In Russ.)

Тимошкин Д.О. Доверие vs дезориентация: экономика русскоязычных «мигрантских» групп в социальных сетях // Экономическая социология. 2019. Т. 20. № 5. С. 53—73. <https://doi.org/10.17323/1726-3247-2019-5-53-73>.

Timoshkin D. (2019) Trust vs. Disorientation: Economy of the Russian-speaking 'Migrant' Groups in Social Media. The Case of VKontakte. *Ekonomicheskaya sotsiologiya*. Vol. 20. No. 5. P. 53—73. <https://doi.org/10.17323/1726-3247-2019-5-53-73>. (In Russ.)

Тимошкин Д. О. Разнообразие «Других»: образ «московского мигранта» на странице результатов поиска Google, Яндекс и DuckDuckgo // Журнал исследований социальной политики. 2020. Т. 18. № 4. С. 641—656. <https://doi.org/10.17323/727-0634-2020-18-4-641-656>.

Timoshkin D.O. (2020) A Variety of 'Others': Images of 'Moscow Migrants' in the Search Results of Google, Yandex and DuckDuckgo. *The Journal of Social Policy Studies*. Vol. 18. No. 4. P.641—656. <https://doi.org/10.17323/727-0634-2020-18-4-641-656>. (In Russ.)

Трегубова Н. Д. Транснациональный мигрант в интернете: теоретические основания исследования транснационализма в режиме онлайн // Мониторинг общественного мнения: экономические и социальные перемены. 2020. № 2. С. 405—419. <https://doi.org/10.14515/monitoring.2020.2.1637>.

Tregubova N. D. (2020) Transnational Migrants in the Internet: Theoretical Foundations for Studying Transnationalism Online. *Monitoring of Public Opinion: Economic and Social Changes*. No. 2. P. 405—419. <https://doi.org/10.14515/monitoring.2020.2.1637>. (In Russ.)

Brubaker R. (2004) Ethnicity Without Groups. Cambridge, MA: Harvard University Press.

Collins R. (2004) Interaction Ritual Chains. Princeton, NJ: Princeton University Press.

DiMaggio P., Bernier C., Heckscher C., Mimno D. (2018) Interaction Ritual Threads: Does IRC Theory Apply Online? In: Weininger E. B., Lareau A., Lizardo O. (eds.) *Ritual, Emotion, Violence: Studies on the Micro-Sociology of Randall Collins*. Abingdon and New York: Routledge. P. 81—124. <https://doi.org/10.4324/9780429464157>.

Hjorth L., Horst H., Galloway A., Bell G. (eds.) (2017) The Routledge Companion to Digital Ethnography. New York, NY; London: Routledge.

Ruget V., Usmanlieva B. (2019) Can Smartphones Empower Labour Migrants? The Case of Kyrgyzstani Migrants in Russia. *Central Asian Survey*. Vol. 38. No. 2. P. 165—180. <https://doi.org/10.1080/02634937.2019.1594170>.

Starikov V. S., Ivanova A. A., Nee M. L. (2018) Transnationalism Online: Exploring Migration Processes With Large Data Sets. *Monitoring of Public Opinion: Economic and Social Changes*. No. 5. P. 213—232. <https://doi.org/10.14515/monitoring.2018.5.17>.

Tregubova N. D., Nee M. L. (2020). Beyond Nations and Nationalities: Discussing the Variety of Migrants' Identifications in Russian Social Media. *Changing Societies & Personalities*. Vol. 4. No. 3. P. 323—349. <https://doi.org/10.15826/csp.2020.4.3.104>.

Urinboyev R. (2017) Establishing an «Uzbek Mahalla» via Smartphones and Social Media: Everyday Transnational Lives of Uzbek Labour Migrants in Russia. In: Laruelle M. (ed.) *Constructing the Uzbek State: Narratives of Post-Soviet Years*. Boulder, CO: Lexington Books. P. 119—149.

DOI: [10.14515/monitoring.2021.1.1760](https://doi.org/10.14515/monitoring.2021.1.1760)



М. Б. Богданов, И. Б. Смирнов

ВОЗМОЖНОСТИ И ОГРАНИЧЕНИЯ ЦИФРОВЫХ СЛЕДОВ И МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ В СОЦИОЛОГИИ

Правильная ссылка на статью:

Богданов М. Б., Смирнов И. Б. Возможности и ограничения цифровых следов и методов машинного обучения в социологии // Мониторинг общественного мнения: экономические и социальные перемены. 2021. № 1. С. 304—328. <https://doi.org/10.14515/monitoring.2021.1.1760>.

For citation:

Bogdanov M. B., Smirnov I. B. (2021) Opportunities and Limitations of Digital Footprints and Machine Learning Methods in Sociology. *Monitoring of Public Opinion: Economic and Social Changes*. No. 1. P. 304–328. <https://doi.org/10.14515/monitoring.2021.1.1760>. (In Russ.)

ВОЗМОЖНОСТИ И ОГРАНИЧЕНИЯ ЦИФРОВЫХ СЛЕДОВ И МЕТОДОВ МАШИННОГО ОБУЧЕНИЯ В СОЦИОЛОГИИ

БОГДАНОВ Михаил Богданович — младший научный сотрудник Центра социологии культуры, Институт образования, аспирант департамента социологии, Национальный исследовательский университет «Высшая школа экономики», Москва, Россия
E-MAIL: bogdanovmikle@mail.ru
<https://orcid.org/0000-0001-6245-7178>

СМИРНОВ Иван Борисович — ведущий научный сотрудник, заведующий лабораторией вычислительных социальных наук, Институт образования, Национальный исследовательский университет «Высшая школа экономики», Москва, Россия
E-MAIL: ibsmirnov@hse.ru
<https://orcid.org/0000-0002-8347-6703>

Аннотация. В статье описываются возможности и ограничения использования в социологии новых источников данных и методов их сбора, обработки и анализа, а именно — цифровых следов и методов машинного обучения. Сначала обсуждаются недостатки классических источников данных — опросов, а затем, в контексте этих недостатков, на основе релевантных исследований анализируются возможности их преодоления с помощью цифровых следов. В качестве главных недостатков опросных данных, которым, в свою очередь, меньше подвержены цифровые следы, выделяются: реактивность, небольшой объем данных и редкая частотность. В контексте описания этих недостатков и способов их преодоления с помощью цифровых следов мы приводим типы исследова-

OPPORTUNITIES AND LIMITATIONS OF DIGITAL FOOTPRINTS AND MACHINE LEARNING METHODS IN SOCIOLOGY

Mikhail B. BOGDANOV¹ — Junior Research Fellow at the Centre for Cultural Sociology, Institute of Education; PhD student in Sociology
E-MAIL: bogdanovmikle@mail.ru
<https://orcid.org/0000-0001-6245-7178>

Ivan B. SMIRNOV¹ — Leading Research Fellow, Head of the Computational Social Science Lab, Institute of Education
E-MAIL: ibsmirnov@hse.ru
<https://orcid.org/0000-0002-8347-6703>

¹ National Research University Higher School of Economics, Moscow, Russia

Abstract. The article discusses the opportunities and limitations of using new data sources and methods of its collection, processing and analysis, namely, digital traces and machine learning in Sociology. At first, we examine the disadvantages of traditional data sources (surveys) and then, based on relevant and recent empirical studies, we discuss how these disadvantages can be overcome using digital traces. The main drawbacks of survey data are the reactivity, a small sample size, and rare frequency of surveys. Based on these drawbacks we identify types of research questions that can only be answered with digital traces. Finally, we also explore the disadvantages of digital traces: lack of representativeness, construct validity, external and internal interfering factors, and non-stationarity. Relying on recent

тельских вопросов, на которые можно ответить только с помощью цифровых следов. После этого рассматриваем ограничения цифровых следов: непрепрезентативность, конструктную валидность, внешние и внутренние вмешивающиеся факторы, нестационарность. Затем, на основе актуальных методологических статей, мы описываем, как учитывать эти ограничения и по возможности корректировать их.

Ключевые слова: цифровые следы, большие данные, машинное обучение, предсказательное моделирование, вычислительные социальные науки, вычислительная социология, анализ данных, анализ текстов

Благодарность. Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 20-311-90056.

methodological developments the paper explains how to take into consideration these limitations and how to adjust for them wherever possible..

Keywords: digital footprints, big data, machine learning, forecasting modeling, computational social sciences, computational sociology, data analysis, text analysis

Acknowledgments. The study was funded by the Russian Foundation for Basic Research (RFBR), project No. 20-311-90056.

Введение

Развитие цифровых технологий и широкое распространение интернета привели к стремительным изменениям в социальных науках, активному использованию новых источников данных, а также появлению новых методов их сбора, обработки и анализа. Возникшее таким образом направление исследований получило название «вычислительные социальные науки» (Computational Social Science) [Lazer et al., 2009]. Вычислительные социальные науки опираются на цифровые следы, то есть данные, получаемые из социальных сетей [Garcia et al., 2018; Garcia, Rimé, 2019; Jaidka et al., 2020; Sivak, Smirnov, 2019], поисковых систем [Bail, Brown, Wimmer, 2019; Stephens-Davidowitz, 2014], логи звонков [Blumenstock, Cadamuro, On, 2015] и других веб-сайтов [Abrahao et al., 2017; Lewis, 2013], показаний датчиков GPS [Stopczynski et al., 2014], оцифрованных текстов [Hills et al., 2019] и масштабных административных данных [Pierson et al., 2020]. Эти новые источники данных зачастую характеризуются большими объемами, высокой разрешающей способностью и слабой структурированностью, а для их анализа обычно используются алгоритмы машинного обучения и другие вычислительные методы. Исторически этим направлением занимались не социологи, большинство публикаций в этой сфере выходили в несоциологических журналах, и только в последние несколько лет это стало меняться [Edelmann et al., 2020; Golder, Macy, 2014; Lazer, Radford, 2017; Molina, Garip, 2019], поэтому для многих социологов

в России остается непонятным, какие именно новые возможности открывают перед ними цифровые следы и вычислительные методы, а также какие трудности могут встретиться при работе с такого рода данными.

Задача этой статьи — ответить, опираясь на последние исследования в области вычислительных социальных наук, на следующие вопросы, которые современные социологи могут задавать сами себе: в каких случаях могут быть полезны цифровые следы? На какие новые исследовательские вопросы они позволяют ответить? Если я захочу использовать такие данные, то что нужно учитывать при планировании своего исследования?

В этом обзоре сначала кратко описываются ограничения таких классических источников данных для количественной социологии, как массовые опросы. Затем на примерах современных исследований описывается, как новые данные и подходы к их обработке и анализу могут помочь преодолеть эти ограничения. После этого мы описываем и обсуждаем типы исследовательских вопросов в количественной социологии, на которые нельзя было ответить с помощью опросов, но можно ответить с помощью новых данных. В конце приводятся и анализируются ограничения новых данных, а также даются некоторые общие рекомендации по сбору, обработке и анализу такого рода данных.

Ограничения опросных данных и возможности цифровых следов по их преодолению

Одним из главных ограничений классических массовых опросов является их реактивность — респонденты и испытуемые практически всегда знают, что участвуют в исследовании, а это, в свою очередь, может оказать эффект на результаты исследования [Lavrakas, 2008: 694]. Проблема в том, что оценить величину этого эффекта невозможно — мы не можем знать, как бы человек ответил на вопросы анкеты или как бы вел себя во время экспериментального воздействия, не дав ему или ей заполнить анкету или не поместив в экспериментальные условия. Цифровые же следы по своей природе *нереактивны* [Salganik, 2019: 23] или *малореактивны* [Девятко, 2018: 30], то есть не производятся специально в целях исследований, а представляют собой побочный продукт фиксации поведения людей в цифровых системах.

Проблема реактивности особенно актуальна в исследованиях сенситивных тем, так как реактивные методы, такие как массовые опросы, могут давать смещенные результаты из-за социально одобряемого поведения респондентов. В случае использования цифровых следов эта проблема становится менее выраженной. Например, данные о поисковых запросах в Google позволили проанализировать уровень расизма в разных штатах и административных единицах Америки. Затем эти данные использовались для предсказания результатов президентских выборов в США в 2008 и 2012 гг. Оказалось, что связь между уровнем расизма и результатами Барака Обамы на данных поисковых запросов была в 1,5—3 раза выше, чем по данным опросов [Stephens-Davidowitz, 2014].

С помощью этих агрегированных данных впоследствии была исследована взаимосвязь между расизмом и уровнем смертности среди темнокожего населения США [Chae et al., 2015], а также взаимосвязь между расизмом и неблагопри-

ятными исходами родов среди темнокожих женщин — преждевременные роды и маленький вес ребенка при рождении [Chae et al., 2018].

На агрегированных по административным округам США данных поисковых запросов в Google была проанализирована причинно-следственная связь между антимусульманскими и протеррористическими настроениями в локальных сообществах [Bail, Merhout, Ding, 2018]. Исследование взаимосвязи этих феноменов с помощью опроса столкнулось бы с проблемами с обеих сторон: экстремисты не стали бы заявлять о своих радикальных взглядах в опросе, а люди с антимусульманскими установками могли бы преуменьшать их в своих ответах или не выражать совсем из-за социальной нежелательности таких взглядов [ibid.: 1].

В другом исследовании похожей тематики расовые предрассудки полицейских в США были исследованы на данных о почти 100 млн проверок машин дорожной полицией [Pierson et al., 2020].

Еще один пример связан с исследованием расовых предрассудков при выборе романтических партнеров на данных о поведении 126 тысяч пользователей онлайн-сервиса для знакомств «OkCupid» за 2,5 месяца [Lewis, 2013]. В этой работе анализировалась вероятность проявления и получения романтического интереса между двумя людьми в зависимости от их расовых и других социально-демографических характеристик. Подобное исследование было бы крайне сложно провести с помощью таких реактивных методов, как опрос или эксперимент: они позволили бы измерить намерения и интерес, а не реальное поведение в естественных условиях [ibid.: 18814].

Цифровые следы также могут служить валидацией результатов, полученных с помощью реактивных методов. Так, на данных сервиса краткосрочной сдачи и аренды жилья Airbnb было проведено исследование социальных предрассудков, в котором результаты онлайн-эксперимента валидировались с помощью анализа реальных поведенческих данных на сервисе [Abraha et al., 2017].

Цифровые следы использовались и для изучения влияния социальной политики в области здравоохранения на уровень абортов [Reis, Brownstein, 2010], а также для изучения потребления наркотиков [Enghoff, Aldridge, 2019].

На данных 635 000 пользователей социальной сети «ВКонтакте» из Санкт-Петербурга было проанализировано гендерное неравенство в упоминании дочерей и сыновей родителями в социальных сетях [Sivak, Smirnov, 2019]. Оказалось, что родители, как матери, так и отцы, чаще упоминают в постах сыновей, чем дочерей. Более того, посты с упоминанием сыновей в среднем набирают на 50 % большей лайков [ibid.: 2040]. Авторы отмечают, что упоминание детей в постах представляет собой простую и напрямую измеримую поведенческую метрику гендерного неравенства, которую было бы сложно измерить с помощью опроса из-за социальной желательности таких вопросов [ibid.: 2040].

Другое существенное ограничение традиционных подходов — относительно небольшой по сравнению с цифровыми следами объем изучаемых выборок. Теоретически можно опросить всех людей, но на практике это сделать невозможно по множеству причин, среди которых ключевая — значительные затраты ресурсов, как финансовых, так и временных. Поэтому ученые ограничиваются выборочными опросами, размер выборки которых редко превышает несколько

тысяч, иногда десятков тысяч, человек и не сравним с потенциальным размером данных цифровых следов.

Неоднородные зависимости

Сравнительно небольшой размер выборок в опросах ограничивает поиск неоднородных зависимостей и паттернов между изучаемыми концептами. Например, было показано, что использование социальных медиа не связано с удовлетворенностью жизнью у подростков [Orben, Dienlin, Przybylski, 2019]. Однако может оказаться, что связь существует, но она либо нелинейна, либо гетерогенна для разных групп подростков, либо нелинейна и гетерогенна одновременно. Возможность разбивать выборку на подгруппы по разным критериям и использовать модели, замеряющие нелинейные и гетерогенные по своей природе связи принципиально ограничена доступным размером выборки [Molina, Garip, 2019: 37]. Это ограничение может преодолеваться за счет использования как больших данных обособленно, так и вкупе с опросами.

Даже если какая-то информация не доступна в явном виде для большого количества людей, ее часто можно оценивать, используя модели машинного обучения на объединенных данных цифровых следов и анкет респондентов. Таким образом можно обогащать данные за счет предсказания значения переменных, измеренных с помощью опроса для тех людей, которые не участвовали в опросе, но чьи цифровые следы можно собрать. Такой подход некоторые исследователи называют усиленным опросом (от англ. *amplified asking*) [Salganik, 2019: 122]. К примеру, многие исследования показали, что черты личности могут быть предсказаны по различным цифровым следам [Azucar, Marengo, Settann, 2018; Huang, 2019; Kosinski, Stillwell, Graepel, 2013; Settanni, Azucar, Marengo, 2018]. Также было показано, что образовательные результаты тоже могут быть предсказаны с помощью цифровых следов [Smirnov, 2020, 2018]. Другими словами, большие объемы данных позволяют анализировать данные более детализированно.

Кроме того, данные могут обогащаться не только на индивидуальном, но и на контекстуальном уровне, то есть на уровне групп, районов, регионов и т. д. Для социологии это представляет особый интерес, так как позволяет получить на агрегированном уровне информацию по тем переменным, которые не измеряются государственной статистикой. Так, на данных о дружеских связях в Facebook был создан индекс социальной связанности административно-территориальных регионов США, который измеряет относительное количество дружеских связей на Facebook между жителями разных административных округов США [Bailey et al., 2018].

С помощью агрегированных данных сервиса Facebook по созданию кампаний таргетированной рекламы было проанализировано гендерное неравенство в 217 странах [Garcia et al., 2018]. В похожем исследовании показано, что отношения количества женщин к количеству мужчин среди пользователей Facebook и Google в разных странах выступают сильными предикторами цифрового неравенства в стране [Kashyap et al., 2020].

На данных 1,53 млрд постов в Twitter было проанализировано географическое распределение субъективного благополучия в 1208 округах США [Jaidka et al., 2020]. При помощи методов обработки естественного языка и машинного обуче-

ния исследователи оценили на данных Twitter показатели субъективного благополучия, агрегированные на уровне округов. Наилучшая оценка коррелировала с данными опроса 1,73 млн респондентов об удовлетворенности жизнью, счастье, беспокойстве и печали на уровне 0,51—0,64 [ibid.: 3]. Более того, эта модель показала свою устойчивость во времени: анализ подвыборок за 2012—2013 и 2015—2016 гг. продемонстрировал схожие результаты [ibidem]. Относительная стационарность во времени позволяет в будущем использовать эту модель для получения агрегированных данных о благополучии на уровне административных округов, а также о динамике этого феномена.

Другой хрестоматийный пример использования новых данных — это исследование географической гетерогенности межпоколенческой мобильности в США [Chetty et al., 2014]. В этом исследовании на индивидуальных данных налоговой статистики о 40 млн американцев была выявлена существенная гетерогенность в межпоколенческой мобильности по административным округам США. Если для Америки в целом вероятность попасть в верхний квинтиль по распределению доходов для ребенка 1980—1985 годов рождения из семьи из нижнего квинтиля равна 7,8 %, то для разных округов эта вероятность может изменяться в несколько раз: от 4 % до 13 % [ibid.: 1596].

Также на данных административной статистики о смертности в штате Калифорния и данных 12 млн профилей на Facebook было выявлено, что социальная интеграция в сети, измеренная через социальные взаимодействия на Facebook (добавление в друзья, размещение постов и фотографий, отправка сообщений), связана со смертностью [Hobbs et al., 2016]. Более того, большой объем данных позволил проследить гетерогенность взаимосвязи для разных причин смертности и оказалось, что для рака такой связи нет, а для суицида и передозировки наркотиками есть [ibid.: 12983].

Предсказательное моделирование

Кроме того, наличие существенно больших по размеру данных позволяет ставить исследовательские вопросы другого характера. Речь идет прежде всего о задачах и вопросах предсказания.

Статистическое моделирование можно разделить на два типа: объяснительное (от англ. *explanatory*) и предсказательное (от англ. — *predictive*) [Shmueli, 2010: 290—291]. Если объяснительное моделирование — это применение статистических моделей для проверки гипотез о корреляциях и причинно-следственных связях между теоретическими конструктами [ibid.: 291], то предсказательное моделирование — это применение статистических моделей для предсказания новых (вневыборочных) наблюдений [ibid.: 292].

В социальных науках доминируют объяснительные модели, а вопросам предсказания уделяется существенно меньше внимания [Cranmer, Desmarais, 2017; Hofman, Sharma, Watts, 2017; Shmueli, 2010]. Вероятно, одна из причин заключается в том, что для использования более мощных с точки зрения предсказательной силы моделей машинного обучения (нейронные сети, бустинги и т. п.) необходимы существенно большие выборки [Yarkoni, Westfall, 2017], чем те, которые обычно собираются с помощью опросов.

Например, недавний конкурс по предсказательному моделированию на данных когортного лонгитюдного исследования «Fragile Families» показал, что на опросных данных, пусть даже очень хорошего качества, возможности предсказания существенно ограничены [Salganik et al., 2020]. Это исследование примечательно тем, что в нем сотни исследовательских команд со всего мира пытались предсказать с помощью имеющихся данных и самых разнообразных моделей машинного обучения различные образовательные и жизненные показатели 15-летних школьников. R-квадрат лучшей предсказательной модели среднего балла по школьным предметам составил всего 0,19, то есть лучшая модель смогла предсказать всего 19 % вариации зависимой переменной [Salganik и др., 2020: 8340].

Некоторые исследователи связывают такой невысокий результат с тем, что возможности классических, как по размеру, так и по качеству, опросных данных для предсказательного моделирования ограничены и следует апробировать этот подход на больших по размеру данных [Garip, 2020: 2]. Кроме того, вероятно, что те теоретические концепты, которые операционализируются в анкетные вопросы, недостаточно хорошо схватывают реальность [*ibidem*]. В этом контексте использование цифровых следов может дать заметно больше информации о тех нишах и тех аспектах реальности, которые не удается измерить с помощью анкетных вопросов.

Например, на данных 852 млн твитов 51,6 млн пользователей исследователям удалось предсказать количество ретвитов (репостов), которое наберет конкретный твит в зависимости от характеристик автора и текста этого твита [Martin et al., 2016]. Лучшая модель предсказывала почти половину дисперсии зависимой переменной — R-квадрат = 0,48.

Большие выборки необходимы еще и потому, что для корректного измерения предсказательной силы моделей машинного обучения требуется разбиение выборки на подвыборки. Это связано с тем, что современные методы машинного обучения, такие как искусственные нейронные сети, используют модели, содержащие большое количество параметров, что может приводить к переобучению моделей [Yarkoni, Westfall, 2017: 1110—1111]. Чтобы избежать переобучения, необходимо разбивать исходную выборку как минимум на три части, тренировочный набор, валидационный набор и тестовый набор. Тренировочный набор используется для построения модели, предсказывающей зависимую переменную по информации о независимых. Валидационный набор используется для проверки, насколько точно эта модель предсказывает значения зависимой переменной на новых данных [Molina, Garip, 2019: 28—30; Yarkoni, Westfall, 2017: 1103—1104]. Исследователи обычно стремятся выбрать модель, которая максимизирует точность на валидационном наборе данных, но это в свою очередь может тоже привести к завышению оценки предсказательной точности модели. Поэтому для финальной оценки используется тестовый набор данных, который откладывается в самом начале исследования и используется только один раз для тестирования предсказательной силы финальной модели [Molina, Garip, 2019: 31—32; Yarkoni, Westfall, 2017: 1113].

Тем не менее в силу необходимости разделения выборки на обучающую и тестовую очевидно, что для задач исследовательских вопросов предсказательного

характера требуется больше данных, чем для классического статистического моделирования, в котором модели обучаются и оцениваются на одних и тех же данных.

Взаимодействия людей и социальные сети

С помощью выборочных опросов значительно сложнее исследовать взаимодействия людей и реконструировать масштабные социальные сети. Это связано с тем, что для точного воспроизведения структуры сети необходимы данные о каждом узле, в противном случае, если нет данных о ключевых узлах, например, соединяющих разные клики сети, вся структура реконструированной сети может быть некорректной [Kossinets, 2006].

Кроме того, люди плохо помнят, с кем они взаимодействовали, и могут по-разному воспринимать такие вопросы [Lazer, Radford, 2017: 23]. Тогда как социальные сети, построенные на цифровых следах (лайки, репосты, комментарии и т. п.), могут охватывать значительно большие сети. Например, на данных о дружбе санкт-петербургских школьников из «ВКонтакте» была воссоздана сеть цифровой близости между школами Санкт-Петербурга и показано, что академические успехи школы сильно связаны с показателями ее цифровых школ-соседей, но совершенно не связаны с географическими [Smirnov, 2019: 5]. Очевидно, что невозможно было бы опросить всех или почти всех школьников города об их взаимодействиях друг с другом.

Цифровые репрезентации социальных сетей, построенных на данных популярных сетевых онлайн-платформ вроде Facebook, «ВКонтакте», Twitter и других, предоставляют гигантские возможности для тестирования классических социологических теорий и гипотез.

Например, теория Марка Грановеттера о силе слабых связей была апробирована на сетевых данных 6 млн американских пользователей Facebook [Gee, Jones, Burke, 2017]. В этой работе сравнивалось влияние сильных и слабых связей на трудоустройство. В качестве показателя связи выступала дружба на Facebook, а сила связей измерялась через количество совместных фотографий, постов друг у друга и общих друзей [ibid.: 493—496]. С помощью этих данных было показано, что, несмотря на важность слабых связей, наличие хотя бы одной сильной связи более ценно с точки зрения поиска работы и трудоустройства [ibid.: 485].

Впоследствии авторы использовали этот подход и сетевые данные Facebook о 17 млн дружеских связей пользователей из 55 стран и показали, что эта зависимость сохраняется во всех странах [Gee et al., 2017]. Более того, сила эффекта зависит от экономического неравенства в стране — чем оно выше, тем выше вероятность, что человек найдет следующую работу через сильную связь [ibid.: 370].

Также на данных о 1,5 млн обменов подарками на Facebook протестирована классическая теория Марселя Мосса о природе дарообмена [Mauss, 2000] и выявлено, что получение подарка увеличивает вероятность дарения в будущем на 56 % [Kizilcec et al., 2018].

Социальные сети влияют и на принятие решений на рынке недвижимости [Bailey et al., 2018]. На данных о дружбе в Facebook выявлено, что люди, чьи географически отдаленные друзья недавно испытали повышение цен на недвижимость в их районе, более склонны к переходу от съема жилья к покупке собственного, причем большего по размеру [Bailey et al., 2018: 2224].

Это исследование также интересно тем, что в нем используются опросные данные для валидации и уточнения результатов, полученных на данных цифровых следов. После анализа данных из Facebook исследователи пришли к выводу, что рост цен на недвижимость друзей, живущих в других районах, может воздействовать на экономическое поведение в сфере недвижимости через улучшение ожиданий относительно покупки недвижимости в районе индивида как способа финансовых инвестиций. Для валидации этого вывода в Facebook был проведен вспомогательный опрос, в котором в том числе оценивалась привлекательность покупки недвижимости в районе проживания респондента как способ финансового инвестирования, а также собирались данные о дружеских сетях респондентов на Facebook. Вспомогательный анализ на данных опроса подтвердил вывод: увеличение цен на недвижимость в районах проживания друзей связано с более оптимистичными установками относительно покупки недвижимости [Bailey, Cao, Kuchler, Stroebel, 2018: 2266—2269].

На данных фотографий и дружеских сетей из Facebook также было проанализировано распространение культурных трендов. С помощью нейронных сетей из изображений была извлечена информация о предметах, местах и категориях, изображенных на них, затем композиция характеристик изображений сравнивалась с фотографиями из профилей друзей, было выявлено, что друзья склонны выкладывать фотографии со схожими культурными трендами. Сетевые данные также позволили отделить эффект социального влияния от эффекта гомофилии¹ и показать, что культурные тренды склонны распространяться по социальным сетям за счет воздействия одних людей на других.

В онлайн-эксперименте на 61 млн пользователей Facebook было выявлено, что сообщения с политической мобилизацией оказывают воздействие не только на самих получателей, но также и на их друзей и друзей друзей [Bond et al., 2012]. Очевидно, что исследование третичных эффектов, то есть воздействия сообщений на друзей друзей получателя, было бы невозможно осуществить с помощью опросных данных из-за проблем с неответами, неполной и неточной информацией о социальных связях и т. п.

Сетевые данные Facebook также использовали для изучения доверия. На основе опроса 6 тыс. пользователей и данных о группах на Facebook, в которых они состоят, и их характеристик было проанализировано общее доверие и доверие по отношению к участникам группы, из которой их рекрутировали [Ma et al., 2019]. Оказалось, что позиция индивида в дружеской сети участников группы связана с уровнем его/ее общего доверия [ibid.: 8].

Труднодоступные и маленькие группы

С помощью опросов сложно исследовать труднодоступные, закрытые и маленькие группы. Если размер группы небольшой и нет конкретных локаций ее обитания (как территориальных, так и виртуальных), по которым можно таргетировать опрос, то количество контактов, необходимых для достижения выборки нужного размера, может быть очень большим и, как следствие, стоимость проведения

¹ Гомофилия — свойство социальной сети, при котором узлы с похожими атрибутами склонны формировать между собой связи.

такого опроса также может быть очень высокой. Кроме этого, члены закрытых групп могут быть не склонны участвовать в опросе. Например, вряд ли удастся исследовать с помощью опроса радикальных футбольных фанатов, состоящих в фанатских группировках и участвующих в организованных драках.

Цифровые же следы позволяют наблюдать за труднодоступными, закрытыми и маленькими группами через социальные медиа. Так, на данных 10 млн мигрантов в США и их сетях дружбы на Facebook был проанализирован уровень интеграции мигрантов из разных стран в американское общество [Herdağdelen et al., 2016]. В другом исследовании схожей тематики анализировалась культурная ассимиляция мигрантов из Мексики в США на данных сервиса Facebook для таргетированной рекламы о музыкальных предпочтениях той или иной аудитории [Stewart et al., 2019]. Размер и охват данных Facebook позволил проанализировать феномен культурной ассимиляции в разных социально-демографических группах [ibid.: 3259]. Кроме того, поскольку данные о музыкальных предпочтениях вносятся пользователями самостоятельно, это, в отличие от ограниченных анкетных вопросов, позволяет зафиксировать больший спектр музыкальных предпочтений и вкусов [ibidem].

Также цифровые следы можно использовать для таргетирования опросов на труднодоступные группы. Например, с помощью сервиса Facebook для создания таргетированной рекламы можно рекрутировать респондентов не только по их демографическим характеристикам, но и по поведению в сети, интересам и другим данным, собирающимся в интернете [Iannelli et al., 2018].

Редкая частотность

Еще одно существенное ограничение опросных данных — их относительно редкая частотность. Проведение опросов трудозатратно: нужно разработать анкету, сконструировать выборку и провести полевой этап сбора данных, длительность которого также зависит от метода сбора данных. Это, в свою очередь, ограничивает частоту проведения опросов и, следовательно, делает невозможным достаточно детализированное во времени изучение динамики некоторых феноменов. Опросные компании проводят еженедельные опросы (так называемые омнибусы), а иногда и ежедневные, но здесь возникают ограничения, связанные с объемами таких ежедневных выборок. При сравнительно небольших объемах выборочных показатели страдали бы от случайных флуктуаций, то есть от случайной ошибки выборки.

Кроме того, в долгосрочной перспективе на это может накладываться усталость людей от опросов, которая сейчас наблюдается в индустрии [Wojcik, Hughes, 2019]. А если говорить о лонгитюдном исследовании, то респонденты могут уставать от участия в опросе и отказываться от участия в исследовании. Если такое осыпание панели (от англ. panel attrition) происходит не случайным образом, то это может приводить к смещениям в выборке. Кроме того, сам опыт участия в лонгитюдном исследовании влияет на ответы респондентов (в англ. panel conditioning) [Warren, Halpern-Manners, 2012].

Преимущество цифровых следов заключается в том, что они всегда доступны и их можно собрать ретроспективно за весь интересующий промежуток времени. Таким образом, цифровые следы могут быть доступны в намного более частотном разрезе, чем опросы: в разрезе недель, дней и даже часов. Так, посты из Twitter

позволили проанализировать динамику коллективных эмоций (то, что Дюркгейм называл «effervescence») до и после террористических атак в Париже в 2015 г. в разрезе месяцев, недель и даже дней [Garcia, Rimé, 2019]. Очевидно, что такое исследование было бы практически невозможно осуществить с помощью опроса по следующим причинам.

Во-первых, для замера эмоций до и после ключевого события было бы необходимо, чтобы уже на протяжении какого-то времени проводился опрос, замеряющий релевантные показатели. Ретроспективный опрос о динамике и изменениях коллективных эмоций видится слабым инструментом в силу давно известной неточности ретроспективных оценок респондентов, особенно оценок таких эфемерных концептов, как установки, отношение и восприятие [Smith, 1984]. Во-вторых, даже если бы такие замеры проводились, они все-равно не могли бы быть настолько детализированными во времени по описанным выше причинам. В-третьих, с помощью опроса было бы сложнее адекватно измерить взаимодействие людей в контексте обсуждения террористических атак и таким образом измерить коллективную составляющую эмоций.

В другой работе на данных 509 млн постов в Twitter 24 млн людей из 84 стран исследовалась динамика индивидуального настроения в течение дня и недели [Golder, Macy, 2011: 1879—1880]. Наличие таких данных позволило рассмотреть суточную и недельную динамику позитивных и негативных аффектов не просто для выборки в целом, но и в разрезе разных социокультурных сред [ibid.: 1879].

Такое исследование было бы невозможно провести с помощью опроса не только потому, что для этого потребовалось бы опрашивать существенное количество людей об их эмоциях каждый час на протяжении года, но это пришлось бы делать сразу в нескольких странах. Кроме того, неизвестно, как связано намерение участвовать в опросе и собственно настроение, если, например, люди в момент плохого настроения менее склонны участвовать в опросах, то это влечет за собой смещения в оценке изучаемого феномена. Однако стоит отметить, что и в случае постов в Twitter могут быть смещения — например, пользователи в зависимости от настроения могут быть более или менее склонны писать посты.

Цифровые следы использовались в качестве данных лонгитюдного формата и для ретроспективных страновых исследований. С помощью текстовых данных из оцифрованных книг, опубликованных за последние 200 лет, исследовалось субъективное благополучие на уровне страны и связанные с ним факторы [Hills et al., 2019]. На основе эмоционального оттенка слов, встречавшихся в этих текстах, был подсчитан национальный индекс валентности для США, Англии, Германии и Италии. Этот показатель валидировался с помощью данных Европарометра — опроса, проводящегося с 1970 г. в этих странах. Корреляция между анкетным вопросом об удовлетворенности жизнью и индексом валентности составила 0,53 [ibid.: 1271—1272]. Это позволило проследить динамику субъективного благополучия в четырех странах с 1820 г. Также были построены модели, оценивающие влияние различных исторических, социальных и экономических факторов на уровень субъективного благополучия в стране [ibid.: 1274].

В похожем по дизайну исследовании на данных миллионов оцифрованных книг, опубликованных за последние 100 лет, изучалось, как изменились ассоциации

и смыслы, вкладываемые в понятие «социальный класс», а также как эти изменения накладывались на социально-экономические трансформации [Kozlowski, Taddy, Evans, 2019]. В еще одном исследовании со схожим дизайном и методами исследовалась динамика и изменение гендерных и этнических стереотипов в США за последние 100 лет, а результаты валидировались с помощью данных переписей населения и государственной статистики [Garg et al., 2018].

Еще один пример использования цифровых следов для исследования динамики и распространения социальных феноменов — работа Кристофер Бейла и коллег про глобальную диффузию вкусов, интересов и потребительских предпочтений на данных поисковых запросов в Google [Bail, Brown, Wimmer, 2019: 1496]. Это исследование примечательно еще и тем, что проверяет классическую теорию социальной имитации французского социального психолога Габриэля Тарда. В работе собраны данные о десяти самых популярных ежемесячных поисковых запросах в Google в 199 странах мира в период с 2004 по 2014 г. На этих данных построена сеть диффузии культурных вкусов и интересов между странами, а также модель, оценивающая взаимосвязь интенсивности диффузии культурных вкусов и интересов из одной страны в другую в конкретный месяц в зависимости от переменных, отражающих статус стран в различных областях (экономика, спорт, искусство, наука, политика, образование и т. д.), истории политического взаимодействия между странами, их географической, культурной и социальной близости и некоторых других факторов [ibid.: 1530—1539].

Для организации такого рода исследования с помощью опроса нужно было бы проводить ежемесячные опросы в 199 странах мира на протяжении десяти лет. Помимо огромных затрат просто на сбор данных, потребовались бы огромные усилия ученых и экспертов со всего мира по стандартизации и унификации анкет для 73 языков, разработке и постоянному обновлению списков культурных практик, предпочтений и интересов для самых разнообразных социокультурных контекстов, а также обработке и подготовке баз данных. Объем работы и финансовые затраты на организацию и проведение такого исследования можно сравнить с запуском нового «Всемирного исследования ценностей». Вышеупомянутое исследование К. Бейла и коллег было сделано усилиями трех ученых и девяти специально обученных ассистентов [ibid.: 1513].

Похожее исследование распространения культурных трендов, а именно моды, между городами всего мира было проведено на данных проекта GeoStyle, который собирает общедоступные фотографии, опубликованные в Instagram и Flickr [Al-Halah, Grauman, 2020].

Сотрудники Microsoft предложили подход к использованию цифровых следов как к несовершенным непрерывным панельным исследованиям [Diaz et al., 2016]. Поскольку сбор данных из социальных сетей не так трудозатратен, как проведение панельного опроса, по цифровым следам из социальных сетей можно наблюдать за некоторой выборкой пользователей на протяжении какого-то времени, таким образом, по сути образуя панельное исследование. Более того, поскольку, как мы уже говорили, цифровые следы можно собирать ретроспективно, то организация и проведение таких цифровых панельных исследований возможны постфактум.

В качестве примера такого подхода можно привести работу Э. Кикимана и коллег, в которой на данных 658 млн постов в Twitter 63 тыс. студентов вузов за пять лет исследовалось влияние раннего употребления алкоголя в университете на различные жизненные и образовательные показатели [Kiciman, Counts, Gasser, 2018]. С помощью анализа текста авторам удалось вычленить среди всех англоязычных пользователей Twitter студентов, недавно поступивших в университеты, а также определить посты о потреблении алкоголя. Затем при помощи методов причинно-следственного анализа было выявлено, что при статистическом контроле многочисленных факторов, также измеренных на данных твитов, студенты, упоминавшие алкоголь во время первого семестра обучения, затем реже упоминали работу и успехи в учебе, но чаще упоминали рискованное поведение, проблемы с законом, их твиты выражали меньше позитивных эмоций [ibid.: 178].

Ограничения цифровых следов и способы их преодоления

Так как цифровые следы — сравнительно новый для современных социологов источник данных, особенно в России, освещение их недостатков и особенностей необходимо для лучшего понимания того, как работать с подобными данными. В этом разделе мы перечислим типичные проблемы, с которыми сталкиваются ученые при построении и реализации исследований с помощью цифровых следов и других новых источников данных.

Нерепрезентативность

Пользователи социальных медиа, таких как Facebook, Twitter, «ВКонтакте», Instagram, которые зачастую служат провайдерами цифровых следов для исследователей, не репрезентируют те популяции, к которым привыкли социологи, конструируя выборки массовых опросов: население страны, региона, города, определенные социальные группы и т. д. [Salganik, 2019: 29—33; Lewis, 2015: 1—2]. Зачастую пользователи таких платформ более молодые, образованные и обеспеченные по сравнению с населением страны, в которой они живут [Hargittai, 2020]. Пользователи Facebook в Америке значительно отличаются от всего населения США по данным переписи населения по возрасту, образованию и доходу [Ribeiro, Benevenuto, Zagheni, 2020]. В Европе были выявлены серьезные межпоколенческие различия в использовании сайтов социальных сетей среди пожилых людей, а также существенная межстрановая вариация в этих различиях [Sala, Gaia, Cerati, 2020].

В России также существуют значительные различия между пользователями социальных сетей [Богданов, Лебедев, 2017]. Так, использование «Одноклассников» выше в более возрастных группах и в малых городах и селах, а «ВКонтакте» популярнее среди более молодых поколений и в более крупных городах [там же: 138—139]. Доля пользователей Facebook больше в крупных городах, а также среди людей с высшим образованием [там же].

Но для разных платформ могут быть характерны разные смещения, варьируются смещения и от страны к стране [Hargittai, 2020: 11—19]. Кроме того, некоторые пользователи платформ могут быть более опытными и активными пользователями интернета и соответствующих платформ [ibid.: 16—19]. Неравномерное распре-

деление активности пользователей на тех или иных платформах выражается, например, в том, что на 10 % активных пользователей Twitter из США приходится 80 % всех постов на этой платформе в этой стране [Wojcik, Hughes, 2019]. Это может приводить к тому, что анализ, основанный на таких данных, может в большей степени отражать закономерности поведения наиболее активных пользователей.

Что с этим можно сделать?

Методы корректировки смещений данных цифровых следов во многом основываются на методах, используемых для взвешивания невероятностных выборок [Beręsewicz et al., 2018: 17—89; Wang et al., 2019]. В качестве наиболее типичного подхода можно назвать постстратификацию, когда данные делятся по некоторым переменным на группы, а затем эти группы перевзвешиваются так, чтобы их размер (доля в выборке) совпадал с размером в генеральной совокупности [Kolenikov, 2016].

Например, для данных из Twitter была разработана модель, которая с помощью нейронных сетей предсказывает принадлежность аккаунта человеку или организации, а также пол и возраст на основе фотографии профиля, имени, никнейма и небольшого описания профиля, причем делает это для различных европейских языков [Wang et al., 2019]. На предсказанных данных, используется многоуровневое регрессионное моделирование для расчета постстратификационных весов для полу-возрастных групп по каждому региону внутри большинства европейских стран [ibidem]. Этот подход позволяет оценить вероятность присутствия человека определенного пола и возраста из определенного региона и страны в Twitter и на основе этих данных скорректировать смещенность выборки по полу и возрасту для конкретного региона или страны [ibid.: 2065—2066]. Предобученные нейронные сети для определения вышеупомянутых характеристик пользователя, а также код для использования этих моделей находятся в открытом доступе и могут быть свободно использованы² [ibid.: 2066].

Для корректировки смещений в данных из Facebook можно использовать подход, апробированный на американской популяции пользователей в [Ribeiro, Benevenuto, Zagheni, 2020].

Однако эти методы применимы только если склонность к присутствию или конкретной активности не связана с исследуемым феноменом. В противном случае методы корректировки смещенности данных, полученных с этой платформы, не исправят положение вещей. В этом смысле чрезвычайно важно различать две категории исследовательских вопросов: 1) исследование феноменов, происходящих на платформе социальных медиа, и производство выводов о пользователях платформы, 2) исследование с помощью данных о пользователях той или иной платформы феноменов, существующих вне этой платформы и, вероятно, охватывающих не только пользователей конкретной платформы [Olteanu и др., 2019: 3—4].

По-видимому, как в случае и с опросными данными, для цифровых следов пока не существует идеальных инструментов, позволяющих корректировать произвольные смещения в данных. Поэтому детальная и дотошная фиксация потенциальных

² Все материалы доступны по ссылке: <https://github.com/euagendas/>.

смещений является витальным аспектом работы, в том числе и с данными цифровых следов [Ruths, Pfeffer, 2014: 1063].

Один из способов проверить надежность результатов исследования — апробировать исследование на данных других онлайн-платформ [ibidem].

Конструктная валидность

Цифровые следы и другие новые данные представляют собой побочный продукт функционирования различных онлайн-сервисов, мобильных приложений и интернет-платформ. Природа происхождения этих данных зачастую никак не связана с теми исследовательскими вопросами, на которые с помощью них могут пытаться ответить ученые [Salganik, 2019: 25—29].

Это влечет за собой проблему, связанную с конструктной валидностью, — теоретические концепты, на которых основывается исследование, могут не быть напрямую отражены в данных, процесс генерации которых никак не связан с целями исследования [Salganik, 2019: 25—29; Lazer, 2015; Rafaeli, Ashtar, Altman, 2019]. Если в случае анкетных вопросов операционализация, то есть переход от теоретического концепта к его измеримым показателям, происходит до исследования и контролируется исследователем, то в случае данных цифровых следов это происходит *ad hoc*, после того как данные уже собраны.

Очевидно, что проблема конструктной валидности в большей степени относится к таким латентным, ненаблюдаемым и неповеденческим концептам, как, например, дружба, ценности, социальные установки, предубеждения и т.д., и в меньшей степени к более однозначным переменным вроде возраста, пола, семейного статуса, дохода и т. п. [Lazer, Radford, 2017: 30; Olteanu et al., 2019: 4].

Так, было показано, что основанные на данных из Twitter метрики политической поддержки кандидатов и партий скорее говорят о временном повышении интереса общественности к политике и политическим фигурам, а не о намерении голосовать за тех или иных политических акторов [Jungherr et al., 2017].

Социальные сети дружбы, измеренные через поведение на онлайн-платформах, сильно разнятся в зависимости от используемых метрик интенсивности «дружбы» [Golder, Macy, 2014: 142], а также заметно отличаются от дружбы, измеренной через анкетные вопросы [Gilbert, Karahalios, 2009].

Что с этим можно сделать?

Один из способов проверки конструктной валидности — использование и сравнение нескольких метрик, потенциально измеряющих один и тот же конструкт. Например, в описанном в предыдущем разделе исследовании территориального распределения благополучия на данных Twitter сравнивалось сразу нескольких моделей предобработки и анализа текста [Jaïdka et al., 2020: 2—3]. Причем сравнение проводилось с условно истинными значениями изучаемых концептов [разными аспектами благополучия], измеренными с помощью анкетных вопросов [ibid.: 3—4].

Быстро и без траты дополнительных ресурсов также можно проверить конструктную валидность исследования, сформулировав результаты не в терминах изучаемых концептов, а напрямую в тех метриках, которые использовались в анализе [Salganik, 2019: 25—30].

Беспрогрышный способ добавить в анализ нужные переменные — собрать дополнительные данные [ibidem]. Например, можно попросить респондентов дать согласие на сбор их данных из социальных сетей и таким образом дополнить данные опроса цифровыми следами респондентов. Подробнее о возможностях, проблемах и способах соединения опросных данных и цифровых следов можно прочесть в [Baghal et al., 2019; Stier et al., 2019].

Внешние и внутренние вмешивающиеся факторы

Работая с цифровыми следами, нужно понимать, что эти данные в подавляющем большинстве случаев являются побочным продуктом функционирования коммерческих платформ, основная цель которых — это коммерческая деятельность [Salganik, 2019: 60—80; Olteanu et al., 2019: 10]. Поэтому процесс производства цифровых следов во многом определяется алгоритмами функционирования этих платформ и их изменениями. Если некоторые алгоритмы функционирования платформ и изменения в этих алгоритмах видны и доступны для исследователей и обычных пользователей, то другие нет и составляют коммерческую тайну компании.

Например, давно известный феномен социальных сетей людей к формированию закрытых триад — друг моего друга с большей вероятностью тоже будет моим другом — используется онлайн-платформами для рекомендации потенциальных друзей пользователям [Salganik, 2019: 60—80; Ruths, Pfeffer, 2014: 1063]. Однако мы не знаем, как именно онлайн-платформы используют это и на что еще опираются, рекомендуя друзей и подписки своим пользователям. В этом смысле сетевая структура, полученная на данных такой платформы, прямым образом зависит от алгоритмов, используемых на платформе, и вопрос об отделении эффекта алгоритмов от социальных эффектов при формировании сети остается открытым [Ruths, Pfeffer, 2014: 1063].

Еще один пример влияния алгоритмов онлайн-платформ — это персонализированные результаты поисковой выдачи в поисковиках [Olteanu et al., 2019: 11]. Такие алгоритмы являются собственностью компаний и засекречены, поэтому, например, исследования цифрового поведения на основе результатов выдачи по поисковым запросам могут быть смещены, и измерить это смещение будет чрезвычайно сложно, если вообще возможно.

Недавно Instagram при поиске фотографий по хештегу #море выдавал пользователям предупреждение о том, что «публикации со словами или тегами, которые вы ищете, зачастую поддерживают поведение, которое может привести к причинению вреда или даже смерти. Если вы столкнулись с какими-то трудностями, мы всегда готовы помочь вам»³. Такое предупреждение может быть результатом тестирования алгоритмов определения суицидальных намерений и их профилактики. Очевидно, что это нововведение может серьезным образом изменить поведение людей на онлайн-платформе и, как следствие, повлиять на процесс генерации данных.

Что с этим можно сделать? К сожалению, мы мало что можем сделать с теми факторами, о которых мало что знаем и которые зачастую составляют коммер-

³ Инстаграм отобрал последнюю радость — возможность смотреть на фотографии моря. Приложение пишет, что такие снимки могут «привести к смерти» // Meduza. 2020. 5 мая. URL: <https://meduza.io/shapiro/2020/05/05/instagram-otobral-poslednyyu-radost-vozmozhnost-smotret-na-fotografiu-morya> [дата обращения: 20.02.2021].

ческую тайну. Однако что касается открытых и заметных для пользователей изменений в функционировании онлайн-платформ, то мы можем фиксировать эти изменения и учитывать при интерпретации результатов анализа. Кроме того, если это позволяют данные, то можно сравнивать результаты анализа при различных состояниях онлайн-платформ — например, до и после какого-то нововведения.

Нестационарность

Еще одно ограничение цифровых следов и других источников новых данных — нестационарность, то есть постоянно или периодически меняющийся процесс генерации данных и зависимостей внутри них. Это может быть связано как с изменениями внутренних алгоритмов функционирования онлайн-платформ, так и с другими факторами [Salganik, 2019: 70—80]. Например, композиция пользователей по их характеристикам на онлайн-платформах может меняться за счет роста или спада популярности той или иной платформы среди определенных пользователей. Также может меняться не только состав пользователей, но и паттерны поведения пользователей на платформе. Например, популяризация стикеров, эмоджи и GIF-изображений представленных на платформе, способна изменить то, как и какие тексты люди пишут на платформе.

Классический пример нестационарности цифровых следов — предсказание эпидемии гриппа в США на данных поисковых запросов в Google [Butler, 2013; Lazer et al., 2014]. Сначала точность предсказаний была высокая, то есть между поисковыми запросами и эпидемией наблюдалась связь, однако затем эти феномены перестали быть связанными и точность предсказаний серьезно снизилась [Lazer et al., 2014].

Что с этим можно сделать? Как и в случае с внутренними изменениями функционирования платформы, с нестационарностью процессов генерации данных на таких платформах мы мало что можем сделать. Фиксация изменений в характеристиках пользователей онлайн-платформ, из которых собираются данные, а также фиксация изменений в практиках пользования этими платформами могут быть разумными и прозрачными практиками для осмыслиения и сравнения результатов исследований с помощью цифровых следов.

Заключение

В этой работе, опираясь на актуальные исследования и публикации в передовых социологических и мультидисциплинарных журналах в области вычислительных социальных наук, мы описали возможности использования цифровых следов в контексте недостатков и ограничений массовых опросов. К новым источникам данных и методам их обработки и анализа пока существует некоторое недоверие, вызванное не столько их ограничениями (ибо они свойственны и другим данным и методам), сколько тем, что они новы и непривычны, а у многих социологов нет навыков работы с такими данными [Edelmann et al., 2020: 75]. При этом за последние несколько лет набралось уже достаточно исследований, которые мы частично описали в этой статье, позволяющих понять, для каких исследовательских задач эти данные и методы подходят, а для каких нет, а также как их нужно использовать. Правильное использование для решения релевантных исследовательских задач

позволит преодолеть многие ограничения традиционных источников данных и откроет большие возможности для получения нового знания. Мы надеемся, что наш обзор вдохновит российских исследователей на то, чтобы активнее использовать новые источники количественных данных и методы их обработки и анализа.

Список литературы (References)

- Богданов М. Б., Лебедев Д. В. Пользование сетью интернет в России в 2003—2015 гг. // Вестник Российской мониторинга экономического положения и здоровья населения НИУ ВШЭ (RLMS-HSE) / отв. ред.: П. М. Козырева. Вып. 7. М.:НИУ ВШЭ, 2017. С. 129—145.
- Bogdanov M. B., Lebedev D. V. (2017) The use of the Internet in Russia from 2003 to 2015. *Russian Longitudinal Monitoring Survey—HSE*. No. 7. P. 129—145. (In Russ.)
- Девятко И. Ф. Новые данные, новая статистика: от кризиса воспроизводимости к новым требованиям к анализу и представлению данных в социальных науках // Социологические исследования. 2018. № 12. С. 30—38.
- Devyatko I. F. (2018) New Data, New Statistics: from Reproducibility Crisis toward New Requirements to Data Analysis and Presentation in Social Sciences. *Sociological Studies*. No. 12. P. 30—38. (In Russ.)
- Abrahao B., Parigi P., Gupta A., Cook K. S. (2017) Reputation Offsets Trust Judgments Based on Social Biases Among Airbnb Users. *Proceedings of the National Academy of Sciences*. Vol. 114. No. 37. P. 9848—9853. <https://doi.org/10.1073/pnas.1604234114>.
- Al-Halah Z., Grauman K. (2020) From Paris to Berlin: Discovering Fashion Style Influences Around the World. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. P. 10136—10145. <http://arxiv.org/abs/2004.01316>.
- Azucar D., Marengo D., Settanni M. (2018) Predicting the Big 5 Personality Traits From Digital Footprints on Social Media: A Meta-Analysis. *Personality and Individual Differences*. Vol. 124. P. 150—159. <https://doi.org/10.1016/j.paid.2017.12.018>.
- Baghal T. A., Sloan L., Jessop C., Williams M. L., Burnap P. (2019) Linking Twitter and Survey Data: The Impact of Survey Mode and Demographics on Consent Rates Across Three UK Studies. *Social Science Computer Review*. P. 517—532. <https://doi.org/10.1177/0894439319828011>.
- Bail C. A., Brown T. W., Wimmer A. (2019) Prestige, Proximity, and Prejudice: How Google Search Terms Diffuse across the World. *American Journal of Sociology*. Vol. 124. No. 5. P. 1496—1548. <https://doi.org/10.1086/702007>.
- Bail C. A., Merhout F., Ding P. (2018) Using Internet Search Data to Examine the Relationship Between Anti-Muslim and Pro-ISIS Sentiment in U. S. Counties. *Science Advances*. Vol. 4. No. 6. eaao5948. <https://doi.org/10.1126/sciadv.aao5948>.
- Bailey M., Cao R., Kuchler T., Stroebel J. (2018) The Economic Effects of Social Networks: Evidence From the Housing Market. *Journal of Political Economy*. Vol. 126. No. 6. P. 2224—2276. <https://doi.org/10.1086/700073>.

Bailey M., Cao R., Kuchler T., Stroebel J., Wong A. (2018) Social Connectedness: Measurement, Determinants, and Effects. *Journal of Economic Perspectives*. Vol. 32. No. 3. P. 259—280. <https://doi.org/10.1257/jep.32.3.259>.

Beręsewicz M., Lehtonen R., Reis F., Di Consiglio L., Karlberg M. (2018) An Overview of Methods for Treating Selectivity in Big Data Sources. Publications Office of the European Union. URL: <https://ec.europa.eu/eurostat/documents/3888793/9053568/KS-TC-18-004-EN-N.pdf/52940f9e-8e60-4bd6-a1fb-78dc80561943> (accessed: 26.02.2021).

Blumenstock J., Cadamuro G., On R. (2015) Predicting Poverty and Wealth From Mobile Phone Metadata. *Science*. Vol. 350. No. 6264. P. 1073—1076. <https://doi.org/10.1126/science.aac4420>.

Bond R. M., Fariss C. J., Jones J. J., Kramer A. D. I., Marlow C., Settle J. E., Fowler J. H. (2012) A 61-Million-Person Experiment in Social Influence and Political Mobilization. *Nature*. Vol. 489. No. 7415. P. 295—298. <https://doi.org/10.1038/nature11421>.

Butler D. (2013) When Google Got Flu Wrong. *Nature News*. Vol. 494. No. 7436. P. 155—156. <https://doi.org/10.1038/494155a>.

Chae D. H., Clouston S., Hatzenbuehler M. L., Kramer M. R., Cooper H. L. F., Wilson S. M., Stephens-Davidowitz S. I., Gold R. S., Link B. G. (2015) Association Between an Internet-Based Measure of Area Racism and Black Mortality. *PLOS ONE*. Vol. 10. No. 4. e0122963. <https://doi.org/10.1371/journal.pone.0122963>.

Chae D. H., Clouston S., Martz C. D., Hatzenbuehler M. L., Cooper H. L. F., Turpin R., Stephens-Davidowitz S., Kramer M. R. (2018) Area Racism and Birth Outcomes Among Blacks in the United States. *Social Science & Medicine*. Vol. 199. P. 49—55. <https://doi.org/10.1016/j.socscimed.2017.04.019>.

Chetty R., Hendren N., Kline P., Saez E. (2014) Where is the land of Opportunity? The Geography of Intergenerational Mobility in the United States. *The Quarterly Journal of Economics*. Vol. 129. No. 4. P. 1553—1623. <https://doi.org/10.1093/qje/qju022>.

Cranmer S. J., Desmarais B. A. (2017) What Can We Learn From Predictive Modeling? *Political Analysis*. Vol. 25. No. 2. P. 145—166. <https://doi.org/10.1017/pan.2017.3>.

Diaz F., Gamon M., Hofman J. M., Kıcıman E., Rothschild D. (2016) Online and Social Media Data As an Imperfect Continuous Panel Survey. *PLOS ONE*. Vol. 11. No. 1. e0145406. <https://doi.org/10.1371/journal.pone.0145406>.

Edelmann A., Wolff T., Montagne D., Bail C. A. (2020) Computational Social Science and Sociology. *Annual Review of Sociology*. Vol. 46. P. 61—81. <https://doi.org/10.1146/annurev-soc-121919-054621>.

Enghoff O., Aldridge J. (2019) The Value of Unsolicited Online Data in Drug Policy Research. *International Journal of Drug Policy*. Vol. 73. P. 210—218. <https://doi.org/10.1016/j.drugpo.2019.01.023>.

Garcia D., Mitike Kassa Y., Cuevas A., Cebrian M., Moro E., Rahwan I., Cuevas R. (2018) Analyzing Gender Inequality Through Large-Scale Facebook Advertising Data. *Proceedings of the National Academy of Sciences*. Vol. 115. No. 27. P. 6958—6963. <https://doi.org/10.1073/pnas.1717781115>.

Garcia D., Rimé B. (2019) Collective Emotions and Social Resilience in the Digital Traces After a Terrorist Attack. *Psychological Science*. Vol. 30. No. 4. P. 617—628. <https://doi.org/10.1177/0956797619831964>.

Garg N., Schiebinger L., Jurafsky D., Zou J. (2018) Word Embeddings Quantify 100 Years of Gender and Ethnic Stereotypes. *Proceedings of the National Academy of Sciences*. Vol. 115. No. 16. P. E3635—E3644. <https://doi.org/10.1073/pnas.1720347115>.

Garip F. (2020) What Failure to Predict Life Outcomes Can Teach Us. *Proceedings of the National Academy of Sciences*. Vol. 117. No. 15. P. 8234—8235. <https://doi.org/10.1073/pnas.2003390117>.

Gee L. K., Jones J., Burke M. (2017) Social Networks and Labor Markets: How Strong Ties Relate to Job Finding on Facebook's Social Network. *Journal of Labor Economics*. Vol. 35. No. 2. P. 485—518. <https://doi.org/10.1086/686225>.

Gee L. K., Jones J. J., Fariss C. J., Burke M., Fowler J. H. (2017) The paradox of weak ties in 55 countries. *Journal of Economic Behavior & Organization*. Vol. 133. P. 362—372. <https://doi.org/10.1016/j.jebo.2016.12.004>.

Gilbert E., Karahalios K. (2009) Predicting Tie Strength With Social Media. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. P. 211—220. <https://doi.org/10.1145/1518701.1518736>.

Golder S. A., Macy M. W. (2011) Diurnal and Seasonal Mood Vary with Work, Sleep, and Daylength Across Diverse Cultures. *Science*. Vol. 333. No. 6051. P. 1878—1881. <https://doi.org/10.1126/science.1202775>.

Golder Scott A., Macy M. W. (2014) Digital Footprints: Opportunities and Challenges for Online Social Research. *Annual Review of Sociology*. Vol. 40. No. 1. P. 129—152. <https://doi.org/10.1146/annurev-soc-071913-043145>.

Hargittai E. (2020) Potential Biases in Big Data: Omitted Voices on Social Media. *Social Science Computer Review*. Vol. 38. No. 1. P. 10—24. <https://doi.org/10.1177/0894439318788322>.

Herdağdelen A., State B., Adamic L., Mason W. (2016) The social ties of immigrant communities in the United States. *Proceedings of the 8th ACM Conference on Web Science — WebSci '16*. P. 78—84. <https://doi.org/10.1145/2908131.2908163>.

Hills T. T., Proto E., Sgroi D., Seresinhe C. I. (2019) Historical Analysis of National Subjective Wellbeing Using Millions of Digitized Books. *Nature Human Behaviour*. Vol. 3. No. 12. P. 1271—1275. <https://doi.org/10.1038/s41562-019-0750-z>.

Hobbs W. R., Burke M., Christakis N. A., Fowler J. H. (2016) Online Social Integration Is Associated With Reduced Mortality Risk. *Proceedings of the National Academy*

of Sciences. Vol. 113. No. 46. P. 12980—12984. <https://doi.org/10.1073/pnas.1605554113>.

Hofman J. M., Sharma A., Watts D. J. (2017) Prediction and Explanation in Social Systems. *Science*. Vol. 355. No. 6324. P. 486—488. <https://doi.org/10.1126/science.aai3856>.

Huang C. (2019) Social Network Site Use and Big Five Personality Traits: A Meta-Analysis. *Computers in Human Behavior*. Vol. 97. P. 280—290. <https://doi.org/10.1016/j.chb.2019.03.009>.

Iannelli L., Giglietto F., Rossi L., Zurovac E. (2018) Facebook Digital Traces for Survey Research: Assessing the Efficiency and Effectiveness of a Facebook Ad — Based Procedure for Recruiting Online Survey Respondents in Niche and Difficult-to-Reach Populations. *Social Science Computer Review*. P. 462—476. <https://doi.org/10.1177/0894439318816638>.

Jaidka K., Giorgi S., Schwartz H. A., Kern M. L., Ungar L. H., Eichstaedt J. C. (2020) Estimating Geographic Subjective Well-Being From Twitter: A Comparison of Dictionary and Data-Driven Language Methods. *Proceedings of the National Academy of Sciences*. P. 10165—10171. <https://doi.org/10.1073/pnas.1906364117>.

Jungherr A., Schoen H., Posegga O., Jürgens P. (2017) Digital Trace Data in the Study of Public Opinion: An Indicator of Attention Toward Politics Rather Than Political Support. *Social Science Computer Review*. Vol. 35. No. 3. P. 336—356. <https://doi.org/10.1177/0894439316631043>.

Kashyap R., Fatehkia M., Al Tamime R., Weber I. (2020) Monitoring Global Digital Gender Inequality Using the Online Populations of Facebook and Google. *Demographic Research*. Vol. 43. No. 27. P. 779—816. <https://doi.org/10.4054/DemRes.2020.43.27>.

Kiciman E., Counts S., Gasser M. (2018) Using Longitudinal Social Media Analysis to Understand the Effects of Early College Alcohol Use. *Proceedings of 12th International Conference on Web and Social Media (ICWSM-18)*. Vol. 12. No. 1. P. 171—180.

Kizilcec R. F., Bakshy E., Eckles D., Burke M. (2018) Social Influence and Reciprocity in Online Gift Giving. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems — CHI '18*. P. 1—11. <https://doi.org/10.1145/3173574.3173700>.

Kolenikov S. (2016) Post-Stratification or Non-Response Adjustment? *Survey Practice*. Vol. 9. No. 3. P. 1—12. <https://doi.org/10.29115/SP-2016-0014>.

Kosinski M., Stillwell D., Graepel T. (2013) Private Traits and Attributes Are Predictable From Digital Records of Human Behavior. *Proceedings of the National Academy of Sciences*. Vol. 110. No. 15. P. 5802—5805. <https://doi.org/10.1073/pnas.1218772110>.

Kossinets G. (2006) Effects of Missing Data in Social Networks. *Social Networks*. Vol. 28. No. 3. P. 247—268. <https://doi.org/10.1016/j.socnet.2005.07.002>.

- Kozlowski A. C., Taddy M., Evans J. A. (2019) The Geometry of Culture: Analyzing the Meanings of Class through Word Embeddings. *American Sociological Review*. Vol. 84. No. 5. P. 905—949. <https://doi.org/10.1177/0003122419877135>.
- Lavrakas P. (2008) *Encyclopedia of Survey Research Methods*. Sage Publications, Inc. <https://doi.org/10.4135/9781412963947>.
- Lazer D. (2015) Issues of Construct Validity and Reliability in Massive, Passive Data Collections. *The City Papers: An Essay Collection from the Decent City Initiative*. URL: <http://citiespapers.ssrc.org/issues-of-construct-validity-and-reliability-in-massive-passive-data-collections/> (accessed 26.02.2021).
- Lazer D., Kennedy R., King G., Vespignani A. (2014) The Parable of Google Flu: Traps in Big Data Analysis. *Science*. Vol. 343. No. 6176. P. 1203—1205.
- Lazer D., Pentland A., Adamic L., Aral S., Barabási A.-L., Brewer D., Christakis N., Contractor N., Fowler J., Gutmann M. (2009) Computational Social Science. *Science*. Vol. 323. No. 5915. P. 721—723.
- Lazer D., Radford J. (2017) Data ex Machina: Introduction to Big Data. *Annual Review of Sociology*. Vol. 43. No. 1. P. 19—39. <https://doi.org/10.1146/annurev-soc-060116-053457>.
- Lewis K. (2013) The Limits of Racial Prejudice. *Proceedings of the National Academy of Sciences*. Vol. 110. No. 47. P. 18814—18819. <https://doi.org/10.1073/pnas.1308501110>.
- Lewis K. (2015) Three Fallacies of Digital Footprints. *Big Data & Society*. Vol. 2. No. 2. P. 1—4. <https://doi.org/10.1177/2053951715602496>.
- Ma X., Cheng J., Iyer S., Naaman M. (2019) When Do People Trust Their Social Groups? *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems—CHI ’19*. P. 1—12. <https://doi.org/10.1145/3290605.3300297>.
- Martin T., Hofman J. M., Sharma A., Anderson A., Watts D. J. (2016) Exploring Limits to Prediction in Complex Social Systems. *Proceedings of the 25th International Conference on World Wide Web—WWW ’16*. P. 683—694. <https://doi.org/10.1145/2872427.2883001>.
- Mauss M. (2000) *The Gift: The Form and Reason for Exchange in Archaic Societies*. New York. NY: W. W. Norton.
- Molina M., Garip F. (2019) Machine Learning for Sociology. *Annual Review of Sociology*. Vol. 45. P. 27—45.
- Olteanu A., Castillo C., Diaz F., Kıcıman E. (2019) Social Data: Biases, Methodological Pitfalls, and Ethical Boundaries. *Frontiers in Big Data*. Vol. 2. P. 1—33. <https://doi.org/10.3389/fdata.2019.00013>.
- Orben A., Dienlin T., Przybylski A. K. (2019) Social Media’s Enduring Effect on Adolescent Life Satisfaction. *Proceedings of the National Academy of Sciences*. Vol. 116. No. 21. P. 10226—10228. <https://doi.org/10.1073/pnas.1902058116>.

- Pierson E., Simoiu C., Overgoor J., Corbett-Davies S., Jenson D., Shoemaker A., Ramachandran V., Barghouty P., Phillips C., Shroff R., Goel S. (2020) A Large-Scale Analysis of Racial Disparities in Police Stops Across the United States. *Nature Human Behaviour*. Vol. 4. No. 7. P. 1—10. <https://doi.org/10.1038/s41562-020-0858-1>.
- Rafaeli A., Ashtar S., Altman D. (2019) Digital Traces: New Data, Resources, and Tools for Psychological-Science Research. *Current Directions in Psychological Science*. Vol. 28. No. 6. P. 560—566. <https://doi.org/10.1177/0963721419861410>.
- Reis B. Y., Brownstein J. S. (2010) Measuring the Impact of Health Policies Using Internet Search Patterns: The Case of Abortion. *BMC Public Health*. Vol. 10. No. 1. P. 1—5. <https://doi.org/10.1186/1471-2458-10-514>.
- Ribeiro F. N., Benevenuto F., Zagheni E. (2020) How Biased is the Population of Facebook Users? Comparing the Demographics of Facebook Users with Census Data to Generate Correction Factors. *12th ACM Conference on Web Science*. P. 325—334. <http://arxiv.org/abs/2005.08065>.
- Ruths D., Pfeffer J. (2014) Social Media for Large Studies of Behavior. *Science*. Vol. 346. No. 6213. P. 1063—1064. <https://doi.org/10.1126/science.346.6213.1063>.
- Sala E., Gaia A., Cerati G. (2020) The Gray Digital Divide in Social Networking Site Use in Europe: Results From a Quantitative Study. *Social Science Computer Review*. P. 1—18. <https://doi.org/10.1177/0894439320909507>.
- Salganik M. (2019) *Bit by Bit: Social Research in the Digital Age*. Princeton, NJ: Princeton University Press. URL: <https://www.bitbybitbook.com/en/1st-ed/preface/> (accessed: 26.02.2021).
- Salganik M. J., Lundberg I., Kindel A. T., Ahearn C. E., Al-Ghoneim K., Almaatouq A., Altschul D. M., Brand J. E., Carnegie N. B., Compton R. J., Datta D., Davidson T., Filippova A., Gilroy C., Goode B. J., Jahani E., Kashyap R., Kirchner A., McKay S., ... McLanahan S. (2020) Measuring the Predictability of Life Outcomes With a Scientific Mass Collaboration. *Proceedings of the National Academy of Sciences*. Vol. 117. No. 15. P. 8398—8403. <https://doi.org/10.1073/pnas.1915006117>.
- Settanni M., Azucar D., Marengo D. (2018) Predicting Individual Characteristics from Digital Traces on Social Media: A Meta-Analysis. *Cyberpsychology, Behavior, and Social Networking*. Vol. 21. No. 4. P. 217—228. <https://doi.org/10.1089/cyber.2017.0384>.
- Shmueli G. (2010) To Explain or to Predict? *Statistical Science*. Vol. 25. No. 3. P. 289—310. <https://doi.org/10.1214/10-STS330>.
- Sivak E., Smirnov I. (2019) Parents Mention Sons More Often Than Daughters on Social Media. *Proceedings of the National Academy of Sciences*. Vol. 116. No. 6. P. 2039—2041. <https://doi.org/10.1073/pnas.1804996116>.
- Smirnov I. (2019) Schools Are Segregated by Educational Outcomes in the Digital Space. *PLOS ONE*. Vol. 14. No. 5. e0217142. <https://doi.org/10.1371/journal.pone.0217142>.

Smirnov I. (2020) Estimating Educational Outcomes From Students' Short Texts on Social Media. *EPJ Data Science*. Vol. 9. No. 1. P. 1—11. <https://doi.org/10.1140/epjds/s13688-020-00245-8>.

Smirnov I. (2018) Predicting PISA Scores from Students' Digital Traces. *Proceedings of the Twelfth International AAAI Conference on Web and Social Media (ICWSM 2018)*. Vol. 3. P. 360—365.

Smith T. W. (1984) Recalling Attitudes: An Analysis of Retrospective Questions on the 1982 GSS. *Public Opinion Quarterly*. Vol. 48. No. 3. P. 639—649. <https://doi.org/10.1086/268865>.

Stephens-Davidowitz S. (2014) The Cost of Racial Animus on a Black Candidate: Evidence Using Google Search Data. *Journal of Public Economics*. Vol. 118. P. 26—40. <https://doi.org/10.1016/j.jpubeco.2014.04.010>.

Stewart I., Flores R. D., Riffe T., Weber I., Zagheni E. (2019) Rock, Rap, or Reggaeton?: Assessing Mexican Immigrants' Cultural Assimilation Using Facebook Data. *The World Wide Web Conference*. P. 3258—3264. <https://doi.org/10.1145/3308558.3313409>.

Stier S., Breuer J., Siegers P., Thorson K. (2019) Integrating Survey Data and Digital Trace Data: Key Issues in Developing an Emerging Field. *Social Science Computer Review*. Vol. 38. No. 5. P. 503—516. <https://doi.org/10.1177/0894439319843669>.

Stopczynski A., Sekara V., Sapiezynski P., Cuttome A., Madsen M. M., Larsen J. E., Lehmann S. (2014) Measuring Large-Scale Social Networks with High Resolution. *PLOS ONE*. Vol. 9. No. 4. e95978. <https://doi.org/10.1371/journal.pone.0095978>.

Wang Z., Hale S., Adelani D. I., Grabowicz P., Hartman T., Flöck F., Jurgens D. (2019) Demographic Inference and Representative Population Estimates from Multilingual Social Media Data. *The World Wide Web Conference*. P. 2056—2067. <https://doi.org/10.1145/3308558.3313684>.

Warren J. R., Halpern-Manners A. (2012) Panel Conditioning in Longitudinal Social Science Surveys: *Sociological Methods & Research*. Vol. 41. No. 4. P. 491—534. <https://doi.org/10.1177/0049124112460374>.

Wojcik S., Hughes A. (2019) How Twitter Users Compare to the General Public. *Pew Research Center: Internet, Science & Tech*. URL: <https://www.pewresearch.org/internet/2019/04/24/sizing-up-twitter-users/> (accessed: 26.02.2021).

Yarkoni T., Westfall J. (2017) Choosing Prediction Over Explanation in Psychology: Lessons From Machine Learning. *Perspectives on Psychological Science*. Vol. 12. No. 6. P. 1100—1122. <https://doi.org/10.1177/1745691617693393>.

DOI: [10.14515/monitoring.2021.1.1756](https://doi.org/10.14515/monitoring.2021.1.1756)



М. Ю. Александрова

МЕТОДЫ МАШИННОГО ОБУЧЕНИЯ В СОЦИОЛОГИЧЕСКОМ ИССЛЕДОВАНИИ: ПРЕДСКАЗАНИЕ ЧАСТИЧНОГО НЕОТВЕТА С ИСПОЛЬЗОВАНИЕМ НАИВНОГО БАЙЕСОВСКОГО КЛАССИФИКАТОРА

Правильная ссылка на статью:

Александрова М.Ю. Методы машинного обучения в социологическом исследовании: предсказание частичного неответа с использованием наивного байесовского классификатора // Мониторинг общественного мнения: экономические и социальные перемены. 2021. № 1. С. 329—350. <https://doi.org/10.14515/monitoring.2021.1.1756>.

For citation:

Aleksandrova M. Y. (2021) Machine Learning in Social Research: Predicting Item Nonresponse Error Using Naive Bayes Classifier. *Monitoring of Public Opinion: Economic and Social Changes*. No. 1. P. 329–350. <https://doi.org/10.14515/monitoring.2021.1.1756>. (In Russ.)

МЕТОДЫ МАШИННОГО ОБУЧЕНИЯ В СОЦИОЛОГИЧЕСКОМ ИССЛЕДОВАНИИ: ПРЕДСКАЗАНИЕ ЧАСТИЧНОГО НЕОТВЕТА С ИСПОЛЬЗОВАНИЕМ НАИВНОГО БАЙЕСОВСКОГО КЛАССИФИКАТОРА

АЛЕКСАНДРОВА Марина Юрьевна — преподаватель, аспирант кафедры методов сбора и анализа социологической информации, Национальный исследовательский университет «Высшая школа экономики», Москва, Россия
E-MAIL: myaleksandrova@hse.ru
<https://orcid.org/0000-0002-7683-7750>

Аннотация. Пропущенные данные в социологических исследованиях могут быть связаны с различными причинами, и в данной статье рассматриваются те из них, что появляются в результате незнания, нежелания или затруднения с поиском ответа на отдельные вопросы анкеты у респондента, — частичные неответы (*item nonresponse*). Остро стоит вопрос о предсказании частичных неответов, решение которого позволило бы сократить вероятность появления пропусков в собираемых данных.

В статье показано, как возникновение частичного неответа можно прогнозировать с помощью современных методов текст-майнинга и машинного обучения на примере данных Европейского социального исследования (European Social Survey) по Великобритании. Для решения поставленной задачи использовался метод наивного байесовского классификатора (Naive Bayes Classifier) — популярный метод предсказания класса зависимой переменной на основе текстовых данных. С опорой на научную литературу

MACHINE LEARNING IN SOCIAL RESEARCH: PREDICTING ITEM NONRESPONSE ERROR USING NAIVE BAYES CLASSIFIER

Marina Yu. ALEKSANDROVA¹—Lecturer,
Doctoral Student at the Department of
Collection and Analysis of Sociological
Information

E-MAIL: myaleksandrova@hse.ru
<https://orcid.org/0000-0002-7683-7750>

¹ National Research University Higher School
of Economics, Moscow, Russia

Abstract. Various reasons may cause missing data in social research. The article highlights the non-response errors caused by ignorance, the lack of desire, or difficulty searching for answers to specific questionnaire questions. Predicting item nonresponse, which would help reduce missing data, poses particular concerns. Based on the data from the European Social Survey (UK respondents) this article shows how text mining and machine learning can predict item nonresponse. The study employs the Naive Bayes Classifier, a popular method to predict the class of dependent variables based on textual data. It relies on scientific literature to show how this method performs. The author provides a database combining full wordings of questions, answers, and instructions, and the ESS survey results in the UK. The paper shows how separate models for predicting the occurrence of item nonresponse were trained using the Naive Bayes technique based on the word frequency and TF — IDF weights (their calculations are also provided). The authors evaluated each model for the frequency of error occurrence. As a result,

туру показываем, как работает этот метод. Мы подготовили базу данных, объединяющую полные формулировки вопросов, ответов, инструкций и результатов опросов исследования European Social Survey по Великобритании. Нами показано, как отдельные модели для предсказания появления частичных неответов были обучены с помощью метода наивного байесовского классификатора на основе частот слов и метрики важности слов TF-IDF, процессу расчета которых мы также приводим подробное описание. Каждая из моделей предсказания частичного неответа оценивалась нами с точки зрения частоты возникновения ошибок при получении прогнозов с их помощью. Мы получили списки слов, наличие в вопросах которых статистически чаще сопровождается или не сопровождается частичными неответами. Наши результаты показали, что респонденты менее охотно отвечают на сенситивные вопросы, а некоторые слова, имеющие отношение к процедуре получения ответа на вопрос, статистически чаще пропускаются респондентами.

Ключевые слова: частичный неответ, отказ от ответа, отсутствие ответа, «затрудняюсь ответить», наивный байесовский классификатор, текст-майнинг, европейское социальное исследование, машинное обучение, качество измерения

lists of words causing or not causing item nonresponse errors were obtained. The results show that respondents are less likely to answer sensitive questions; certain words related to the procedure of getting an answer to a question can also lead to high levels of item nonresponse.

Keywords: item nonresponse, refusal to answer, no answer, “Don't know” option, naive Bayes classifier, text-mining, European Social Survey, ESS, machine learning, measurement quality

Введение. Проблема и ее актуальность

Проблема пропусков в данных часто рассматривается в литературе в силу понятных причин, и ее решение сохраняет актуальность для социологических исследований. В качестве пропусков в данных в социологии чаще всего подразумевается отсутствие ответа респондента в ситуации, когда этот ответ должен был бы присутствовать — на отдельные вопросы анкеты (частичный неответ, item

nonresponse) или на анкету целиком (полный неответ, unit nonresponse). В данной статье мы рассматриваем только частичные неответы.

Исследование причин возникновения отказа от ответа необходимо для того, чтобы понимать, что именно в разработанной анкете может быть связано с отказами от ответа и какими способами возможно снизить предсказываемую долю таких отказов. Мы поставили перед собой цель построить модель прогнозирования связи между появлением частичных неответов и формулировками вопросов. Модель прогнозирования — это представленная в математической форме информация о влиянии независимых переменных на зависимую в данных, которыми располагает исследователь. Эта информация представляет ценность, так как позволяет строить предположения о поведении зависимой переменной в ситуациях, данные о которых отсутствуют. Модели прогнозирования различаются в зависимости от методов, с помощью которых они могут быть построены (для построения прогнозных моделей используются различные виды регрессий, деревья решений, нейронные сети и многие другие методы), поэтому они могут быть достаточно разнообразны. Благодаря построению модели прогнозирования появления частичных неответов в связи с определенными формулировками вопросов можно понять, какие именно слова в этих вопросах способны повышать вероятность возникновения неответа, а какие, наоборот, могут снижать эту вероятность. Под частичным неответом мы понимаем отказ от ответа, затруднение с ответом и отсутствие ответа на вопросы анкет. Для того чтобы прогнозировать связи между появлением частичных неответов и формулировками вопросов, был выбран активно применяемый в последнее время для работы с текстовыми данными метод — наивный байесовский классификатор.

В качестве данных для нашего исследования использовались результаты опроса Европейского социального исследования (European Social Survey, ESS) — известное научное сравнительное межстрановое исследование, проводящееся каждые два года во многих европейских странах. Для анализа мы использовали данные опроса ESS по Великобритании. При подготовке к анализу перед нами всталась задача объединения данных, собирающихся ESS в ходе опросов, — ответов респондентов с самими формулировками вопросов, так как в исходной базе данных ESS содержатся только краткие формулировки, позволяющие получить лишь примерное представление о хранимой информации в соответствующих им переменных, в то время как для прогнозирования связи между появлением частичных неответов и формулировками вопросов нам были нужны точные формулировки вопросов в том виде, в котором они задавались респондентам. Для этого мы воспользовались технологией веб-скрейпинга, с помощью которой автоматически собрали открытые данные с веб-сайтов, содержащие формулировки вопросов и данные с ответами на эти вопросы, и потом объединили их.

Текстовые данные переводились в числовой формат, который необходим для того, чтобы анализ текста стал возможен, с помощью расчета частоты встречаемости слов (сколько раз то или иное слово встречается в одном тексте) и меры важности слов TF — IDF (частота слова — обратная частота документа, «term frequency — inverse document frequency»). Например, слово «она» может часто встречаться в тексте, а «шапочка» — реже, но последняя все равно будет обладать

высоким значением меры важности TF — IDF для текста сказки «Красная шапочка» в корпусе текстов детских сказок, потому что лучше описывает содержание этой сказки. Модели предсказания частичных неответов сравнивались с точки зрения их точности — доли ошибок первого и второго рода в получаемых с их помощью прогнозах возникновения неответов. Нами были получены списки слов, наличие которых в вопросах анкеты показывает связь с одним из типов частичного неответа, что позволило сделать выводы о том, какие темы могут статистически чаще встречаться вместе с частичными неответами.

Работа метода наивного байесовского классификатора

Интуитивно ясно, что лучше всего задачу прогнозирования связи между частичными неответами и формулировками вопросов было бы решать с помощью анализа текстов, заложенных в анкете: формулировок вопросов, вариантов ответов и т. д. Под «формулировкой» вопросов мы будем понимать наличие или отсутствие определенных слов в вопросе — для выражения одной и той же мысли можно выбирать разные слова, и именно благодаря выбору слов вопрос может быть понят по-разному, он может вызвать разную реакцию. Это прекрасно иллюстрируется примером того, как по-разному можно сформулировать сенситивный вопрос «Вы убили свою жену?», приведенным со ссылкой на А. Бартона [Barton, 1958] С. Садменом и Н. Брэдберном, авторами книги «Как правильно задавать вопросы» [Садмен, Брэдберн 2002: 65].

Для решения поставленной нами задачи необходимо исследовать связь формулировок вопросов с наличием или отсутствием частичных неответов у этих вопросов. Например, в вопросе А присутствует слово «пожалуйста», и у этого вопроса нет частичных неответов, а в вопросе Б есть слово «возраст», и на этот вопрос некоторые респонденты не дали ответ. Мы можем предположить, что вопросы со словом «пожалуйста» будут статистически реже оставаться без ответа, в то время как вопрос со словом «возраст» — чаще оказываться неотвеченным. Конечно, не все вопросы со словом «пожалуйста» обязательно будут иметь ответ, в то время как вопросы со словом «возраст» — неответ, поэтому мы можем говорить только о какой-то вычисляемой вероятности, показывающей, что вопрос с одним словом может статистически чаще оказываться отвеченным, а другой вопрос с другим словом будет чаще пропускаться респондентом. Проанализировав все слова из вопросов таким образом, мы можем в итоге разделить слова на две группы: те, которые скорее связаны с неответами, и те, которые скорее будут связаны с ответами. Можно предположить, что внутри этих совокупностей также какие-то слова будут с большей или с меньшей вероятностью связаны с неответом (например, это могут оказаться слова, связанные с сенситивными темами [Sakshaug, Yan, Tourangeau, 2010]).

Мы описали примерную логику условной вероятности, которая свойственна статистике Байеса [Айвазян и др., 1989: 70]. Например, мы могли сделать такой вывод: при условии, что в вопросе присутствует слово «пожалуйста», вероятность пропуска этого вопроса респондентом будет низкой. Более того, мы можем получить классификацию. На основе всех проанализированных нами вопросов, по которым имелись данные о наличии частичных неответов, мы получаем два

класса: слова, с которыми вопросы анкеты статистически чаще оказываются неответченными, и слова, с которыми вопросы анкеты оказываются статистически чаще отвеченными.

Среди байесовских методов классификации наиболее известны наивный байесовский классификатор и знакомая социологам логистическая регрессия. Мы решили использовать в данной работе метод наивного байесовского классификатора, так как он часто используется в задачах классификации текстов. Большинство методов анализа текстовых данных, как упоминается в научной литературе, требуют большого количества данных [Zhang, 2005: 187]. Одним из достоинств наивного байесовского классификатора называют то, что ему не требуются большие обучающие выборки для оценки параметров, важных для классификации [*ibid.*: 191]. Наивный байесовский классификатор основывается на модели условной вероятности. Условная вероятность предполагает измерение вероятности возникновения какого-то события при условии возникновения другого события. Данный метод исследуется уже достаточно давно — начиная с 1960-х годов, и он начал использоваться в первую очередь для решения задач информационного поиска (*information retrieval*) [Maron, 1961: 410]. Информационный поиск — это автоматизированный поиск в коллекции текстовых документов тех, которые подходят по интересующей теме [Manning, Raghavan, Schütze, 2008: 114] (этоту задачу выполняют, например, современные поисковые системы — Google, Yandex, Yahoo и т.д.). Байесовские модели использовались в известном исследовании Ф. Мостеллера и Д. Уоллеса, посвященном определению авторов американского сборника статей в поддержку утверждения Конституции США под названием «Записки Федералиста» [Mosteller, Wallace, 1964]. Долгое время оставалось неизвестным, кто авторы этих статей, так как все они выходили в газетах под псевдонимом «Публий». Благодаря расчету самых употребляемых слов удалось определить реальных авторов «Записок» [Stine, 2019: 294]. Сегодня наивный байесовский классификатор используется также для решения проблемы распределения текстов по различным категориям (поиск спама, определение научных и публицистических статей, тональности публикаций и т.д.) (см., например, [Rennie, 2003; Vadivukarassi, Puviarasan, Aruna, 2017; Ting, Ip, Tsang, 2011]). Таким образом, наивный байесовский классификатор используется для решения весьма разнообразного круга задач, требующих применения коллекций текстов.

Опыт говорит, что читатель-социолог не всегда готов воспринимать научные статьи, подобные цитируемым нами ниже (описывающие работу интересующего нас алгоритма). Поэтому мы решили уделить внимание «переводу» содержания рассматриваемых статей на «язык» задачи о пропусках в ответах на вопросы социологической анкеты.

В основе наивного байесовского классификатора лежит следующий алгоритм. Допустим, есть некий классифицируемый объект (корпус текстов). Объект или текст может быть отнесен к одному из двух каких-то классов (класс выступает в качестве зависимой переменной) — например, это электронные письма, которые являются или не являются спамом, рецензии на книги — положительные и отрицательные, а также вопросы из анкет социологических исследований, при ответе на которые возникали или не возникали частичные неответы. Для

каждого из классов объекта вычисляются функции правдоподобия, на основе которых рассчитываются апостериорные вероятности этих классов [Ting, Ip, Tsang, 2011: 39].

Наивный байесовский классификатор назван наивным, так как в его основе лежит допущение о независимости признаков объектов (признаки объектов — это, иными словами, независимые переменные): предполагается, что наличие одного признака у какого-то класса не влияет на наличие другого признака у этого же класса. Признаки могут зависеть друг от друга или от других признаков, но их вклад в вероятность отнесения объекта к одному из классов остается независимым [Zulfikar et al., 2017: 3]. Например, в случае прогнозирования частичного неответа на основе формулировок вопросов признаками будут выступать слова, из которых состоят анкетные вопросы, а классом — наличие или отсутствие частичного неответа на соответствующий анкетный вопрос с определенными признаками — словами, с помощью которых он был сформулирован. Так, на вопрос А («Укажите, пожалуйста, все марки мороженого, которые вы приобретали в течение последнего месяца хотя бы один раз»), в котором есть слово «мороженое», ответили все респонденты, в то время как на вопрос Б («Скажите, пожалуйста, приходилось ли вам когда-либо пробовать наркотические вещества?»), в котором присутствовало слово «наркотический», респонденты часто отказывались отвечать. При этом наличие признака (слова) «Укажите» в вопросе А или «Скажите» в вопросе Б никак не влияет на то, что в вопросе А присутствовало слово «мороженое», а в вопросе Б — «наркотический». Именно это и является допущением о независимости признаков, в связи с которым рассматриваемый нами метод называется наивным байесовским классификатором.

Исходя из теоремы Байеса, которая лежит в основе наивного байесовского классификатора, апостериорная вероятность того, что наш объект будет отнесен к определенному классу при соответствующем значении признака, рассчитывается следующим образом [Lynch, Bartlett 2019: 55]:

$$P(c_i|d_i) = \frac{P(d_i|c_i)P(c_i)}{P(d_i)},$$

где $P(c_i|d_i)$ — условная вероятность того, что объект i с признаком d принадлежит к классу c ;

$P(d_i|c_i)$ — условная вероятность того, что объект i , принадлежащий к классу c , обладает признаком d ;

$P(c_i)$ — априорная вероятность класса c ;

$P(d_i)$ — априорная вероятность признака d ;

i — некий объект, у которого могут быть признаки c и который может быть отнесен к какому-то классу d .

Для нашей задачи прогнозирования появления частичных неответов на основе формулировок вопросов наивный байесовский классификатор будет работать следующим образом. Мы имеем набор данных — формулировок вопросов, в котором содержится один признак — отдельные слова из этих вопросов, и два класса — «частичный неответ есть» («Да») и «частичного неответа нет» («Нет»). Мы предполагаем, что наличие определенных слов в вопросе (признаков), будет вли-

ять на то, появится ли частичный неответ в данном вопросе или не появится. Таким образом, мы имеем информацию по каждому вопросу (см. табл. 1)

Таблица 1. Набор данных о формулировках вопросов

Слово из вопроса	Наличие частичного неответа в вопросе с указанным словом
Птица	Да
Доверять	Нет
Доверять	Да
Президент	Нет
Птица	Да
Доверять	Да
Наркотик	Нет
Президент	Да
Наркотик	Нет

Теперь имеющийся набор данных необходимо преобразовать в частотную таблицу, резюмирующую, какое число вопросов вызвало неответы, а какое — не вызвало, для каждого из списка слов (см. табл. 2).

Таблица 2. Частотное распределение вопросов по словам и наличию частичного неответа

Слово из вопроса	Частичный неответ есть	Частичного неответа нет
Птица	2	0
Доверять	2	1
Президент	1	1
Наркотик	0	2
Всего	5	4

Так как данных по всем вопросам в нашем примере немного, то уже на этапе подготовки частотной таблицы можно предположить, что есть слова, которые принадлежат вопросам с большим числом неответов, и слова, которые принадлежат вопросам с меньшим числом неответов.

Теперь могут быть рассчитаны вероятности значений признака (наличие определенного слова в вопросе) при соответствующем классе (частичный неответ есть/нет). Данная вероятность также называется правдоподобием [Lynch, Bartlett, 2019: 55] и рассчитывается следующим образом:

$$P(\text{Есть частичный неответ} \mid \text{Доверять}) = \\ = \frac{P(\text{Доверять} \mid \text{Есть частичный неответ}) \times P(\text{Есть частичный неответ})}{P(\text{Доверять})};$$

Таким образом, апостериорная вероятность класса «Есть частичный неответ» при значении признака «Доверять», будет равна:

$$P(\text{Доверять} | \text{Есть частичный неответ}) = 2/5 = 0,4;$$

$$P(\text{Есть частичный неответ}) = 5/9 = 0,56;$$

$$P(\text{Доверять}) = 3/9 = 0,33.$$

Теперь можно подставить полученные значения для расчета правдоподобия для слова «доверять» быть связанным с частичным неответом в вопросе, в котором это слово встретится:

$$P(\text{Есть частичный неответ} | \text{Доверять}) = 0,4 \times 0,56 / 0,33 = 0,68.$$

Следовательно, слово «доверять» в нашем примере окажется связанным с возникновением не ответа на вопрос, если оно присутствует в его формулировке.

Таким образом с помощью наивного байесовского классификатора для каждого слова может быть рассчитана вероятность того, что его наличие в тексте будет связывать появление данного слова скорее с одним классом, чем с другим. В качестве подготовительного этапа выступает только преобразование текстовой информации в числовую, которая может происходить различным образом (например, на основе частот встречаемости слов или метрики TF — IDF) [Sharma, Singh, 2016], о чём мы расскажем в следующей части нашей статьи.

Процесс перевода текстовых данных в числовой формат для прогнозирования частичных неответов

Чтобы перевести наши формулировки вопросов в формат, который позволит применить наивный байесовский классификатор, необходимо пройти подготовительный этап. Вопрос в виде текста нам как людям, владеющим языком, на котором он написан, понятен, однако для компьютерного анализа он не подходит — это язык, который в первозданном виде будет непонятен и для компьютера, и для наивного байесовского классификатора. Поэтому нам нужно каким-то образом перевести формулировки вопросов в цифры. Вопрос в том, как это можно сделать. Для решения этого вопроса существуют различные методы «перевода» текстовой информации в числовую. Чаще всего для этого применяются расчет частот встречаемости слов и расчет метрики TF — IDF. TF — IDF расшифровывается как «term frequency — inverse document frequency», дословно перевести можно так: «частота слова — обратная частота документа» — она позволяет находить часто упоминаемые и специфичные в анализируемом тексте слова. Например, в уже упоминавшемся примере со сказкой «Красная шапочка» слово «она» будет часто встречаться, как и во многих других текстах, из-за чего мера важности этого слова будет низкой. Зато даже если слово «шапочка» встретится в этой сказке меньше число раз, чем слово «она», мы с трудом найдем текст с такой же частотой упоминания «шапочки». Поэтому данное слово будет обладать высоким значением меры важности TF — IDF.

Упомянутые метрики мы использовали для оцифровки формулировок вопросов ESS.

Описание расчета частот встречаемости слов

Частоты слов, встречающихся в формулировках вопросов, которые были собраны из анкет ESS, рассчитываются следующим образом. На основе всего анализируемого корпуса — коллекции текстов (текстами в нашем случае выступают отдельные вопросы) создается словарь слов, содержащихся в нем,— набор слов, которые встречались во всех наших вопросах. Далее строится матрица, где в столбцах располагаются слова, встретившиеся в нашем массиве, а в строках — все вопросы, имеющиеся в массиве. На пересечении строки и столбца размещаются цифры — количество раз, которое каждое слово встречалось в каждом вопросе [Baayen 2002: 38]. Пример можно увидеть ниже (см. табл. 3).

Таблица 3. Матрица встречаемости слов в массиве текстовых данных

	Are	you	a	Citizen	of	the	UK	...
Are you a citizen of the UK?	1	1	1	1	1	1	1	...
in your main job are you...	1	1	0	0	0	0	0	...
do you have any friends who have come to live in the UK from another country?	0	1	0	0	0	0	0	...
...

Описание расчета меры важности TF — IDF слов

Частота встречаемости слова позволяет определить наиболее и наименее часто встречающиеся слова в изучаемых текстах. Наиболее часто встречающиеся слова могут быть достаточно важными для описания текстов, в которых они встретились. Но в расчете частот встречаемости слов кроется определенный риск, который заключается в том, что самыми часто встречающимися могут оказаться наиболее общеупотребимые, распространенные слова, которые на самом деле не позволяют понять, в чем отличие одного текста от всех остальных. Преобразование текста в числовой формат с учетом различия тематик может помочь нам определить, существуют ли темы, связанные с частичными неответами. Для поиска одновременно и часто употребляемых, но при этом достаточно специфичных слов, используется метрика важности слова TF — IDF [Robertson, 2004].

TF — IDF — это статистическая мера важности отдельного слова в тексте, который является частью некой коллекции текстов — корпуса текстов [Evans, Aceves, 2016: 41]. Рассчитывается данная мера по следующей формуле, представляющей собой произведение частоты слова и обратной частоты документа [Hirschberg, Manning, 2015: 263]:

$$TF - IDF(t,k) = TF(t,k) \times IDF(t,k),$$

где $TF - IDF(t,k)$ — в обозначении метрики подразумевается не минус, а тире;
 t — отдельно взятое слово из какого-то текста из корпуса текстов. Это слово может встречаться и в других текстах данного корпуса тоже, но может и оказаться уникальным для какого-то одного текста;
 k — отдельно взятый текст, в котором встретилось данное слово t ;
 $TF(t,d)$ — частота слова (term frequency) — как часто слово t встречалось в тексте d , которая вычисляется по формуле:

$$TF(t,d) = n_{t,d} / (\sum_k n_{k,d}),$$

где $n_{t,d}$ — количество упоминаний слова t в тексте d ,

$\sum_k n_{k,d}$ — сумма всех слов k , которые есть в тексте d .

Делитель в данной формуле вводится, чтобы избежать смещения из-за того, что анализируемые тексты могут сильно различаться своей длиной: в более длинных текстах количество упоминаний какого-то слова может быть больше просто потому, что сам текст длинный, а не потому, что это слово важно для данного текста.

IDF — обратная частота документа (inverse document frequency), с помощью которой оценивается, насколько редко слово встречается во всем корпусе текстов (если слово присутствует во всех текстах, то данная метрика будет равна нулю). Обратная частота документа вычисляется по формуле:

$$IDF(t,d) = \log \frac{n_d}{|\{d_i \in D | t \in D_i\}|},$$

где n_d — количество всех текстов k в корпусе, а выражение $|\{d_i \in D | t \in D_i\}|$ означает количество только тех текстов d_i из всего корпуса текстов D , в которых обязательно встретилось слово t как минимум один раз [Manning, Raghavan, Schütze, 2008: 3].

Таким образом, $TF - IDF$ использует функцию частоты встречаемости термина в тексте, деленную на логарифмически масштабированную обратную долю текстов, содержащих конкретное слово — общее количество текстов в корпусе, деленные на те, в которых данное слово упоминается хотя бы один раз [Hirschberg, Manning, 2015: 264].

Рассмотрим на примере расчет по данной формуле. Предположим, что у нас есть корпус, состоящий из 100 000 вопросов. Слово «доверять» присутствует в 1000 вопросах. В одном из вопросов, состоящем из 100 слов, слово «доверять» встретилось 5 раз. Частота слова «доверять» в данном вопросе (например, это некий седьмой вопрос из абстрактной анкеты) будет равна 0,05, если посчитать по приведенной выше формуле:

$$TF(\text{доверять, вопрос 7}) = 5 / 100 = 0,05.$$

Обратная частота текста — соотношение количества всех вопросов и количества вопросов, в которых встретилось слово «доверять», будет равна 2, исходя из уже приведенной формулы для ее расчета:

$$IDF(\text{доверять, вопрос 7}) = \log \frac{100000}{1000} = 2.$$

Вес $TF-IDF$ для слова «доверять» в нашем корпусе из 100 000 вопросов составит 0,1 в седьмом вопросе как результат произведения уже посчитанных частоты слова и обратной частоты текста:

$$TF-IDF(t,k) = 0,05 \times 2 = 0,1.$$

Таким образом, с помощью меры важности $TF-IDF$ можно присвоить каждому слову в корпусе текстов вес, который:

- тем выше, чем чаще слово t встречается в небольшом количестве текстов (таким образом, данное слово помогает лучше понять отличие данных текстов от остальных);
- тем ниже, чем реже слово t встречается в отдельно взятом тексте или встречается в большом количестве текстов (из-за чего это слово плохо отличает данные тексты от остальных);
- ниже всего у тех слов t , которые встретились во всех текстах (самые общие, часто употребляемые и неспецифичные слова).

Мы рассмотрели два способа перевода текстовой информации в числовую. После того как этот переход был совершен, можно приступить к использованию наивного байесовского классификатора на полученных данных для прогнозирования связи между появлением частичных неответов и формулировками вопросов (наличия определенных слов в содержании вопроса).

Особенности подготовки данных для дальнейшего обучения модели предсказания частичных неответов

В нашем исследовании использовались данные Европейского социального исследования (European Social Survey, ESS), собранные в Великобритании на английском языке и включающие не только ответы и неответы респондентов, но и полные формулировки вопросов анкет в том виде, в котором они задавались респондентам. В этих данных содержится информация о трех вариантах частичного неответа: отказ от ответа, «нет ответа» и «затрудняюсь ответить».

Для подготовки к анализу мы объединили данные, собранные во время опроса: ответы респондентов с самими формулировками вопросов, так как в исходной базе данных ESS содержатся только краткие формулировки, позволяющие лишь получить представление о хранимой информации в соответствующих им переменных, в то время как для прогнозирования связи между появлением частичных неответов и формулировками вопросов нам нужны именно формулировки вопросов в том виде, в котором они задавались респонденту. Такие исходные формулировки вопросов можно найти на сайте ESS, однако содержащие их анкеты хранятся в неудобном для автоматической обработки формате pdf, а ручная обработка файлов происходила бы значительно медленнее и не смогла бы застраховать от ошибок при вводе данных. Поэтому формулировки вопросов разных волн ESS мы автоматически собирали с сайта инструмента SQP (Survey Quality Predictor)¹. Данный инструмент был разработан для прогнозирования качества измерения анкетных вопросов при участии ESRA — организации, занимающейся проведением

¹ SQP 2.1. URL: <http://sqp.upf.edu> (дата обращения: 12.01.2021).

исследования ESS. В настоящий момент SQP используется в ESS при разработке социологических анкет и в процессе ее перевода на другие языки.

Информация по вопросам хранится в SQP в структурированном виде (см. рис. 1): вопросы могут быть отфильтрованы по источнику — исследованию, в котором была использована та или иная формулировка, по языку, на котором сформулированы вопросы, и по стране, в которой проводилось само социологическое исследование. Основную часть окна занимает таблица, содержащая информацию по каждому из вопросов интересующего исследования, при нажатии на который появляется подробная информация о формулировке соответствующего вопроса и вариантов ответа на него.

The screenshot shows the SQP interface. At the top, there's a navigation bar with links for Home, Database, Studies, Help, FAQ, Limits, Publications, Presentations & Courses, and About us. The main content area has a dark header with the text 'A1 / TvTot / Media use , tv total' and 'ESS Round 1 United Kingdom - English'. Below this, there's a 'Request for Answer Text' section with instructions: 'On an average weekday, how much time, in total, do you spend watching television? Please use this card to answer.' To the right, there's a 'Answer options:' list with six items: 'Less than ½ hour', '½ hour to 1 hour', 'More than 1 hour, up to 2 hours', 'More than 1 ½ hours, up to 2 ½ hours', 'More than 2 hours, up to 2 ½ hours', and 'More than 2 ½ hours, up to 3 hours'. At the bottom, there's a 'Information' section with three entries: 'Authorized Prediction' (0.651), 'User Prediction by User 145' (0.596), and 'My Quality Prediction'. A 'Code question to create my own quality prediction' link is also present.

Рис. 1. Вид хранящихся формулировок вопросов в SQP на официальном сайте

После того как формулировки вопросов из исследования ESS были собраны, они объединялись с собранными ответами респондентов. Для данного исследования в ответах респондентов наиболее важны наличие или отсутствие частичного неответа: выбор респондента между вариантами отказом от ответа (refusal), отсутствием ответа (no answer) и затруднением с ответом (don't know).

Тем не менее на этапе объединения данных с ответами респондентов и соответствующих им формулировок из анкеты было выявлено, что на сайте SQP представлены не все вопросы волн исследования ESS. В результате была получена база с 1456 наблюдениями-вопросами.

Сводная информация по подготовленным данным представлена в таблице 4.

Таблица 4. Сводная информация о данных

	Частичный неответ: есть	Частичный неответ: нет	Обучающая подвыборка	Тестовая подвыборка
Затрудняюсь ответить	1274	182	981	484
Отказ от ответа	453	1003	981	484
Отсутствие ответа	422	1034	981	484

Обучающая и тестовая подвыборки нам нужны для того, чтобы сначала позволить наивному байесовскому классификатору «проанализировать» наши данные и найти закономерности в формулировках вопросов, которые могут влиять на появление или непоявление частичных неответов, а затем проверить полученные результаты. Приведем пример. Пусть в нашей выборке вопросов есть десять вопросов, в которых встретилось слово «доверять». Между обучающей и тестовой подвыборками эти вопросы распределились случайным образом: ничто не влияло на то, что в обучающую подвыборку попали семь вопросов о доверии, а в тестовую — оставшиеся три вопроса о доверии (влияло только соотношение величин подвыборок, которое мы определили в данном случае — 70: 30, то есть 70% — обучающая и 30% — тестовая подвыборка). В результате мы будем получать информацию о том, связано или не связано статистически слово «доверять» с частичным неответом на основе тех семи вопросов, которые попали в обучающую подвыборку. Получив эту информацию, мы сможем проверить ее на трех вопросах — будет ли предсказание о частичных неответах на основе семи вопросов подходить для трех вопросов, которые оказались в тестовой подвыборке.

Обучающая подвыборка представляет собой данные, на основе которых производится поиск связей между независимыми и зависимыми переменными. Тестовая подвыборка — это совокупность данных, на которых проверяются связи, выявленные в обучающей подвыборке [James et al., 2013: 73]. Если тестовая подвыборка подтверждает связи между переменными, которые показала обучающая подвыборка, то это позволяет, во-первых, сравнивать между собой построенные модели, а во-вторых, таким образом осуществляется проверка устойчивости полученных на обучающей подвыборке результатов. Процесс обучения модели проходит следующим образом: на обучающей подвыборке производится поиск связей между переменными. Далее на основе этих связей формируются предположения о поведении зависимой переменной — предсказания. Полученные предсказания проверяются на основе тестовой подвыборки — подтверждают ли данные в ней предсказания о поведении переменных или нет [Kuhn, Johnson, 2013].

Данные между обучающей и тестовой выборкой распределяются случайным образом, то есть мы не можем повлиять на то, в какую подвыборку попадают наблюдения из всей выборки. Это необходимо нам для того, чтобы тестовая и обучающая подвыборки не имели каких-то латентных особенностей, которые влияли бы на получаемые нами результаты. Мы можем проконтролировать, какую долю составят эти две подвыборки. У нас это соотношение получилось следующее: 70% — обучающая подвыборка и 30% — тестовая.

Результаты обучения моделей прогнозирования частичных неответов с помощью наивного байесовского классификатора

Для всех моделей рассчитывались матрицы ошибок (confusion matrix) и коэффициент точности предсказаний (accuracy score), которые являются распространенными способами оценки качества полученных прогнозных моделей. Матрица ошибок — двумерная матрица, показывающая распределение правильных и ошибочных предсказаний, сделанных с помощью обученной модели. Данная матрица показывает количество правильных предсказаний наличия признака (закоди-

рованный как «1») и отсутствия признака (закодированный как «0»), сделанных моделью с использованием отобранных независимых переменных. Коэффициент точности предсказаний резюмирует и обобщает матрицу ошибок, он представляет собой долю тестовой выборки, предсказания для которой оказались верными. Следовательно, чем выше его значение, тем более точные предсказания могут быть получены с помощью построенной модели.

Таблица 5 показывает коэффициенты точности предсказаний отказов от ответа, затруднений с ответом и отсутствия ответов для моделей, обученных на частотах слов и на TF — IDF. Коэффициенты точности показывают, что для предсказания отказов от ответа и затруднений с ответом точнее оказались модели, основанные на метрике TF — IDF, в то время как отсутствие ответа несколько лучше прогнозируется моделью, построенной на частотах слов.

Таблица 5. Качество предсказания частичного неответа

	Отказ от ответа	«Затрудняюсь ответить»	Отсутствие ответа
Counts	0,686	0,843	0,764
TF — IDF	0,740	0,878	0,762

Таблицы 6, 7, 8, демонстрируют матрицы ошибок предсказания отказов от ответа, затруднений с ответом и отсутствия ответа для моделей, обученных на частотах слов и на коэффициентах TF — IDF. В целом можно заметить, что хотя доля ошибок первого и второго рода суммарно становится меньше для моделей, рассчитанных на основе TF — IDF, тем не менее улучшение это происходит неравномерно, более того — с увеличением по какому-то одному из типов ошибок. Например, хотя доля ложноположительных предсказаний отказа от ответа уменьшилась (см. табл. 6) с 78 до 31, доля ложноотрицательных предсказаний увеличилась — с 74 до 95. Данное увеличение произошло за счет снижения доли верных предсказаний наличия отказа от ответа на вопрос (с 65 верных предсказаний сократилось до 44).

Таблица 6. Матрица ошибок предсказания отказа от ответа

Counts	Отсутствие неответа	Наличие неответа	TF — IDF	Отсутствие неответа	Наличие неответа
Отсутствие неответа	267	78	Отсутствие неответа	314	31
Наличие неответа	74	65	Наличие неответа	95	44

Доля ложноположительных предсказаний отсутствия ответа также уменьшилась (с 57 до 29), доля ложноотрицательных предсказаний увеличилась — с 57 до 86 за счет снижения доли верных предсказаний наличия отказа от ответа на вопрос (с 74 верных предсказаний сократилось до 45).

Таблица 7. Матрица ошибок предсказания отсутствия ответа

Counts	Отсутствие неответа	Наличие неответа	TF — IDF	Отсутствие неответа	Наличие неответа
Отсутствие неответа	296	57	Отсутствие неответа	324	29
Наличие неответа	57	74	Наличие неответа	86	45

Предсказание затруднений ответов у респондентов показывает несколько иную картину. Количество верных предсказаний наличия затруднений увеличилось, в то время как отсутствие затруднений с ответом в модели, построенной на TF — IDF, стало предсказываться значительно хуже — только 1 случай был верно угадан моделью (см. табл. 8).

Таблица 8. Матрица ошибок предсказания «Затрудняюсь ответить»

Counts	Отсутствие неответа	Наличие неответа	TF — IDF	Отсутствие неответа	Наличие неответа
Отсутствие неответа	18	40	Отсутствие неответа	1	57
Наличие неответа	36	390	Наличие неответа	2	424

Исходя из примеров слов, представленных в таблице 9, можно заметить, что с более частым появлением частичного неответа демонстрируют связь слова — маркеры сенситивных тематик, так как респондентам может быть некомфортно, неловко отвечать на какие-то вопросы анкеты: вопросы, затрагивающие темы развода (*divorced*), работы и безработицы (*unemployed* — безработный, *money* — деньги, *income* — доход), этнополитические вопросы (*England*, *Ireland*, *Scotland*). Это подтверждается результатами исследований других авторов, которые обращают в целом внимание на то, что сенситивные вопросы могут быть связаны с нежеланием отвечать или желанием отвечать неискренне на такие вопросы (см., например, [Chou, Imai, Rosenfeld, 2020; O'Brien et al., 2006; Sakshaug, Yan, Tourangeau, 2010; Ипатова, Рогозин, 2019 и т.д.]). Более того, отдельные исследования подтверждают получаемые нами результаты. Так, К. Митчелл подтверждает, что именно пропуски, связанные с частичными неответами, являются основной причиной смещений в данных [Mitchell, 2010: 899]. Вопросы, связанные с доходом, нередко ассоциируются со значительным объемом пропущенных данных [Yan, Curtin, Jans, 2010: 152]. Можно предположить, что степень сенситивности этнополитической тематики для британцев тоже высока в связи с историей конфликтов между Англией и Ирландией (знаменитый Ольстерский конфликт, война за независимость Ирландии, теракты Ирландской республиканской армии), а также популярности идей о выходе Шотландии из состава Великобритании (на референдуме о независимости Шотландии 2014 г. мнения

поделились почти поровну — только 55,3% опрошенных пожелали остаться в составе Великобритании²).

Кроме того, с более частым появлением частичных неответов показывают связь слова, поясняющие, как ответить на поставленный вопрос: оценить что-либо «в целом» (overall), использовать вспомогательную карточку (card) — визуальные материалы, помогающие понять и дать ответ на анкетный вопрос (например, карточка с перечислением возможных вариантов ответа на вопрос, какие-то изображения, иллюстрации, с опорой на которые респонденту необходимо ответить).

Таблица 9. Примеры слов, которые связаны с частичными неответами

Тип частичного неответа	Слова в вопросах, на которые статистически чаще оставляют без ответа	Слова в вопросах, на которые отвечают статистически чаще
Отказ от ответа	accommodation accomplishment actively automobile behaviour charitable cheerful childcare church citizenship consumer friendly grandmother hobby housework	divorced examinations economy unemployed responsibility politicians religion England Ireland Scotland overall democracy mother father card
Затруднение с ответом	aid apprenticeship church farmers hobby languages medication participated relationship sports students unable vacation street transport	woman better unemployed difficult democracy police trust old housework area good family supervising health card

² Референдум: Шотландия решила остаться в составе Британии // BBC News. Русская служба. 19.09.2014. URL: https://www.bbc.com/russian/uk/2014/09/140919_scotland_wrap_up.shtml (дата обращения: 21.01.2021).

Тип частичного неответа	Слова в вопросах, на которые статистически чаще оставляют без ответа	Слова в вопросах, на которые отвечают статистически чаще
Отсутствие ответа	able accept accommodation accomplishment achievements adult bank buy calm careful criticise discussed discussions expect expenses	answer competent pity overall employed unemployed child retire money hours husband education mother father income

Можно также заметить, что примерно одни и те же слова в вопросах показывают связь с частичными неответами — предсказывается ли отказ от ответа, затруднение с ответом или отсутствие ответа, а также модели, обученные на основе частоты встречаемости слов и TF — IDF дают примерно одинаковые результаты и схожим образом разделяют слова, которые демонстрируют связь с частичным неответом (вопросы с подобными словами статистически чаще оказываются неотвеченными или, наоборот, чаще даются ответы на вопросы с определенными словами).

Выводы и рекомендации

Результаты анализа данных показывают, что респонденты менее охотно отвечают на вопросы, связанные сенситивными темами, что подтверждает выводы других исследований (см., например, [Chou, Imai, Rosenfeld 2020; O'Brien et al., 2006; Sakshaug, Yan, Tourangeau, 2010; Ипатова, Рогозин, 2019] и т.д.). Это может показаться очевидным результатом, но его польза состоит в том, что его получение подтверждает качество полученных предсказательных моделей: они оказались способны делать то, что делает социолог при разработке анкеты, а именно видеть темы, которые в перспективе могут вызывать у респондентов нежелание отвечать на связанные с ними вопросы.

Более того, обучение моделей предсказания частичных неответов показывает, что некоторые слова, имеющие отношение к самой процедуре ответа на вопрос (использование карточек, необходимость что-то оценить «в целом»), могут быть статистически связаны с более частым возникновением частичных неответов. Данные слова могут маркировать сложные вопросы — требующие использования вспомогательных средств для помощи респонденту в поиске ответа на вопрос (использование карточек) или те, которые требуют размышлений, рефлексии, обращения к памяти респондента.

Тем не менее у данного исследования существуют и некоторые ограничения. Во-первых, построенные модели не учитывают помимо формулировок вопросов иные параметры, которые, безусловно, тоже влияют на возникновение частичных неответов. Это и способ сбора данных [Hansen, Hurwitz, 1946], и наличие интервьюера [Groves, 1989], и характеристики самих респондентов — пол [Francis, Busch, 1975], возраст и уровень образования [Groves, 1979; Schuman, Presser, 1980; Herzog, Dielman, 1985], наличие или отсутствие работы [Bell, 1984] и даже национальность, расовая принадлежность [Sicinski, 1970]. Один и тот же вопрос может быть задан разными интервьюерами совершенно разным респондентам: мужчине или женщине, молодому или пожилому, с высшим образованием или без него, британцу или испанцу. И этот вопрос может показывать разную долю частичных неответов при наличии этих характеристик. При этом данный вопрос не меняется, в анкете он был сформулирован совершенно одинаково — даже в межстранных опросах существуют процедуры, направленные на достижение эквивалентности переводов; и все равно в условной Великобритании неответы будут, например, практически отсутствовать для 80 % анкет, в то время как в условной Испании частичных неответов не будет только в 15 %—20 % анкет. Наше исследование не исключает данные исследовательские направления и не противоречит им, но несет в себе цель попытаться посмотреть на проблему работы с частичными неответами с еще одной стороны. Мы также не исключаем возможность и ценность проведения сравнительного исследования, которое учитывало бы влияние характеристик респондентов, имеющейся информации о проведении опроса и формулировок самих вопросов и соотносило бы их влияние между собой.

Во-вторых, анализ слов из анкетных вопросов по отдельности создает риск потери контекста вопроса. Так, слово «возраст» может встретиться в разных по степени своей сенситивности вопросах: вопрос «В каком возрасте вы пошли в школу?» и вопрос «В каком возрасте у вас был первый секс?» будут иметь разный уровень отвечаемости. Проведенный нами анализ позволяет получить усредненное значение отвечаемости между этими вопросами. Как это ограничение может быть снято? Методы автоматической обработки текстовых данных позволяют разбивать текст не только на отдельные слова (униграммы), но и на словосочетания, состоящие из двух (биграммы) и трех слов (триграмммы) [Bird, Klein, Loper, 2009: 141]. Анализ словосочетаний, а не отдельных слов, может помочь уйти от проблемы «средней температуры» отвечаемости при анализе только отдельных слов («униграмм»).

Для повышения качества результатов также предполагается выполнение следующей работы. Во-первых, увеличение выборки за счет добавления вопросов, удаленных в процессе веб-скрейпинга (автоматизированный сбор данных со страниц веб-сайтов [Mitchell, 2018]) анкетных вопросов данного исследования, позволит разнообразить данные, увеличив разнообразие формулировок вопросов и тем, которые они затрагивают. Во-вторых, увеличение выборки за счет добавления всех респондентов, проходивших опросы ESS на английском языке, позволит нам проверить наличие влияния межнациональных особенностей на частичные неответы, а также проверить, насколько формулировка вопроса сильнее или слабее может влиять на возникновение частичных неответов, чем межнациональные различия.

Более того, возможно провести сравнение работы предсказательных моделей — наивного байесовского классификатора с другими методами машинного обучения для определения лучшего метода для прогнозирования частичного неответа, взяв в качестве основания сравнения точность предсказания. Возможны дальнейшие кластеризация и классификация полученных слов, которые могут показывать связь с частичным неответом, с помощью лучше всего показавшей себя предсказательной модели для формирования рекомендаций по составлению вопросов в анкетах социологических исследований.

Список литературы (References)

- Айвазян С. А., Бухштабер В. М., Енуков И. С., Мешалкин Л. Д. Прикладная статистика: классификация и снижение размерности. М.: Финансы и статистика, 1989.
Aivazyan S.A., Bukhshtaber V.M., Enyukov I.S., Meshalkin L.D. (1989) Applied Statistics: Classification and Dimensionality Reduction. Moscow: Finance and Statistics. (In Russ.)
- Ипатова А. А., Рогозин Д. М. Способы преодоления коммуникативных затруднений в стандартизированном телефонном интервью // Вестник Российского университета дружбы народов. Серия: Социология. 2019. Т. 19. № 1. С. 141—166. <https://www.doi.org/10.22363/2313-2272-2019-19-1-144-166>.
- Ipatova A.A., Rogozin D.M. (2019) Techniques for Communication Repair in the Standardized Telephone Interview. *RUDN Journal of Sociology*. Vol. 19. No. 1. P. 141—166. (In Russ.)
- Садмен С., Брэдберн Н. Как правильно задавать вопросы: введение в проектирование массовых обследований. М.: Институт Фонда «Общественное мнение», 2002.
Sudman S., Bradburn N. (2002) Asking Questions: A Practical Guide to Questionnaire Design. Moscow: Institute of the Public Opinion Foundation. (In Russ.)
- Baayen R. H. (2002) Word Frequency Distributions. Dordrecht: Springer. <https://www.doi.org/10.1007/978-94-010-0844-0>.
- Barton A.J. (1958) Asking the Embarrassing Question. *Public Opinion Quarterly*. Vol. 22. No. 1. P. 67—68. <https://www.doi.org/10.1086/266761>.
- Bell R. (1984) Item Nonresponse in Telephone Surveys: An Analysis of Who Fails to Report Income. *Social Science Quarterly*. Vol. 65. No. 1. P. 207—215.
- Bird S., Klein E., Loper E. (2009) Natural Language Processing With Python: Analyzing Text With the Natural Language Toolkit. Beijing Köln O'Reilly June.
- Chou W., Imai K., Rosenfeld B. (2020) Sensitive Survey Questions With Auxiliary Information. *Sociological Methods & Research*. Vol. 49. No. 2. P. 418—454. <https://www.doi.org/10.1177/0049124117729711>.
- Evans J. A., Aceves P. (2016) Machine Translation: Mining Text for Social Theory. *Annual Review of Sociology*. No. 42. P. 21—50. <https://www.doi.org/10.1146/annurev-soc-081715-074206>.

- Francis J. D., Busch L. (1975) What We Now Know About “I Don’t Knows”. *Public Opinion Quarterly*. Vol. 39. No. 2. P. 207—218.
- Groves R. M. (1979) Actors and Questions in Telephone and Personal Interview Surveys. *Public Opinion Quarterly*. Vol. 43. No. 2. P. 190—205.
- Groves R. M. (1989) Survey Costs and Survey Errors. New York, NY: Wiley.
- Hansen M. H., Hurwitz W. N. (1946) The Problem of Nonresponse in Sample Surveys. *Journal of the American Statistical Association*. No. 41. P. 517—529.
- Herzog A., Dielman L. (1985) Age Differences in Response Accuracy for Factual Survey Questions. *Journal of Gerontology*. Vol. 40. No. 3. P. 350—357.
- Hirschberg J., Manning C. D. (2015) Advances in Natural Language Processing. *Science*. Vol. 349. No. 6245. P. 261—266. <https://www.doi.org/10.1126/science.aaa8685>.
- James G., Witten D., Hastie T., Tibshirani R. (2013) An Introduction to Statistical Learning: With Applications In R. New York, NY: Springer. <https://www.doi.org/10.1007/978-1-4614-7138-7>.
- Kuhn M., Johnson K. (2013) Applied Predictive Modeling. New York, NY: Springer. <https://www.doi.org/10.1007/978-1-4614-6849-3>.
- Lynch S. M., Bartlett B. (2019) Bayesian Statistics in Sociology: Past, Present, and Future. *Annual Review of Sociology*. No. 45. P. 47—68. <https://www.doi.org/annurev-soc-073018-022457>.
- Manning C. D., Raghavan P., Schütze H. (2008) Introduction to Information Retrieval. New York, NY: Cambridge University Press.
- Maron M. E. (1961) Automatic Indexing: An Experimental Inquiry. *Journal of the ACM*. Vol. 8. No. 3. P. 404—417.
- Mitchell C. (2010) Are Divorce Studies Trustworthy? The Effects of Survey Nonresponse and Response Errors. *Journal of Marriage and Family*. Vol. 72. No. 4. P. 893—905. <https://doi.org/10.1111/j.1741-3737.2010.00737.x>.
- Mitchell R. (2018) Web Scraping With Python: Collecting More Data From the Modern Web. Sebastopol, CA: O'Reilly Media.
- Mosteller F., Wallace D. L. (1963) Inference in an Authorship Problem: A Comparative Study of Discrimination Methods Applied to the Authorship of the Disputed Federalist Papers. *Journal of the American Statistical Association*. Vol. 58. No. 302. P. 275—309. <https://www.doi.org/10.2307/2283270>.
- O’Brien E. M., Black M. C., Carley-Baxter L. R., Simon Th. R. (2006) Sensitive Topics, Survey Nonresponse, and Considerations for Interviewer Training. *American Journal of Preventive Medicine*. Vol. 31. No. 5. P. 419—426. <https://www.doi.org/10.1016/j.amepre.2006.07.010>.

Rennie J. D., Shih L., Teevan J., Karger D. R. (2003) Tackling the Poor Assumptions of Naive Bayes Text Classifiers. *Proceedings of the 20th International Conference on Machine Learning (ICML-03)*. P. 616—623.

Robertson S. (2004) Understanding Inverse Document Frequency: On Theoretical Arguments for IDF. *Journal of Documentation*. <https://www.doi.org/10.1108/00220410410560582>.

Sakshaug J. W., Yan T., Tourangeau R. (2010) Nonresponse Error, Measurement Error, and Mode of Data Collection: Tradeoffs in a Multi-Mode Survey of Sensitive and Non-Sensitive Items. *Public Opinion Quarterly*. Vol. 74. No. 5. P. 907—933. <https://www.doi.org/10.1093/poq/nfq057>.

Sharma N., Singh M. (2016) Modifying Naive Bayes Classifier for Multinomial Text Classification. *2016 International Conference on Recent Advances and Innovations in Engineering (ICRAIE)*. IEEE. P. 1—7.

Schuman H., Presser S. (1980) Public Opinion and Public Ignorance: The Fine Line Between Attitudes and Nonattitudes. *American Journal of Sociology*. Vol. 85. No. 5. P. 1214—1225.

Sicinski A. (1970) “Don’t Know” Answers in Cross-National Surveys. *Public Opinion Quarterly*. Vol. 34. No. 1. P. 126—129.

Stine R.A. (2019) Sentiment Analysis. *Annual Review of Statistics and Its Application*. No. 6. P. 287—308. <https://www.doi.org/10.1146/annurev-statistics-030718-105242>.

Ting S. L., Ip W. H., Tsang A. H. C. (2011) Is Naive Bayes a Good Classifier for Document Classification. *International Journal of Software Engineering and Its Applications*. Vol. 5. No. 3. P. 37—46.

Vadivukarassi M., Puviarasan N., Aruna P. (2017) Sentimental Analysis of Tweets Using Naive Bayes Algorithm. *World Applied Sciences Journal*. Vol. 35. No. 1. P. 54—59. <https://www.doi.org/10.5829/idosi.wasj.2017.54.59>.

Yan T., Curtin R., Jans M. (2010) Trends in Income Nonresponse Over Two Decades. *Journal of Official Statistics*. Vol. 26. No. 1. P. 145—164.

Zhang H. (2005) Exploring Conditions for the Optimality of Naive Bayes. *International Journal of Pattern Recognition and Artificial Intelligence*. Vol. 19. No. 2. P. 183—198. <https://www.doi.org/10.1142/S0218001405003983>.

Zulfikar W. B., Irfan M., Alam C. N., Indra M. (2017) The Comparation of Text Mining With Naive Bayes Classifier, Nearest Neighbor, and Decision Tree to Detect Indonesian Swear Words on Twitter. *2017 5th International Conference on Cyber and IT Service Management (CITSM)*. IEEE. P. 1—5. <https://www.doi.org/10.1109/CITSM.2017.8089231>.

DOI: [10.14515/monitoring.2021.1.1750](https://doi.org/10.14515/monitoring.2021.1.1750)



Н. В. Ярцева

АНАЛИЗ БОЛЬШИХ ОБЪЕМОВ ДАННЫХ: ВОЗМОЖНОСТИ GDELT PROJECT ПРИ ИСПОЛЬЗОВАНИИ ЯЗЫКА ПРОГРАММИРОВАНИЯ PYTHON. ОПЫТ ГУМАНИТАРИЯ, РЕШИВШЕГО ПОСТИЧЬ BIG DATA

Правильная ссылка на статью:

Ярцева Н. В. Анализ больших объемов данных: возможности Gdelt Project при использовании языка программирования Python. Опыт гуманитария, решившего постичь Big Data // Мониторинг общественного мнения: экономические и социальные перемены. 2021. № 1. С. 351—367. <https://doi.org/10.14515/monitoring.2021.1.1750>.

For citation:

Yartseva N. V. (2021) Analysing Large Amounts of Data: GDELТ Project's Opportunities Using the Python Programming Language. A Humanities Scholar's Experience With Big Data. *Monitoring of Public Opinion: Economic and Social Changes*. No. 1. P. 351–367. <https://doi.org/10.14515/monitoring.2021.1.1750>. (In Russ.)

АНАЛИЗ БОЛЬШИХ ОБЪЕМОВ ДАННЫХ: ВОЗМОЖНОСТИ GDELT PROJECT ПРИ ИСПОЛЬЗОВАНИИ ЯЗЫКА ПРОГРАММИРОВАНИЯ PYTHON. ОПЫТ ГУМАНИТАРИЯ, РЕШИВШЕГО ПОСТИЧЬ BIG DATA

ЯРЦЕВА Наталья Владимировна — кандидат политических наук, доцент, Самарский университет, Самара, Россия; эксперт-консультант, Всероссийский центр изучения общественного мнения, Москва, Россия

E-MAIL: yartseva.nat@gmail.com
<https://orcid.org/0000-0002-7236-1812>

Аннотация. Научный мир развивается по междисциплинарному пути, одно из самых популярных направлений — соединение возможностей программирования, позволяющего обрабатывать большие объемы данных, и гуманитарного знания. В статье я описываю свой опыт освоения Big Data, анализирую возможности ресурса GDELT и показываю, как с помощью команд на языке программирования Python обрабатывать большие данные. Благодаря этому данная статья может иметь и вполне практическое применение — в ней перечислены и проанализированы шаги, которые позволят ученым, не знакомым с обработкой больших объемов данных, не только разобраться в сути метода, но и самостоятельно сделать первые шаги в обработке данных на Python. Кроме того, работа проиллюстрирована кейсом французских «желтых жилетов», что позволяет лучше разобраться в структуре кода и принципах работы GDELT.

Ключевые слова: GDELT, Big Data, Python, BigQuery

ANALYSING LARGE AMOUNTS OF DATA: GDELT PROJECT'S OPPORTUNITIES USING THE PYTHON PROGRAMMING LANGUAGE. A HUMANITIES SCHOLAR'S EXPERIENCE WITH BIG DATA

Natalia V. YARTSEVA^{1,2} — Cand. Sci. (Polit.), Associate Professor; Expert
E-MAIL: yartseva.nat@gmail.com
<https://orcid.org/0000-0002-7236-1812>

¹ Samara University, Samara, Russia

² Russian Public Opinion Research Center, Moscow, Russia

Abstract. The scientific community is currently following the interdisciplinary path, and one of the most popular directions is a combination of Programming enabling processing large volumes of data and the Humanities. The paper describes the author's experience with Big Data, provides her analysis of the GDELT Project opportunities and shows how large amounts of data can be processed using the Python programming language. The analysis of steps in data processing using Python can help scholars who have not dealt with large amounts of data. The French “yellow vests” is the case study the author uses to illustrate how the GDELT Project works.

Keywords: GDELT, Big Data, Python, BigQuery

В инструментарии современного гуманитарного исследователя не так много технических средств, предназначенных для работы с большими объемами данных. SPSS и «Статистика» прекрасно справляются с обработкой данных, это отличные статистические программы, предназначенные для анализа данных, с широким набором функций и параметров.

Однако и у них есть недостатки — высокие требования к системе компьютера: требуются большие объемы оперативной памяти, память на жестком диске и быстрый процессор; высокая цена.

В этой статье я расскажу:

- что такое GDELT;
- что собой представляет блог GDEL Project;
- что он делает и как повторить его аналитику;
- что может Google Cloud Platform;
- как при помощи языка SQL написать на нем запросы и что-то получить.

Затем попробуем повторить то же самое на Jupiter и Python.

Данные платформы будут представлены на примере движения «желтых жилетов» во Франции. В качестве исследовательской гипотезы я возьму предположение, что французская пресса пишет о «желтых жилетах» в более негативном ключе, чем, например, итальянская. Мотивом этого может служить тот факт, что «желтые жилеты» проводят большее количество демонстраций во Франции, чем в других европейских странах. Поэтому для Франции они условные «бузотеры», выходящие на несанкционированные протесты, а для других стран «желтые жилеты» — граждане, отстаивающие свои права, и тон публикаций на тему данного движения, соответственно, гораздо дружелюбнее.

Обзор GDELT

Платформа данных GDELT представляется принципиально новым направлением, позволяющим обрабатывать большие объемы информации: а) бесплатно и б) не привлекая большой памяти компьютера. Именно эти два критерия, на мой взгляд, и делают GDELT востребованным порталом, позволяющим за короткий срок обрабатывать и систематизировать большие данные.

GDELT¹ — мировая база данных о социальных, политических, экономических и культурологических процессах, происходящих в государствах. GDELT отслеживает мировые вещательные, печатные и веб-новости практически из каждого уголка земного шара на более чем ста языках. Он идентифицирует людей, места, организации, темы, источники, эмоции, цифры, цитаты, изображения и события, которые происходят в мире и влияют на наше общество. GDELT — это бесплатная мировая платформа, предоставляющая доступ к терабайтам самой разной информации.

GDELT в режиме 24/7 отслеживает мировые средства массовой информации практически из каждого уголка земного шара. Данные GDELT привязаны к координатам (широте и долготе), что и позволяет составлять карты: от индекса счастья до карты конфликтов. При этом у GDELT есть несколько сервисов. Сам GDELT — это источник данных, а GDELT Analytical Services — инструмент работы с этими данными.

¹ The GDELT Project. URL: <https://www.gdeltproject.org> (дата обращения: 19.02.2021).

ми. Он и позволяет предсказывать возможность эскалации конфликта (опираясь на совокупность данных и опыт предыдущих конфликтов, которые имели свои особенности) и даже распространения вируса гриппа.

GDELT анализирует в том числе Twitter, популярные телевизионные шоу и даже региональные СМИ, не пишущие на английском языке. Все глобальные новости GDELT отслеживает практически в режиме реального времени, и это составляет 98,4% от ежедневного объема не англоязычных СМИ. Эти материалы переводятся на английский и обрабатываются. В ближайшее время портал намерен расширить свою базу данных до 1800 г.²

Данные GDELT можно скачать в CSV, но поскольку их много (2,5 ТВ/год), то на практике используют либо Google BigQuery, либо GDELT Analytical Services.

Google хранит у себя «копию» GDELT (и не только его). Для анализа GDELT и выборки данных по нужным нам критериям Google предлагает использовать свой сервис — BigQuery public datasets³. Он хорош тем, что данные не нужно загружать на компьютер, а все общение с облаком данных происходит исключительно посредством языка SQL. Данные по указанным критериям отбираются в облаке, а на компьютер исследователя выгружаются только те строки, что подходят под критерии.

Задачи исследования

Обратимся к анализу данных GDELT. Про движение «желтых жилетов» писали все мировые СМИ — с конца 2018 г. и по сей день на эту тему вышло большое количество статей. Давайте попробуем проанализировать через GDELT, как освещались неконтролируемые акции протesta в 2019 г. в Париже и какое место среди этих акций занимали «желтые жилеты»⁴.

Google Cloud Platform — платформа, представляющая собой набор облачных служб и позволяющая проводить разного рода вычисления. Она содержит около 300 различных сервисов, но не предлагает готовых решений. У нее есть русский интерфейс и пояснения для каждой задачи.

У Google Cloud Platform в открытом доступе много полезных наборов данных, облегчающих работу аналитику. В частности, в нем находятся материалы GDELT. Извлечем кусочек информации по «желтым жилетам». Для этого заходим на сервис Google Cloud Public Datasets⁵, где хранятся популярные общедоступные наборы данных в облаке. Здесь можно получить доступ к более чем ста общедоступным наборам данных из разных отраслей и тем. Отсюда можно запрашивать данные непосредственно из GDELT и пользоваться понятным для гуманитария интерфейсом. Нажимаем на кнопку “Explore public datasets” и переходим к анализу общедоступных наборов данных. Далее попадаем на страницу с общедоступными наборами данных, где и находим GDELT.

² Intro // The GDELT Project. URL: <https://www.gdeltproject.org/#intro> (дата обращения: 19.02.2021).

³ В BigQuery public datasets необходимо зарегистрироваться и выбрать по этой ссылке данные с GDELT — URL: <https://console.cloud.google.com/marketplace/product/the-gdelt-project/gdelt-2-events.BigQuery> (дата обращения: 19.02.2021).

⁴ По умолчанию BigQuery API (то, при помощи чего мы будем анализировать данные) может быть выключено. Включить его можно здесь: <https://console.cloud.google.com/apis/library/BigQuery.googleapis.com>.

⁵ Google Cloud Public Datasets. URL: <https://cloud.google.com/public-datasets?hl=ru> (дата обращения: 19.02.2021).

По запросу “gdel” нам выдается несколько баз данных, среди которых выбираем GDEL 2.0 Event Database. По щелчку “Посмотреть набор данных” переходим в редактор запросов. Далее выбираем папку “events” и все дальнейшие действия уже осуществляем в ней.

Все запросы в Google Cloud Platform мы делаем на языке SQL. Этот язык применяется для создания, модификации и управления данными. SQL — это стандарт индустрии при работе с реляционными данными (данными, организованными в таблицы со связями между таблицами)⁶.

В редакторе запросов на языке SQL мы выполняем команду, где после WHERE идут условия фильтрации, а SOURCEURL — это название колонки в нужной нам таблице данных.

```
SELECT *
FROM `gdelt-bq.gdeltv2.events`
WHERE SOURCEURL like "%yellow-vest%"
LIMIT 100
```

Так мы ищем все упоминания «желтых жилетов», но лимит строк ограничиваем 100 штуками. На выходе получается обширная таблица, где собраны: даты, акторы (кто и о ком говорит), очень важный eventcode, который классифицирует событие (о нем ниже), avgTone (в каком тоне) и, наконец, ссылки на те ресурсы, с которых GDEL получает информацию.

EventCode	EventBaseCode	EventRootCode	QuadClass	GoldsteinScale	NumMentions	NumSources	NumArticles	AvgTone	Actor1Geo_Type	Actor1Geo_FullName	Actor1Geo_CountryCode	Actor1Geo_ADMIN1Code
040	040	04	1	1.0	6	1	6	-5.1819184123484	4	Paris, France (general), France	FR	FR00
175	175	17	4	-9.0	2	1	2	-5.1819184123484	4	Paris, France (general), France	FR	FR00
040	040	04	1	1.0	1	1	1	-5.1819184123484	4	Paris, France (general), France	FR	FR00
050	050	05	1	3.5	5	1	5	-5.1819184123484	4	Paris, France (general), France	FR	FR00
0874	087	08	2	10.0	10	1	10	-5.1819184123484	4	Paris, France (general), France	FR	FR00
120	120	12	3	-4.0	5	1	5	-5.1819184123484	4	Paris, France (general), France	FR	FR00
175	175	17	4	-9.0	6	1	6	-5.1819184123484	4	Paris, France (general), France	FR	FR00
100	100	10	3	-5.0	6	1	6	-5.1819184123484	4	London, London, City of, United Kingdom	UK	UKH9
040	040	04	1	1.0	1	1	1	-5.1819184123484	4	London, London, City of, United Kingdom	UK	UKH9

⁶ SQL запросы быстро. Часть 1 //Хабр. 2019. 17 декабря. URL: <https://habr.com/ru/post/480838/> (дата обращения: 19.02.2021).

Полученного количества ответов слишком много, теперь попробуем максимально сузить поиск. Давайте выберем только то, что про «желтые жилеты» писала CNN, например, отфильтровав по URL онлайн-публикации. Для этого выполним команду:

```
SELECT *  
FROM `gdelt-bq.gdeltv2.events`  
WHERE SOURCEURL LIKE 'https://edition.cnn.com/%yellow-vest%'  
LIMIT 100
```

На выходе снова получим таблицу со множеством граф, но уже ту, которая показывает только данные CNN.

ActionGeo_CountryCode	ActionGeo ADM1Code	ActionGeo ADM2Code	ActionGeo_Lat	ActionGeo_Long	ActionGeo_FeatureID	DATEADDED	SOURCEURL
FR	FR00	16282	48.8667	2.33333	-1456928	20190429190000	https://edition.cnn.com/2019/04/25/europe/emmanuel-macron-yellow-vest-ir
FR	FR00	16282	48.8667	2.33333	-1456928	20190429190000	https://edition.cnn.com/2019/04/25/europe/emmanuel-macron-yellow-vest-ir
FR	FR00	16282	48.8667	2.33333	-1456928	20190429190000	https://edition.cnn.com/2019/04/25/europe/emmanuel-macron-yellow-vest-ir
FR	FR00	16282	48.8667	2.33333	-1456928	20190429190000	https://edition.cnn.com/2019/04/25/europe/emmanuel-macron-yellow-vest-ir
FR	FR00	16282	48.8667	2.33333	-1456928	20190429190000	https://edition.cnn.com/2019/04/25/europe/emmanuel-macron-yellow-vest-ir
FR	FR00	16282	48.8667	2.33333	-1456928	20190429190000	https://edition.cnn.com/2019/04/25/europe/emmanuel-macron-yellow-vest-ir
FR	FR00	16282	48.8667	2.33333	-1456928	20190429190000	https://edition.cnn.com/2019/04/25/europe/emmanuel-macron-yellow-vest-ir
n UK	UKH9	40110	51.5	-0.116667	-2601889	20190429190000	https://edition.cnn.com/2019/04/25/europe/emmanuel-macron-yellow-vest-ir
n UK	UKH9	40110	51.5	-0.116667	-2601889	20190429190000	https://edition.cnn.com/2019/04/25/europe/emmanuel-macron-yellow-vest-ir

Обратите внимание на колонку EventBaseCode:

IsRootEvent	EventCode	EventBaseCode
1	175	175
0	175	175
1	175	175
0	175	175
1	175	175
0	175	175
0	175	175

Код 175, так часто присваиваемый репортажам CNN о «желтых жилетах», говорит о действиях неповиновения, подавляемых властью.

Каждую новость GDELT относит к конкретному типу. Это может быть деловая встреча, образовательный проект, уличная акция и еще множество иных вариантов событий. GDELT использует классификацию CAMEO⁷.

Фильтруя по коду 175, мы можем получить ссылки на статьи мировых онлайн-изданий (а не только CNN), освещдающие гражданское неповиновение и применение репрессий. В следующем примере происходит фильтрация по трем полям:

EventBaseCode—что случилось—“175 Use repression”;

ActionGeo_ADM1Code—где случилось—Париж;

Year—когда случилось—2019 г.

⁷ Коды CAMEO с выдержками из медиа см.: <http://data.gdeltproject.org/documentation/CAMEO.Manual.1.1b3.pdf> (дата обращения: 19.02.2021).

Query editor [+ COMPOSE NEW QUERY](#)

```
1 SELECT *
2 FROM `gdel-bq.gdeltv2.events`
3 WHERE EventBaseCode = '175' AND ActionGeo_ADM1Code = 'FR00' AND Year = 2019
4 LIMIT 100
5
```

Run Save query Save view Schedule query More

Query results [SAVE RESULTS](#) [EXPLORE DATA](#)

Query complete (3.0 sec elapsed, 213.5 GB processed)

Job information Results JSON Execution details

20190119234500	https://www.unian.ua/world/10414623-u-franciji-v-protestah-zhovtih-zhileta-vzali-uchast-84-tisyachi-osib.html
20190120091500	http://polit.ru/news/2019/01/20/paris/ 
20190120170000	https://en.trend.az/world/europe/3008150.html

В полученной выборке (см. скриншот) фигурирует статья с сайта polit.ru, в которой говорится о разгоне демонстрантов спецназом с применением слезоточивого газа.

Какие коды нам еще могут быть интересны?

Например, регулярно встречающийся код 170, свидетельствующий о принуждении, репрессиях и насилии в отношении гражданских лиц. Например — предотвращение демонстрации.

Или код 011 — уклонение от комментариев (в любом виде и любой форме). Например, отказ представителей НАТО комментировать те или иные военные события.

Большим преимуществом Google BigQuery в сфере обработки больших объемов данных является то, что он способен (как мы увидели на примерах выше) извлекать из терабайтов данных GDELT только необходимые записи и экспорттировать полученные результаты в таблицу. Эту таблицу можно сохранить на компьютер в формате CSV для дальнейшего анализа в MS Excel.

Если таблички для работы нам хватает, то этот сервис — то, что нам нужно.

GDELT Analytical Services

В процессе освоения GDELT мне очень помог блог GDELT Project⁸, в частности, представленные там графики.

В поисковой строке блога забиваем, скажем, Russia. В ответ выдается большое количество статей блога, внутри которых содержатся графики, где так или иначе фигурирует Россия. Выбираем статью — Russia Fades From Television News («Россия исчезает из телевизионных новостей»). В ней с октября 2016 по июль 2020 г. собрана публикационная активность трех медиапорталов (CNN, Fox News, MSNBC) относительно представленности в американских новостях сообщений о России. Считается упоминание России в процентном соотношении к другим новостям. Если максимальные доли достигали отметки в 12%—14%, то к 2020 г. процент новостей, связанных с Россией, сократился до 2%. Это не так много, учитывая, что одно шоу Рэйчел Мэддоу составляет 5,5% всех упоминаний России.

⁸ The GDELT Project. URL: <https://blog.gdeltproject.org>

О чем говорят подобные графики? В первую очередь — как работает GDELT. По какому из принципов он может собирать и фильтровать данные. И здесь мы видим сразу несколько критериев: данные были собраны по конкретной стране, оценен общий объем материала, выпускаемого СМИ США, и внутри него выделена доля информации о России. При этом на графиках мы видим сравнение нескольких медиапорталов во временном контексте, что наглядно показывает, как изменялся процент упоминания России в трех СМИ.

Мы можем повторить это либо через инструментарий GDELТ Analytical Services, либо самостоятельно, используя Jupiter Notebook и Python (об этом чуть позже).

Здесь действует такой алгоритм.

Открываем страницу инструмента GDELТ Summary — <https://api.gdeltproject.org/api/v2/summary/summary>.

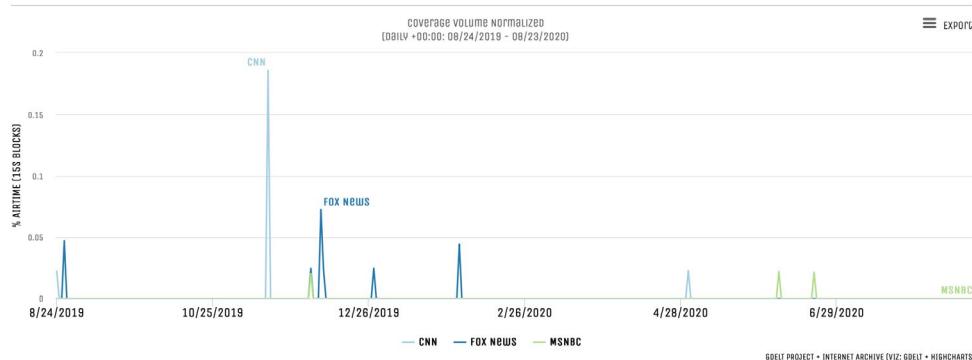
Задаем критерии выборки:

Step 1: Dataset = “Television News”

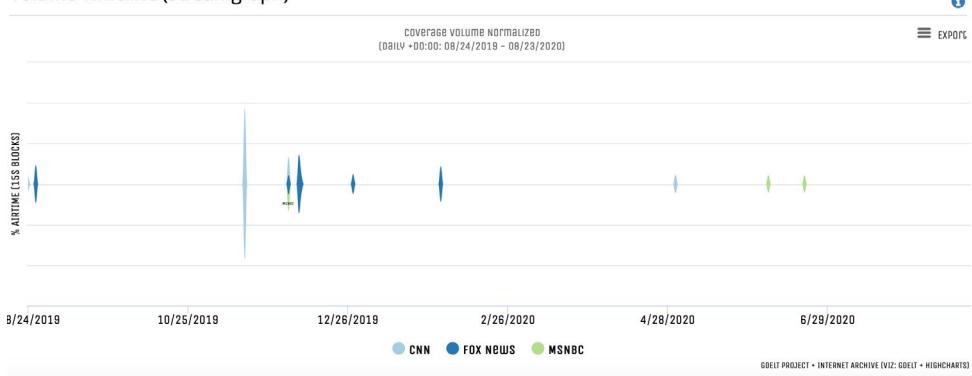
Step 2: Output type = “Summary Overview Dashboard”

Step 3: Search / Keywords = justice–“chief justice”–“justice department”

У меня получилось несколько интересных дашбордов, построенных по тому же принципу, что и блог GDELТ. Например, график за последние два года с процентами эфирного времени, уделяемого «желтым жилетам».

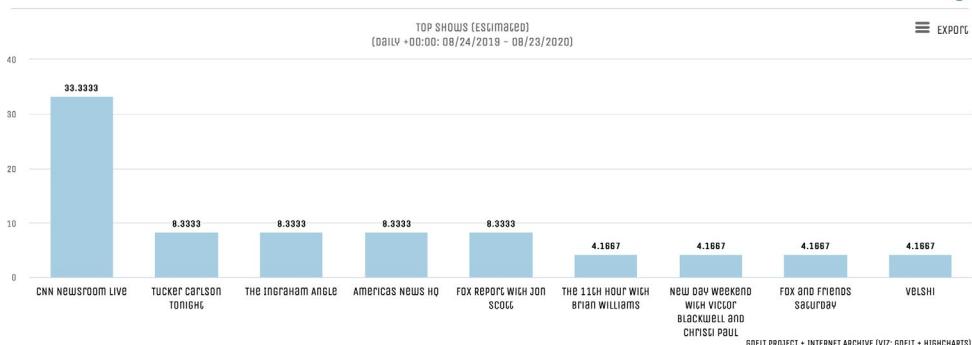


Volume Timeline (Streamgraph)

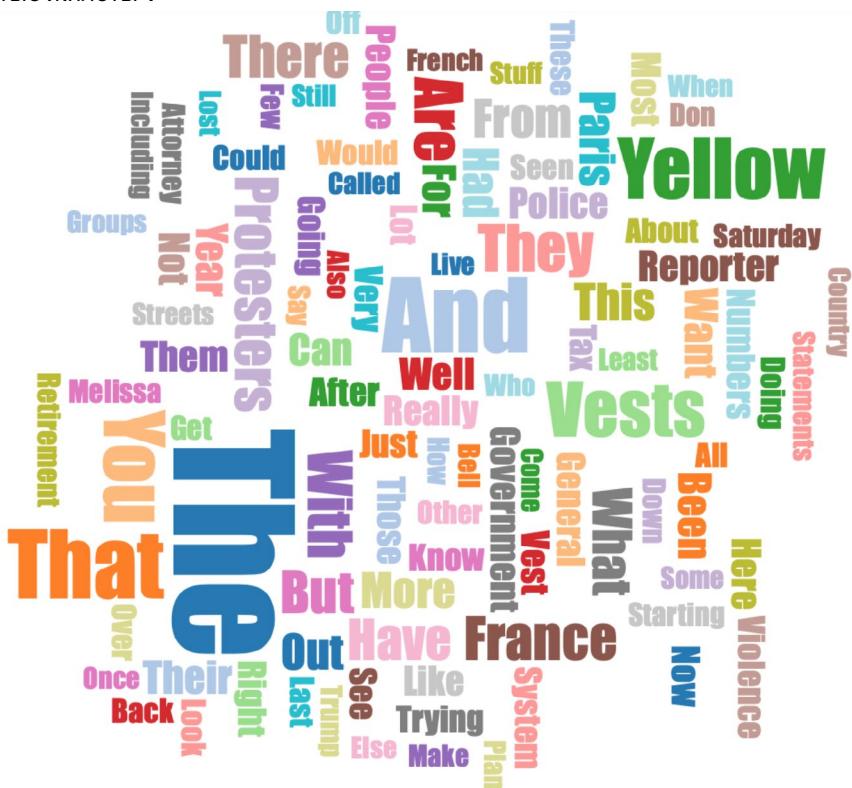


А вот какие шоу были популярны в связи с протестами:

Show Chart (Estimated)



И в завершение — слова, которые чаще всего встречаются со словосочетанием «желтые жилеты»:



Мы использовали только возможности GDELT Analytical Services для того, чтобы получить нужные данные. Попробуем написать код на Python и получить нечто похожее.

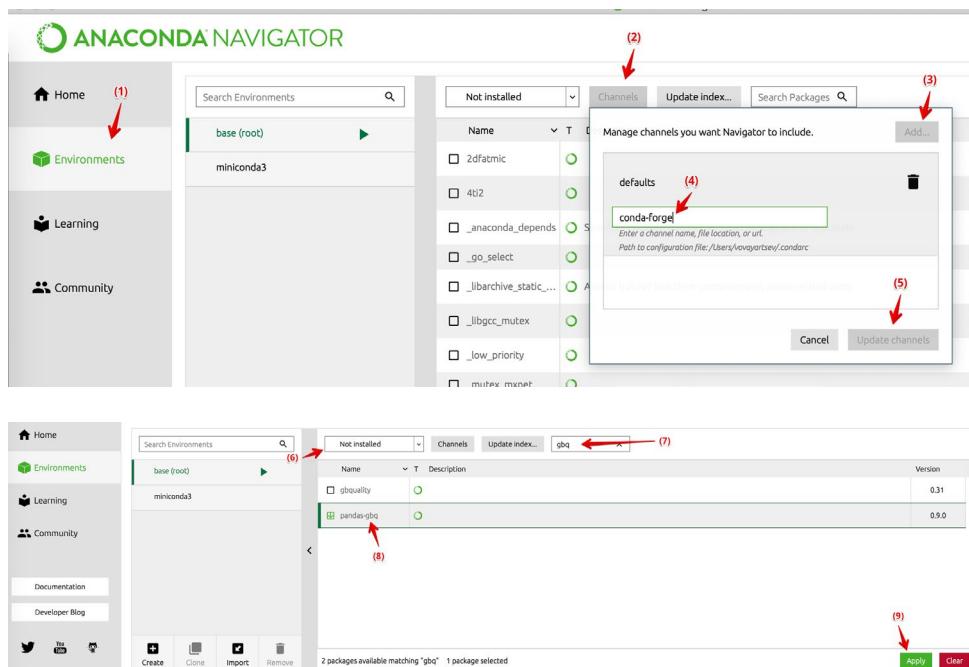
Подготовка окружения для разработки

Jupyter⁹ — это среда для программирования на Python. Jupyter открывается в окне браузера и позволяет создавать и просматривать Jupyter Notebooks — интерактивные документы, содержащие Python-код, данные, графики и поясняющие тексты к ним. В интернете можно найти большое количество руководств по работе с Jupyter¹⁰.

Один из самых быстрых способов настройки рабочей среды — это установка Anaconda¹¹. На сайте в разделе Getting Started приведено 15-минутное обучающее видео (на английском языке).

По шагам это выглядит так:

- скачиваем Anaconda Individual Edition (~450 MB),
- запускаем Anaconda Navigator,
- во вкладке Environments подключаем канал conda-forge и устанавливаем пакет pandas-gbq (он нам потребуется для доступа к Google BigQuery).



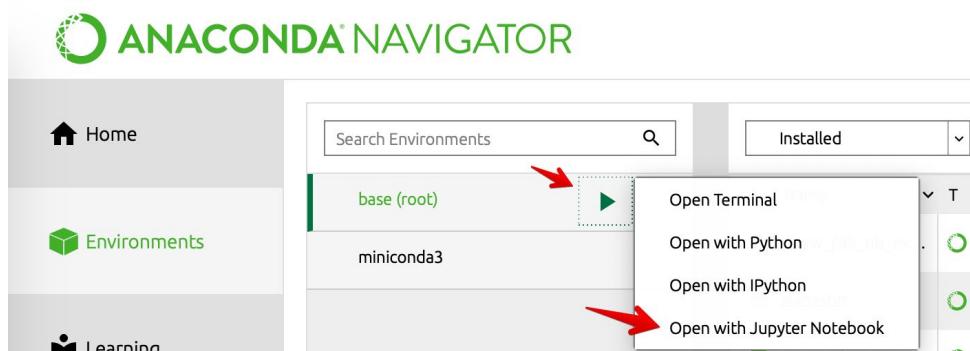
Последнее действие также можно выполнить через командную строку: /opt/anaconda3/bin/conda install -c conda-forge pandas-gbq.

1. Во вкладке Environments щелкаем на треугольной зеленой иконке и выбираем Open with Jupyter Notebook:

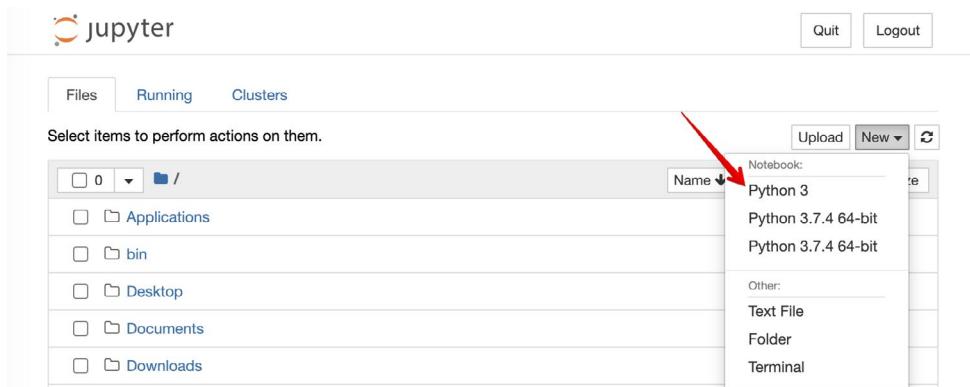
⁹ URL: <https://jupyter.org>.

¹⁰ Особенности Jupyter Notebook, о которых вы (может быть) не слышали // Хабр. 2016. 6 декабря. URL: <https://habr.com/ru/company/wunderfund/blog/316826/> (дата обращения: 19.02.2021).

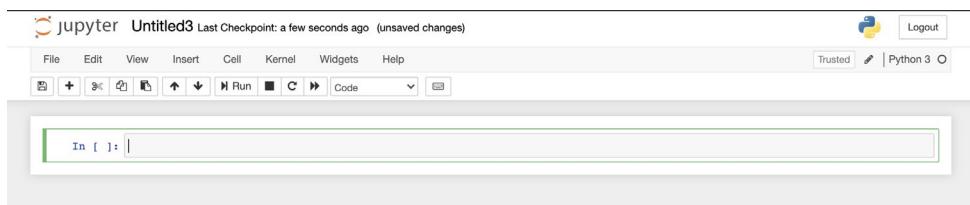
¹¹ URL: <https://www.anaconda.com/>.



2. В открывшемся окне создаем новый Notebook (Python 3):



3. Откроется окно Jupyter, куда можно писать код:



Базовый пример на Python и Jupyter

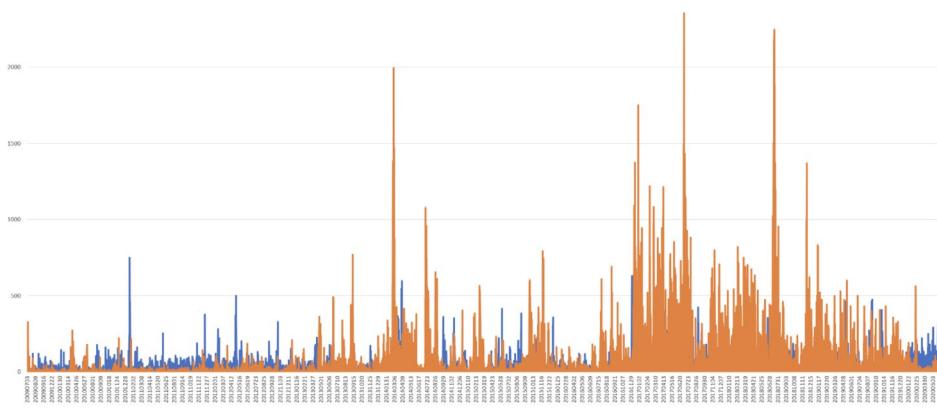
Достоинством подхода, предполагающего написание кода на Питоне, по сравнению с GDEL Analytical Services является то, что он оперирует “сырыми данными” (и может выполнять их предобработку перед построением графиков), и то, что эти графики удобно группируются в блокноты Jupiter Notebook.

Для начала возьмем одну из самых несложных задач для аналитики — упоминания в новостях России и Китая¹²:

¹² Tracking Country Mentions Using Television Ngrams: China vs Russia // The GDEL Project. 2020. 8 June. URL: <https://blog.gdelproject.org/tracking-country-mentions-using-television-ngrams-china-vs-russia/> (дата обращения: 19.02.2021).

Tracking Country Mentions Using Television Ngrams: China vs Russia

© JUNE 8, 2020

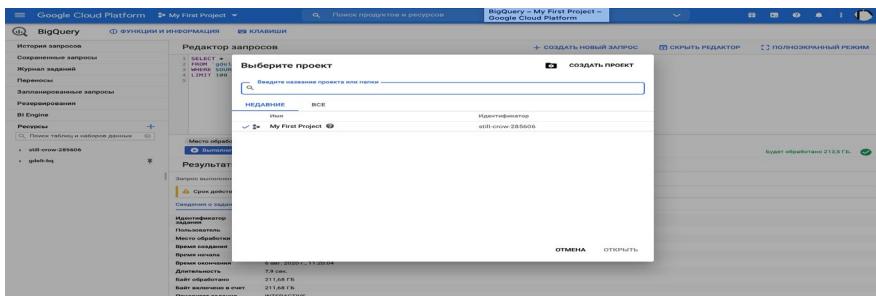


В статье блога приведен SQL-запрос для выборки исходных данных этого графика из Google BigQuery:

```
SELECT DATE, China, Russia from (
SELECT DATE, sum(COUNT) China, 0 Russia FROM `gdelt-bq.gdeltv2.iatv_1gramsv2` where STATION='CNN' and (NGRAM='china' OR NGRAM='chinese' OR NGRAM='beijing') group by DATE
UNION ALL
SELECT DATE, 0 China, sum(COUNT) Russia FROM `gdelt-bq.gdeltv2.iatv_1gramsv2` where STATION='CNN' and (NGRAM='russia' OR NGRAM='russian' OR NGRAM='moscow' OR NGRAM='kremlin' OR NGRAM='putin') group by DATE) order by DATE asc
```

Этот SQL мы возьмем за основу, чтобы воспроизвести график, приведенный в статье GDELT, средствами Python и Jupiter.

Чтобы авторизовать компьютер для работы с Google BigQuery, нам потребуется идентификатор проекта из консоли Google Cloud Platform:



Сначала построим график только по одной стране. Для этого вставим код в окно Jupyter (заменив <ваш-идентификатор-проекта> на значение со скриншота выше) и выполним его, нажав Shift-Enter:

```
[*]: import pandas_gbq
import pandas as pd

# ВАЖНО: это должен быть ID проекта из Вашего аккаунта Google Compute Engine
# (иначе пример не запустится)
project_id = "calcium-bridge-691"

sql = """
SELECT DATE, sum(COUNT) China
FROM `gdelt-bq.gdeltv2.iatv_1gramsv2`
WHERE STATION='CNN' and (NGRAM='china' OR NGRAM='chinese' OR NGRAM='beijing')
GROUP BY DATE
"""
df = pandas_gbq.read_gbq(sql, project_id=project_id)
df['DATE'] = pd.to_datetime(df['DATE'], format='%Y%m%d')
df = df.groupby('DATE').sum()

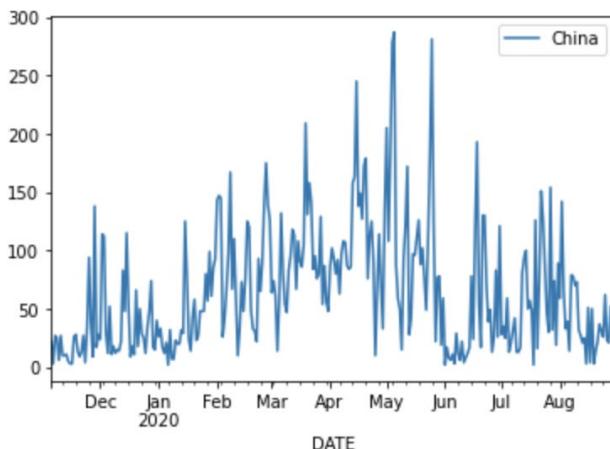
df.tail(300).plot()
```

Please visit this URL to authorize this application: https://accounts.google.com/o/oauth2/auth?response_type=code&scope=https://www.googleapis.com/auth/bigquery&state=urKuyhYYdxee1DD7bMgsLHzlzoTxC&prompt=consent&
Enter the authorization code:

При первом запуске на экране появится запрос авторизации. Нужно пройти по указанной ссылке, в появившемся окне разрешить Pandas GBQ доступ к Google BigQuery, скопировать авторизационный код в окошко ввода в Jupyter и нажать Enter. При последующих запусках запрос авторизации больше появляться не будет.

В результате выполнения приведенного выше кода мы получим график, в котором каждая точка обозначает количество упоминаний Китая в новостях за соответствующий день.

Out[1]: <matplotlib.axes._subplots.AxesSubplot at 0x11c549a00>



Немного усложним SQL-запрос для выборки данных по двум странам и используем функцию resample(), чтобы сгруппировать данные по месяцам:

```
In [7]: import pandas_gbq
import pandas as pd

# ВАЖНО: это должен быть ID проекта из Вашего аккаунта Google Compute Engine
# (иначе пример не запустится)
project_id = "alien-clover-203818"

sql = """
SELECT DATE, sum(China) as chi, sum(Russia) as rus from (
SELECT DATE, sum(COUNT) China, 0 Russia FROM `gdelt-bq.gdeltv2.iatv_lgrams2` where STATION='CNN'
and (NGRAM='china' OR NGRAM='chinese' OR NGRAM='beijing') group by DATE
UNION ALL
SELECT DATE, 0 China, sum(COUNT) Russia FROM `gdelt-bq.gdeltv2.iatv_lgrams2` where STATION='CNN'
and (NGRAM='russia' OR NGRAM='russian' OR NGRAM='moscow' OR NGRAM='kremlin' OR NGRAM='putin') group by DATE
)
GROUP BY DATE
ORDER BY DATE asc
"""

df = pandas_gbq.read_gbq(sql, project_id=project_id)
df['DATE'] = pd.to_datetime(df['DATE'], format='%Y%m%d')

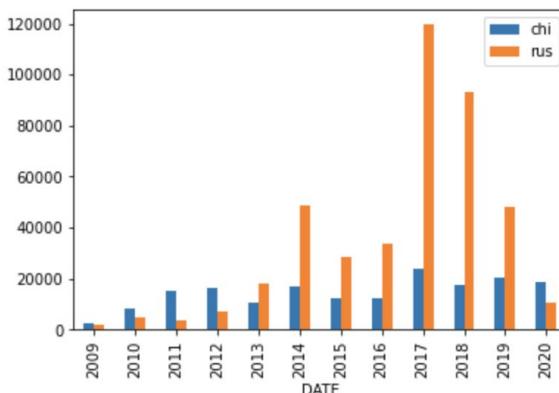
Downloading: 100%|██████████| 3968/3968 [00:01<00:00, 3005.75rows/s]
```

```
In [9]: df.groupby('DATE').sum().resample('M').sum().plot()
Out[9]: <matplotlib.axes._subplots.AxesSubplot at 0x1d4e5760>
```

Этот график совпадает с тем, что мы видели в блоге GDELT, но он более читаемый.
Попробуем сгруппировать упоминания России и Китая по годам:

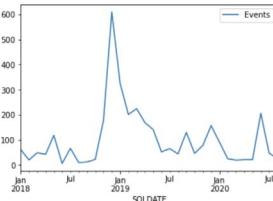
```
In [10]: by_year = df.groupby('DATE').sum().resample('Y').sum()
years = by_year.index.map(lambda dt: dt.year)
by_year.set_index(years).plot.bar()
```

```
Out[10]: <matplotlib.axes._subplots.AxesSubplot at 0x11d2d1490>
```

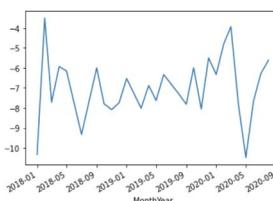
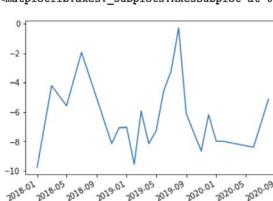


На последнем графике четко виден тренд исчезновения России из мировых новостей.

И в заключение — график количества мировых публикаций, освещавший события в Париже за 2018—2019 гг. с кодом 175 “Use repression” с группировкой по месяцам:

```
In [8]:  
import pandas_gbq  
import pandas as pd  
  
# ВАЖНО: это должен быть ID проекта из Вашего аккаунта Google Compute Engine  
# иначе пример не запустится  
project_id = "alien-clover-203818"  
  
sql = ""  
SELECT SQLDATE, count(*) as Events,  
FROM `gdelt-bq.gdeltv2.events`  
WHERE EventBaseCode = '175' AND ActionGeo_ADM1Code = 'FR00' AND Year > 2017  
GROUP BY SQLDATE  
""  
df = pandas_gbq.read_gbq(sql, project_id=project_id)  
df['SQLDATE'] = pd.to_datetime(df['SQLDATE'], format='%Y%m%d')  
df = df.groupby('SQLDATE').sum()  
  
Downloading: 100%|██████████| 535/535 [00:00<00:00, 1633.52rows/s]  
  
In [9]: df.resample('M').sum().plot()  
Out[9]: <matplotlib.axes._subplots.AxesSubplot at 0x117dalc70>  

```

Также можно проследить изменение тона этих публикаций в отношении демонстрантов и полиции:

```
In [68]:  
sql = ""  
SELECT MonthYear, Actor1Name, avg(AvgTone) as Tone,  
FROM `gdelt-bq.gdeltv2.events`  
WHERE EventBaseCode = '175' AND ActionGeo_ADM1Code = 'FR00' AND Year > 2017  
GROUP BY MonthYear, Actor1Name  
""  
df = pandas_gbq.read_gbq(sql, project_id=project_id)  
  
Downloading: 100%|██████████| 504/504 [00:01<00:00, 498.96rows/s]  
  
In [73]: df_indexed = df.set_index(pd.to_datetime(df['MonthYear'], format='%Y%m'))  
df_indexed[df_indexed.Actor1Name.eq('POLICE')].loc[:, 'Tone'].plot()  
Out[73]: <matplotlib.axes._subplots.AxesSubplot at 0x1187dfa60>  
  
  
In [72]: df_indexed[df_indexed.Actor1Name.eq('DEMONSTRATOR')].loc[:, 'Tone'].plot()  
Out[72]: <matplotlib.axes._subplots.AxesSubplot at 0x1186ed160>  

```

Проверка статистической значимости

После всех выкладок и обилия кода остановимся на проверке статистической значимости моей гипотезы. Напомню, она заключается в том, что французская пресса пишет о «желтых жилетах» в более негативном контексте, чем итальянская.

При проверке статистических гипотез результат имеет статистическую значимость, когда он очень маловероятен при нулевой гипотезе. Нулевая гипотеза — это предположение, что исследуемой закономерности вообще не существует. В данной статье нулевая гипотеза — это предположение, что и французская, и итальянская пресса пишут о желтых жилетах в одинаковом контексте.

Решение о статистической значимости исследования принимают путем сравнения двух значений:

- * р-значение, р — это вероятность получения результата наблюдения при условии, что нулевая гипотеза истинна;
- а — уровень значимости исследования.

По стандартам исследования результат статистически значим, когда $p \leq a$.

Уровень значимости а для исследования выбирается до сбора данных и обычно устанавливается на уровне 5% или ниже — в зависимости от области исследования. В любом эксперименте или наблюдении, включающем выборку из популяции/сочетанности, всегда существует вероятность того, что наблюдаемый эффект произошел бы только из-за ошибки выборки. Но если р-значение наблюдаемого эффекта меньше уровня значимости (или равно ему), исследователь может заключить, что эффект отражает характеристики всей популяции, тем самым отвергая нулевую гипотезу.

В данной статье исходными наблюдениями являются значения AvgTone публикаций на сайтах из доменов .fr и .it за 2019 г. и имеющих в URL фразу «yellow-vest».

```
sql = """
SELECT AvgTone
FROM `gdelt-bq.gdeltv2.events`
WHERE SOURCEURL LIKE 'http%.it/%yellow-vest%'
AND year = 2019
"""
it = pandas_gbq.read_gbq(sql, project_id=project_id)
```

Так как наблюдаемые значения являются нормально распределенными и независимыми, мы можем использовать t-student's test.

После проверки данных на нормальность p-value можно вычислить, используя функцию normaltest из пакета scipy.stats.

```
normaltest(it['AvgTone'])
NormaltestResult(statistic=18,688692012556487,
pvalue=8,74585136912613e-05)

normaltest(fr['AvgTone'])
NormaltestResult(statistic=78,62375287543983,
pvalue=8,454131325695306e-18)
```

```
ttest_ind(fr['AvgTone'], it['AvgTone'], equal_var = False)
Ttest_indResult(statistic= -12,93921478149851,
pvalue=5,768147975657512e-23)
```

Так как $p = 5 \times 10^{-23} < \alpha = 0,05$, гипотезу признаем статистически значимой.

Чтобы проиллюстрировать, что данный метод проверки статистической значимости работает, можно провести А-А тест (подать на вход два фрагмента одной и той же выборки, например сайтов .fr)

```
ttest_ind(fr['AvgTone'].sample(frac=0,5), fr['AvgTone'].sample(frac=0,5), equal_var = False)
```

```
Ttest_indResult(statistic= -0,8931836531373752,
pvalue=0,3720966690912032)
```

В данном случае $p = 0,37 > \alpha = 0,05$, так как обе выборки состоят из публикаций французских СМИ¹³.

Таким образом, выдвинутая гипотеза подтверждается. Французская пресса действительно пишет о «желтых жилетах» в более негативном ключе, чем итальянская.

¹³ Исходные коды всех примеров из данной статьи доступны по адресу <https://github.com/YartsevaNat/GPD>.

DOI: [10.14515/monitoring.2021.1.1899](https://doi.org/10.14515/monitoring.2021.1.1899)



Т. А. Нестик, Н. Н. Седова, Е. Г. Климанова

БУДУЩЕЕ ИССЛЕДОВАТЕЛЬСКОЙ ИНДУСТРИИ: ОТ КОНКУРЕНЦИИ ЗА БЮДЖЕТЫ К ПОИСКУ ПАРТНЕРОВ

Правильная ссылка на статью:

Нестик Т. А., Седова Н. Н., Климанова Е. Г. Будущее исследовательской индустрии: от конкуренции за бюджеты к поиску партнеров // Мониторинг общественного мнения: экономические и социальные перемены. 2021. № 1. С. 368—386. <https://doi.org/10.14515/monitoring.2021.1.1899>.

For citation:

Nestik T. A., Sedova N. N., Klimanova E. G. (2021) The Future of The Research Industry: From Competing for Budgets to Finding Partners. *Monitoring of Public Opinion: Economic and Social Changes*. No. 1. P. 368–386. <https://doi.org/10.14515/monitoring.2021.1.1899>. (In Russ.)

БУДУЩЕЕ ИССЛЕДОВАТЕЛЬСКОЙ ИНДУСТРИИ: ОТ КОНКУРЕНЦИИ ЗА БЮДЖЕТЫ К ПОИСКУ ПАРТНЕРОВ

НЕСТИК Тимофей Александрович —
доктор психологических наук, заве-
дующий лабораторией социальной и
экономической психологии, Институт
психологии РАН, Москва, Россия

E-MAIL: nestik@gmail.com
<https://orcid.org/0000-0002-1410-4762>

СЕДОВА Наталья Николаевна — стар-
ший научный сотрудник Центра ком-
плексных социальных исследований,
Институт социологии ФНИСЦ РАН,
Москва, Россия; советник генераль-
ного директора, Всероссийский центр
исследований общественного мнения
(ВЦИОМ), Москва, Россия

E-MAIL: nnseanova@wciom.com
<https://orcid.org/0000-0001-8492-7847>

КЛИМАНОВА Евгения Геннадьевна —
генеральный директор, Маркетинговое
агентство «Делфи», Омск, Россия
E-MAIL: jane.klimanova@delfi2000.ru
<https://orcid.org/0000-0000-0000-0000>

Аннотация. Стремительные измене-
ния в исследовательской отрасли, вы-
званные цифровой трансформацией
и социальными изменениями, привле-
кают все больше внимания как со сто-
роны академических исследователей,
так и со стороны участников исследо-
вательского рынка. Все более актуаль-
ной задачей становится определение
возможных последствий внедрения
цифровых технологий в исследова-

THE FUTURE OF THE RESEARCH INDUS- TRY: FROM COMPETING FOR BUDGETS TO FINDING PARTNERS

Timofei A. NESTIK¹ — Dr. Sci. (Social Psy-
chol.), Head of the Laboratory of Social
and Economic Psychology
E-MAIL: nestik@gmail.com
<https://orcid.org/0000-0002-1410-4762>

Natalia N. SEDOVA^{2,3} — Senior Research
Fellow, Center for Complex Social Stud-
ies; CEO's councilor
E-MAIL: nnseanova@wciom.com
<https://orcid.org/0000-0001-8492-7847>

Evgeniya G. KLIMANOVA⁴ — General
Director
E-MAIL: jane.klimanova@delfi2000.ru
<https://orcid.org/0000-0000-0000-0000>

¹ Institute of Psychology, Russian Academy of Sciences,
Moscow, Russia

² Institute of Sociology of the FCTAS RAS, Moscow,
Russia

³ Russian Public Opinion Research Center (VCIOM),
Moscow, Russia

⁴ Marketing Agency 'Delfi', Omsk, Russia

Abstract. The rapid changes in the re-
search industry driven by digital trans-
formation and social change attract
more and more attention from academic
researchers and research market partic-
ipants. An increasingly urgent task is to
determine the consequences of intro-
ducing digital technologies into research
processes, including the use of artificial
intelligence systems.

тельские процессы, в том числе с использованием систем искусственного интеллекта.

Цель настоящей статьи — проанализировать технологические, экономические, социально-политические и культурные факторы изменений в исследовательской индустрии и характеристики складывающейся под их влиянием новой экосистемы рынка прикладных исследований. Наш анализ опирается на результаты форсайт-сессии «Исследовательская индустрия: тренды и окна возможностей», проведенной с участием 24 экспертов по инициативе ВЦИОМ 11 сентября 2020 г.

На основе результатов сессии и серии интервью с экспертами нами проанализированы изменения на рынке прикладных социологических исследований, ожидаемые в ближайшие 3—10 лет, а также окна возможностей, открывающиеся для развития участников складывающейся исследовательской экосистемы. Делается вывод, что в ближайшие годы будет расти число исследовательских компаний, перешедших на бизнес-модели, основанные на обеспечении быстрого доступа клиентов к уникальному пользовательскому опыту, а также на встраивании исследовательских процессов во взаимодействие потребителей с продуктами клиентов за счет использования интернета вещей. Форсированное внедрение систем искусственного интеллекта повысит востребованность концептуализации и качественных исследований. Все более важным фактором успеха на рынке станет умение создавать уникальную ценность для клиентов за счет координации обме-

This article aims to analyze technological, economic, socio-political, and cultural factors of change in the research industry and the characteristics of the emerging ecosystem of the applied research market. We base our analysis on the results of the foresight session “Research Industry: Trends and Windows of Opportunity”, organized by the Russian Public Opinion Research Center (VCIOM) on September 11, 2020.

We examined the changes in the applied sociological research market that are expected in the next 3-10 years, and the windows of opportunities opening for the development of participants in the emerging research ecosystem. We conclude that in the coming years, the growing number of research companies would switch to such business models that are based on providing customers with quick access to a unique user experience, and on embedding research processes in the interaction of consumers with customers' products using the Internet of Things. The accelerated introduction of artificial intelligence systems will increase the demand for conceptualization and qualitative research. An increasingly important factor for market success will be creating unique value for customers by coordinating the exchange of resources between different participants in the emerging ecosystem of the research market.

на ресурсами различных участников складывающейся экосистемы исследовательского рынка.

Ключевые слова: форсайт, будущее исследовательской индустрии, исследовательская экосистема, автоматизация, искусственный интеллект

Keywords: foresight, future of research market, research ecosystem, automation, artificial intelligence

Введение

Изменения в проблематике и методологии социологических и маркетинговых исследований, а также в отношениях между исследователями и обществом все чаще становятся предметом обсуждения в научной литературе [Дулина, Звоновский, Токарев, 2017; Au, 2018; Chen et al., 2019; Schwemmer, Wieczorek, 2020; Горшков, 2020; Булычева, Мальцева, 2020], а также в рамках отраслевой аналитики и рефлексии, которую сами социологи и маркетологи ведут в профессиональных сообществах¹.

Все более актуальной задачей становится определение возможных последствий внедрения цифровых технологий в исследовательские процессы, в том числе с использованием систем искусственного интеллекта².

Чтобы осмыслить стремительные изменения на рынке прикладных социологических и маркетинговых исследований и наметить перспективы его развития, ВЦИОМ совместно с партнерами приступил к проведению серии интервью и круглых столов в рамках форсайта «Будущее исследовательской индустрии». Ее первым мероприятием 11 сентября 2020 г. стала форсайт-сессия «Исследовательская индустрия: тренды и окна возможностей», посвященная изменениям на рынке прикладных социологических исследований, ожидаемым в ближайшие 3—10 лет, а также окнам возможностей, открывающимся для развития участников складывающейся исследовательской экосистемы.

Методология форсайт-сессии

При разработке сессии были использованы принципы инновационно-технологического форсайта, которые хорошо иллюстрирует «ромб форсайта» [Popper, 2008]: в программу подготовки к форсайт-сессии, а также в ее сценарий были включены технологии, которые обеспечивают 1) доказательность

¹ См. Задорин И. В. Итоги и перспективы российского рынка прикладной социологии. Интервью Д. Мальцевой для портала Социолог.ру. 2016. 12 декабря. URL: http://www.sociologos.ru/statii/analitika_otrosli/Itogi_i_perspektivy_rossijskogo_rynka_prikladnoj_sociologii (дата обращения: 20.02.2021); Состояние рынка социологических и маркетинговых исследований в 2019. URL: <https://researchweek.ru/report2019.pdf> (дата обращения: 19.02.2021); GRIT Report: Greenbook Research Industry Trends Report. Insights Practice Edition 2020. N.Y.: AMA Communication Services Inc, 2020. URL: <https://www.greenbook.org/mr/grit/insights-practice-edition/> (дата обращения: 19.02.2021); Market Research Global Trends, 2021. URL: <https://www.qualtrics.com/ebooks-guides/market-research-trends> (дата обращения: 20.02.2021) и Poynter R. Predictions for the Next Five Years // NewMR. 2021. 24 January. URL: <https://newmr.org/blog/the-next-five-years/> (дата обращения: 19.02.2021).

² The State of AI in Market Research, 2021. N.Y.: Remesh, 2021. URL: <https://hello.remesh.ai/the-state-of-ai-in-market-research-2021> (дата обращения: 19.02.2021).

(обзорно-аналитические материалы, результаты других форсайт-исследований); 2) опираются на экспертизу участников (серия интервью, экспертный опрос, картирование трендов и анализ их последствий); 3) креативные методы, позволяющие придумывать, а не только предсказывать будущее (проективная технология метафорического прототипирования будущего отрасли, техника «Вещь из будущего» и др.); а также 4) технологии, основанные на взаимодействии между участниками (структурированные мозговые штурмы в подгруппах с использованием различных фасилитационных техник, общегрупповые дискуссии, сопоставление результатов работы участников, работавших в онлайн-режиме и очно в зале). При подборе техник групповой работы мы учитывали социально-психологические механизмы конструирования коллективного образа будущего [Нестик, 2018]. Руководствуясь принципами интегрального форсайта [Slaughter, 2008], мы использовали не только традиционные ракурсы анализа (учет политических, экономических, социокультурных, демографических, технологических и экологических изменений), но и соотношение коллективного и индивидуального, изменения на уровне личности, а также межличностных, внутригрупповых, межгрупповых отношений, организаций и общества в целом.

Еще одним принципом, который лег в основу совместной работы, был подход к обсуждению будущего исследовательской отрасли с позиций современной социологии и антропологии будущего [Lösch et al., 2019; Bryant, Knight, 2019; Beckert, Suckert, 2020 и др.]: во-первых, мы стремились рассматривать воображаемое будущее как ключ к пониманию социальных процессов настоящего; во-вторых, мы учитывали множественность конкурирующих друг с другом образов будущего, а также их включенность в экономику ожиданий и связь с экономическими и политическими интересами различных участников рынка.

При отборе экспертов были использованы несколько основных критериев:

1. Исследовательская компания: руководство федеральной или региональной исследовательской группой, агентством, консалтинговой компанией (или их подразделениями, проектами), входящими в топ-25 «Рейтинга российских исследовательских компаний» («РИК»).

2. Опыт инноваций в исследовательской индустрии: участие в корпоративных проектах и стартапах, основанных на использовании новых технологий прикладных исследований, в том числе Big Data (например, трекинга геолокации, анализа банковских транзакций, поисковых запросов), мобильных опросов (in-app), ботов-интервьюеров, Agile UX, дизайн-мышления, DIY- исследования и т. п. Опыт работы в аналитических подразделениях компаний из сферы финансовых технологий, интернета и телекоммуникаций.

3. Роль заказчика прикладных исследований: руководство подразделениями российских и зарубежных организаций, регулярно заказывающих маркетинговые исследования или исследования общественного мнения.

4. Сфера исследовательских интересов: участие в форсайтах и конференциях по данной тематике; доклады на отраслевых конференциях, посвященные будущему российского общества, будущему технологий, будущему исследовательской отрасли, а также будущему смежных с ней областей — прикладной психологии и коучинга, маркетинга, стратегического менеджмента, консалтинга и др.

5. Социальная активность: участие в органах управления профессиональными и научными ассоциациями в области социальных наук, опыт организации отраслевых конференций и других публичных мероприятий, посвященных настоящему и будущему исследовательской отрасли.

6. Кономинация: эксперт должен быть рекомендован не менее чем двумя ранее отобранными экспертами.

Таким образом, в сессии приняли участие представители международных и федеральных исследовательских компаний, академического сообщества, фрилансеров, аналитических подразделений интернет-компаний и банков, региональных исследовательских компаний, а также компаний — заказчиков исследований. Сессия проходила одновременно в очном и виртуальном режимах, что позволило включиться в обсуждение будущего удаленным участникам из регионов и Нью-Йорка.

В ходе подготовки сессии и анализа ее результатов были использованы несколько источников данных.

1. Опрос представителей индустрии накануне сессии (о трендах как внутри отрасли, так и о внешних трендах, влияющих на ситуацию в ней). Онлайн-опрос был проведен среди участников международной Грушинской социологической конференции (собрано 47 заполненных анкет с открытыми вопросами, база рассылки составила 1396 человек).

2. Экспертные интервью накануне сессии с руководителями ведущих исследовательских компаний о ключевых трендах в исследовательской индустрии (9 интервью).

3. Опрос представителей Ассоциации исследовательских компаний «Группа 7/89» о трендах развития регионального рынка маркетинговых и социологических исследований, тенденциях и трендах, влияющих на бизнес региональных компаний и отрасль в целом, и о прогнозе последствий их влияния (15 представителей компаний — участников Ассоциации³).

Глобальные и российские тренды, влияющие на развитие отрасли

Эксперты проанализировали влияние политических, экономических, социальных, демографических, культурных, технологических и экологических изменений на трансформацию исследовательского рынка.

Среди **политических и социальных трендов**, формирующих контекст развития отрасли, рассматривались прежде всего те процессы, которые стали реакцией общества на растущую неопределенность будущего.

Так, снижение доверия к социальным институтам (в том числе правительству, СМИ, бизнесу, НКО) повлекло за собой снижение доверия к науке и экспертам, а также усилило недоверие к результатам социологических исследований. Прекращение непрерывного экономического роста в развитых странах и очевидная неэффективность государства даже в развитых демократиях вызвали рост социального пессимизма. Все меньше тех, кто считает, что «наши дети будут жить лучше нас».

³ Благодарим за помощь в проведении опроса Анну Булгакову, директора ООО «ДВМЦ „Мониторинг“» (г. Владивосток).

Формирование экономики интернет-платформ подстегнуло рост численности «прекариата», то есть людей, которые не имеют постоянной занятости, стабильного заработка и социальных гарантий. В качестве целевой аудитории исследований все чаще выступают люди, меняющие работу, а также курьеры, таксисты, фрилансеры. Для взаимодействия с ними потребуются новые «языки» исследования, будет востребована экспертиза исследователя в работе с маленькими, постоянно меняющимися выборками, с глубоким погружением в практики их поведения, поиске инсайтов и встраивании их в продукт клиента.

Рост озабоченности социальной несправедливостью сопровождается повышением нетерпимости и поляризации в социальных медиа (эффекты «эхо-камеры»: идеи распространяются в закрытой среде «френдов», адресаты которой соглашаются друг с другом и слушают сами себя).

Ответом на растущую неопределенность и переживание трудноконтролируемой угрозы стали популизм во внешней политике и война санкций (затрудненный доступ к технологиям, сертификационные барьеры, усложненный выход на внешние рынки); антиглобализм (протекционизм, замедление или остановка интеграционных процессов в международных отношениях); консервативная волна (усиление консервативных и авторитарных ценностей). В ряде стран обсуждается переход к базовому социальному доходу, но происходит это на фоне краха общественно-го договора, сложившегося после Второй мировой войны: государство больше не в состоянии обеспечить защиту для всех перед лицом «больших вызовов».

В этих условиях продолжается поиск новых ролевых моделей успеха и «нarrативов надежды», основанных не на экономическом процветании, а на просоциальном поведении. Мы наблюдаем рост значимости экологических ценностей: озабоченность экологическими рисками, солидаризация людей для защиты собственного экологического благополучия. Сегодня символическая связь с условной Гретой Тунберг через тысячи километров приводит к реальным последствиям и для политики, и для власти. Растет значимость социальной ответственности и устойчивого развития, что неизбежно будет отражаться и на характере прикладных социологических исследований. С одной стороны, государственные структуры станут чаще заказывать исследования, чтобы выстраивать диалог с различными сообществами. С другой стороны, будет увеличиваться спрос бизнеса на репутационные исследования, а также социальный инжиниринг, исследовательскую поддержку социальных инноваций.

Среди глобальных **социокультурных и социально-демографических трендов**, влияющих на прикладную социологию, рассматривались прежде всего изменения, связанные с влиянием цифровых технологий на общество, изменением ценностей перед лицом глобальных рисков и «демографическим переходом». Приход поколений «Z» и «Альфа», отличающихся от своих родителей по системе ценностей и способам взаимодействия, усиливает значимость защиты персональных данных, сдвиг к визуальной культуре. Исследовательские агентства в ближайшие годы вынуждены будут искать новые способы рекрутования респондентов, больше опираться на визуальный язык при взаимодействии с ними и при представлении данных. Формирование «серебряной экономики» (старение населения и увеличение продолжительности активной жизни) и позднее взросление («отложенная»

социализация молодежи) повышают интерес заказчиков к исследованиям кредитного, инвестиционного и страхового поведения, поиску новых маркетинговых решений для этих возрастных категорий.

Пандемия COVID-19 подстегнула рост социального неравенства и новые его виды (социально-экономическое, культурное, цифровое и др.), медикализацию массового сознания (повышение ценности здоровья, внимания к гигиене и своему состоянию), переход к бесконтактной экономике (развитие фриланса и удаленной работы, переход к бесконтактной торговле и онлайн-коммерции). Форсированная цифровизация повседневных социальных практик в условиях пандемии ускоряет наметившийся уже переход к CATI, онлайн-опросам и бесконтактным способам сбора данных, отказ от поквартирных опросов лицом к лицу.

С одной стороны, формируется культура квантификации «я» (например, с помощью датчиков носимых устройств и интернет-сервисов, облегчающих анализ собственного поведения и сравнение себя с другими, таких как часы Apple и wellbeing-приложения). Благодаря «цифровым следам» и социальным сетям для человека расширяются возможности сравнения себя с другими, причем по все большему спектру характеристик. Жизненный опыт одних людей становится тут же доступным другим, что подстегивает формирование «префигуративной культуры» (представители младших поколений учатся друг у друга, а не у старших), а также институт потребительских отзывов, экономику обмена впечатлениями. Происходит переход к обучению в реальном времени (переход от тренингов к микрообучению через мобильные приложения, видеоблоги, геймификацию и т. п., настройка обучения 24/7 на решаемые человеком задачи). Мы наблюдаем взрыв авторства, когда большое количество людей получили инструменты для самовыражения, создания и доставки контента. Благодаря социальным сетям закрепляется массовая «культура уникальности»: приоритет единичного над массовым, поиск и публичное отстаивание исключительности мест, знакомых людей, продуктов и услуг, жизненного опыта, событий.

С другой стороны, реакцией на эту растущую цифровую публичность и прозрачность становится «деприватизация» жизни: отказ от самораскрытия, неготовность делиться переживаниями с исследователями. Происходит так называемая «цифровая де- и ресоциализация», когда все больше пользователей пересматривают свои социальные связи в сетях, отказываются от подписок на группы, устав от информационного шума. Стремление россиян защитить свои персональные данные приведет к еще большей монополизации доступа к ним. Данные все труднее получить через прямое обращение к интернет-пользователям, но они будут накапливаться у телекоммуникационных и интернет-компаний. В условиях надзирающего капитализма цифровые следы используются для контроля и манипулирования. Поэтому частные компании и государственные организации захотят «держать поближе к себе» не только большие данные, но любые другие сведения о пользователях, полученные тем или иным способом.

В рамках форсайта эксперты выделили несколько тесно связанных между собой **социокультурных процессов**, имеющих особенно серьезные последствия для исследовательской индустрии. Все они так или иначе связаны с ускорением социальных процессов. Возникает феномен «текущих идентичностей» — быстро

формирующими и распадающими солидарностей на основе событий, вкусов, отношения к проблеме. Все это уже сегодня требует сокращения сроков проведения исследований. Растет фрагментация общества на субкультуры. Эта мозаичность все больше ограничивает эффективность массовой рекламы, сокращая маркетинговые бюджеты и спрос на услуги исследовательских холдингов и полстеров.

После глобальной рецессии, связанной с пандемией, изменится стратификация российского общества. Но традиционная стратификация по социальнодемографическим характеристикам уже сейчас перестает работать. Основаниями для сегментации становятся меняющиеся, требующие постоянной перепроверки границы воображаемых сообществ. Растет запрос на новые методики сегментации, пересматриваются принципы построения выборки в маркетинговых исследованиях, ожидаются изменения в подходах к формированию выборок при изучении общественного мнения. Ускорение социальных процессов заставит заказчиков маркетинговых исследований переходить к еще более тесному контакту со своими потребителями, а также больше внимания уделять согласованию интересов разных сообществ, включенных в сферу их бизнеса.

Среди глобальных трендов наиболее существенное влияние на исследовательскую индустрию в ближайшие годы будет оказывать не только глобальная рецессия в результате пандемии COVID-19, но и ряд других **экономических и технологических изменений**, наметившихся в последние годы. Прежде всего это переход к экосистемам, экономике платформ: компании-платформы, которые связывают производителей и потребителей, как, например, Uber и Airbnb, видоизменяют конкуренцию: от повышения потребительской ценности к максимизации ценности экосистемы, то есть к координации обмена ресурсами ее членов. Происходит переход от обладания к доступности, то есть надежность и доступность услуги становятся важнее обладания ею. Распространение бизнес-моделей, основанных на продукте как услуге, превращает любой, даже производственный бизнес в консалтинговый. Бизнес становится более интеллектуальным, топ-менеджмент стремится принимать решения, опираясь на данные собственных in-house экспериментов, корпоративных, а не чужих исследований.

Развитие экосистем, переход от обладания к доступности и сформировавшаяся благодаря этим трендам экономика обмена имеют и другие последствия. Уже сегодня можно наблюдать рост заинтересованности клиентов в колаборационных проектах и общекатегорийных исследованиях. В результате можно ожидать интеграции заказчиков исследовательских услуг, а также более согласованных, артикулированных, отчетливых требований, которые заказчики из одной отрасли будут предъявлять к исследователям.

Сначала появление интернета, а теперь развитие индустрии 4.0 привели к сокращению дистанции между производством и потребителем. Иными словами, благодаря автоматизации сокращается число посредников. Это заставляет компании постоянно интересоваться своими конечными потребителями, вовлекать их в создание своих продуктов. В результате растет стремление компаний включить исследования в свои бизнес-процессы, центр генерации исследовательских данных смещается внутрь компаний.

Это повышает требовательность к качеству наполнения маркетплейсов данных, возрастает значимость умения формулировать гипотезы. Внутри больших бизнесов возникают маленькие стартапы, которые будут требовать от исследователей той же скорости, с которой работают сами. Иными словами, заказчики будут ожидать быстрой проверки гипотез по низкой стоимости.

Другим следствием тех же самых изменений становится распространение сервисов *DIY* (do it yourself — «сделай сам»). Разработка и настройка продуктов от специалистов переходит к «чайникам», расширяется доступность технологий для любителей, снижается авторитет традиционных экспертов. Частью этого тренда становится развитие исследовательских сервисов *DIY*, позволяющих любому подписчику самостоятельно формировать выборку, рекрутировать респондентов, анализировать в «облаке» и визуализировать данные с применением передовых технологий.

В ближайшие пять лет будет происходить все более глубокая автоматизация исследовательских процессов. Внедрение стандарта связи 5G многократно расширит возможности сбора данных через интернет «умных вещей». Это позволит использовать исследовательские технологии, пока не получившие широкого распространения, такие как айтреинг, дистанционная диагностика состояния клиента по психофизиологическим маркерам, а также технологии дополненной реальности. Развитие систем искусственного интеллекта позволит анализировать и агрегировать данные, полученные на основании опросов, анализа текстов, видео и различных «цифровых следов».

Отраслевые тренды и контуры новой экосистемы исследовательского рынка

В ходе форсайта были выделены продуктивные противоречия, от разрешения которых будут зависеть сценарии развития отрасли в ближайшие годы:

1) противоречие между растущими возможностями анализа цифровых следов и стремлением людей ограничить доступ к своим персональным данным;

2) противоречие между запросом на скорость проведения исследований и потребностью в их персонализации, настройке на уникального клиента;

3) противоречие между автоматизацией процессов, в том числе за счет использования искусственного интеллекта, и спросом на консалтинг, необходимостью концептуализации, творческого осмысления.

Эксперты отмечают целый ряд **отраслевых трендов**, проявляющихся в снижении спроса на исследования, росте недоверия к исследователям, новых требованиях к исследовательскому продукту, изменениях в методах и технологиях, а также в изменениях на рынке труда.

Снижение спроса на исследования связано с несколькими обстоятельствами. Во-первых, конкурентами исследовательских компаний в борьбе за заказчика все чаще становятся неисследовательские структуры, в том числе телекоммуникационные и интернет-компании. Во-вторых, все больше клиентов стремятся проводить исследования самостоятельно, без привлечения профессиональных агентств. Для этого используются альтернативные способы получения информации о потребителе, в том числе в режиме реального времени при взаимодействии

с продуктом, через носимые устройства и интернет вещей. На фоне сокращения бюджетов в условиях пандемии падает маржинальность в отрасли и усиливается конкуренция. Сказывается также и негативный институциональный фон для исследований: при инициации проектов и представлении бизнес-плана не требуются результаты исследования. Кроме того, мы по-прежнему наблюдаем рост недоверия к результатам социологических исследований, связанный с девальвацией статуса науки в России и снижением доверия к экспертурному знанию в целом.

Сами заказчики все еще нечувствительны к различиям в методологии и технологиях, которые используют исследовательские агентства. Они готовы работать как с крупными, так и с мелкими узкоспециализированными компаниями. Статус, география, «размер» исследователя в ближайшие годы не будут иметь существенного значения. Тем не менее желание заказчика четко идентифицировать исполнителей по их специализации будет расти.

Изменяются требования к продукту, прежде всего — к *визуализации результатов исследований и возможности их использования для принятия решений*. Востребованными становятся комплексные продукты, которые включают в себя не только само исследование, но и PR, и консалтинговое сопровождение. Сбор информации становится менее ценным, растет значимость интеллектуальной составляющей, интерпретации.

Меняются методы и технологии (см. рис. 1). Растет популярность смешанных методов, востребованность гибридных методик и стратегий исследования. Усиливается междисциплинарность в интерпретации, все более важным для заказчиков становится умение исследователей корректно согласовывать данные, полученные разными методами (опросы, качественные исследования, экспертные интервью, статистика, Big Data и др.). За счет автоматизации, использования ботов-интервьюеров и онлайн-панелей ускоряется процесс сбора данных.

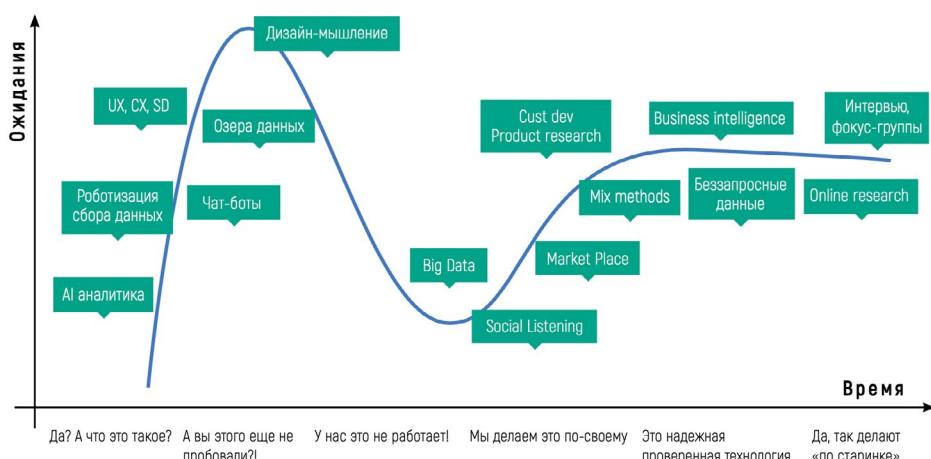


Рис. 1. Исследовательские технологии на кривой ажиотажа

Индустрія продолжит дрейф в сторону неслучайных выборок, при использовании которых нет однозначных правил обобщения, генерализации данных. Это заставит заказчиков ориентироваться в исследованиях прежде всего на собственные базы клиентов.

Рост доступности больших данных обостряет потребность в их интерпретации, понимании установок и мотивов поведения людей, поэтому в ближайшие годы будет расти интерес к качественным исследованиям. Уже сейчас расширяется использование материалов этносоциологических исследований в социальной и политической практике, развиваются немассовые методы — например, глубинные интервью с потребителями.

Продолжится вытеснение контактных методов сбора информации бесконтактными (наблюдение за потребителем, анализ предпочтений через носимые девайсы, контент-анализ соцсетей). Будут шире использоваться мобильные исследования и онлайн-интервью, а онлайн-опросы взамен личных и телефонных станут восприниматься как норма даже государственными органами.

На рынке труда обострится кадровый голод: произойдет переток специалистов из исследовательских агентств на сторону заказчика, в состав управлеченческих команд, к центру генерации данных. Будет происходить интеграция ролей исследователя, программиста, специалиста по большим данным и консультанта, что приведет к изменению требований к компетенциям при найме. На рынок будут приходить специалисты без профильного исследовательского образования, придерживающиеся разных профессиональных стандартов. В этой ситуации инструментом стандартизации и кузницей новых кадров для исследовательской индустрии станут корпоративные университеты. Высшая школа не сможет угнаться за скоростью изменений в инструментарии и организации исследований, что со временем обострит другую проблему — дефицит профессионалов, способных объяснить и научно обосновать используемую ими методологию.

Под влиянием глобальных и российских трендов меняется **экосистема исследовательского рынка** (см. рис. 2). Вызовами для исследовательской индустрии в ближайшие годы станут разнонаправленные процессы. С одной стороны, происходит монополизация и сокращение доступности больших данных, собираемых крупными телеком- и интернет-компаниями в рамках своей экосистемы. С другой стороны, идет развитие горизонтальных связей, когда специализированные исследовательские компании дополняют экспертизу друг друга.

Сегодня исследователь имеет огромное поле интерпретации. В неопределенности получаемых данных он может сделать любые выводы, любые рекомендации, и во многом сакрализирует сферу между реальностью и тем, что рассказывает заказчику. Он достаточно универсален, может одновременно проводить и исследования в социальных сетях, и традиционные, и онлайн-опросы, и стратегии писать, и стратсессии проводить. Маг-универсал, который может интерпретировать что хочешь, и делать что хочешь. А вокруг него в качестве индейцев, аборигенов присутствуют заказчики, которым он вполне может «влепить» любое зеркало, получив в обмен какое-то количество золотых монет.

В перспективе десяти лет в центре экосистемы появляется заказчик, которого обслуживают уже не маги, а солдаты с достаточно четким прорисованным функционалом, зонами компетенций, автоматизированными выводами. Он теряет монополию на сакральность данных, спектр этих вариаций неопределенности гораздо меньше. Заказчик теперь ищет узкую специализацию и узкую компетенцию. На рынке начинается серьезное разделение труда. Заказчик получает эти узкоспециализированные, рафинированные знания и внутри себя выстраивает интегрированную модель. (Кирилл Родин, ВЦИОМ)

Будет расти потребность в консультации специалистов разных профилей. Мы все чаще и чаще сталкиваемся с ситуацией, когда бывает проще, лучше, выгоднее, быстрее и полезнее для клиента и для исследователя привлечь специалиста из какой-то другой отрасли. (Владимир Звоновский, Фонд социальных исследований)

Все более важную роль на рынке социологических исследований будет играть государство. Оно будет одним из крупнейших заказчиков, а также регулятором доступа к данным. Дальнейшее развитие надзирающего капитализма, ускоренное пандемией и переходом к «бесконтактной» экономике, будет способствовать разделению рынка на внутри- и внесистемных игроков. При этом государство, банковские, телекоммуникационные и интернет-кампании станут основными поставщиками данных друг для друга. Иными словами, мы будем наблюдать централизацию и монополизацию в индустрии.

Государство возьмет под контроль крупные ИТ-компании, такие как Яндекс, Mail.ru, Ростелеком, Сбер, и сформирует мощную экосистему. В этой системе источники данных Big Data — чипы, мобильные приложения, госуслуги. На основе этих данных будут формироваться правительственные решения. Все живое будет фактически консолидировано государством под себя. Исследователи должны будут работать внутри структуры либо рядом с ней. Тем не менее вне этой системы останутся еще частные клиенты и исследователи-фрилансеры, например, специализирующие на качественных исследованиях. (Константин Абрамов, ВЦИОМ)

Заказчики переходят к изучению рынка через непрерывное экспериментирование. Такие эксперименты оказываются свободными от «методологического пуританства» прикладной социологии. Они привлекают в отрасль исследователей, не имеющих специального образования или, во всяком случае, не объединенных какими-то общими профессиональными, «цеховыми» стандартами. Сокращение дистанции между производителями и потребителями, переход к экономике платформ будут способствовать «уберизация» нишевых онлайн-панелей и услуг аналитиков. В России в ближайшие два года будет также формироваться рынок сервисов для самостоятельного проведения исследований.

Ландшафт исследовательского рынка и состав его участников будут меняться. Растет риск ухода с рынка компаний, которые используют устаревшие бизнес-модели. Полстеры при госзаказчике будут испытывать трудности, так как государство станет получать необходимые ему данные через телекоммуникационные и интернет-компании. Региональные полевые агентства будут вынуждены спе-

циализироваться или предоставлять комплексные решения, переходить на общенациональный и глобальный уровень. Московские исследовательские компании, ранее обращавшиеся к региональным полевикам для рекрутинга, теперь будут использовать другие каналы, прежде всего клиентские базы регионального и федерального бизнеса, в том числе телекоммуникационных корпораций. Рост доступности данных и автоматизация исследовательских процессов лишат преимущества универсальные исследовательские компании. Создаваемая ими ценность перестанет быть уникальной, им придется специализироваться или уйти с рынка. Их место займут новые игроки — такие как Яндекс, Mail.ru и другие агрегаторы данных, прежде всего телекоммуникационный бизнес и государство. Будет расти доля рынка у CX и UX агентств, так как они будут поддерживать процессы непрерывного экспериментирования внутри компаний-заказчиков. Сами заказчики тоже будут меняться: среди них заметную роль начнет играть малый бизнес, небольшие медиа. В целом глобальная рецессия, связанная с пандемией и ее последствиями, сократит бюджеты на исследования у всех компаний. Тем не менее можно ждать роста заказов от высокотехнологичных компаний, разработчиков систем искусственного интеллекта, государства, а также ритейла.



Отрасль через 5 лет

АКТУАЛЬНЫЕ БИЗНЕС-МОДЕЛИ ИГРОКОВ РЫНКА:



Рис. 2. Складывающаяся экосистема исследовательского рынка

Изменение бизнес-моделей и компетенции, необходимые для успеха на рынке

Изменение экосистемы приведет к изменению бизнес-моделей. С одной стороны, для типовых исследований по-прежнему будут важны высокая скорость и умеренная стоимость. Ценность, как и раньше, будет создаваться за счет понимания бизнеса клиента, умения решать практические задачи и представлять результаты в такой форме, чтобы на их основании сразу можно было строить план действий. Поэтому ценность качественной аналитики и консалтинга тоже не станет ниже. С другой стороны, на рынке будет происходить «миграция ценности», характерная для отраслей, границы которых смещаются под влиянием технологических изменений.

Ценность для клиентов будет создаваться не столько охватом и репрезентативностью выборок, сколько возможностями для персонализации — такого тестирования социально-экономических и продуктовых решений, которое не поддается автоматизации. В этих условиях будет происходить переориентация исследовательских компаний на бизнес-модели, в основе которых будет не только «поле».

Создание уникальной ценности будет все больше обеспечиваться за счет сочетания двух ключевых компонентов. Во-первых, быстрого доступа клиентов к пользовательскому опыту, новым идеям. Во-вторых, за счет встраивания исследовательских процессов во взаимодействие потребителей с продуктами клиентов. Решающую роль в этом процессе сыграет использование интернета вещей, возможности которого многократно возрастут при внедрении стандарта связи 5G. Уже в ближайшие два-три года многие клиенты будут ожидать от исследователей совмещения опросов с информацией о поведении человека в сети, анализом его «цифровых следов», в том числе данных с мобильных устройств, умных приборов и т. п.

В этой ситуации небольшие социологические «бутики» могут оказаться в более выигрышном положении, чем крупные агентства-универсалы. В условиях экономики платформ для доступа к данным уже не нужно будет выстраивать свою собственную онлайн-панель или систему анализа цифровых следов. Кооперируясь друг с другом, небольшие компании получат преимущество в гибкости и глубине экспертизы, которая будет труднее даваться крупным компаниям, привыкшим к большому портфелю проектов. Автоматизация значительной части исследовательских процессов и ориентация на непрерывную поддержку клиента в процессе принятия решений может привести к изменению подходов к управлению и способов монетизации: наряду с проектной организацией работ исследовательские компании все чаще будут использовать лицензирование клиентов и партнеров как источник прибыли.

Окна возможностей будут связаны с поиском и формированием альянсов, с углублением специализации исследовательских компаний и более четким разграничением зон экспертизы. Команды, работающие с разными инструментами, будут дополнять друг друга в решении задачи заказчика. Именно генерация решений, которые заказчик может применить в своей отрасли для своей управленческой задачи, станет основанием для пересборки исследовательского бизнеса. Скорость, с которой это нужно будет делать, уже не позволит специалистам перестраиваться на ходу и наращивать новые компетенции. Чтобы обеспечить

требуемый уровень экспертизы на каждом этапе исследовательского процесса, потребуются междисциплинарные команды, собранные из глубоко специализированных профессионалов.

Клиенты «отращивают» собственную экспертизу по многим вопросам, и в структуре компаний появляются подразделения с глубокими компетенциями в своей зоне. Они ожидают от поставщиков такой же глубины погружения в одну компетенцию. В результате на рынке должны появиться поставщики с четко очерченной сферой экспертизы. Более того, они должны обладать способностью объединяться для создания общего продукта, освобождая клиентов от координации их деятельности. В идеале это прозрачный понятный рынок, где есть четкое позиционирование поставщиков, которые взаимодействуют с отдельными функциями компаний. (Александр Кацуро, ВТБ)

При специализации на рынке компании будут выбирать одну из трех бизнес-моделей. Модель «Агрегатор экспертизы» будет основана на использовании сети партнеров, которая позволит под конкретный проект быстро собирать спектр глубоких экспертиз, направленных на решение конкретной задачи клиента. Модель «Отраслевой эксперт» выберут компании, которые хорошо (лучше, чем клиент) разбираются в специфике отрасли и предлагают решения, привлекая экспертизу со стороны. Ключевой ценностью в данном случае станет глубокая интеграция в процессы клиента. Наконец, востребованность модели «Вдыхателей смыслов» будет связана с ростом доступности данных и одновременно — с дефицитом концептуализации, необходимостью новых моделей, фреймворков. Вероятными клиентами в данном случае станут компании — агрегаторы больших данных, интернет-платформы, разработчики систем искусственного интеллекта, а также государственные структуры.

Возможны два варианта специализации. Первый — компании, которые глубоко погружены в отрасль, в рынок; они хорошо, даже лучше, чем сам заказчик, понимают, что ему нужно. Такие компании обладают глубокой экспертизой, знают конкурентов и целевую аудиторию, работают на один конкретный рынок. Второй вариант — исследователь под проект для решения задачи клиента собирает разных специалистов: стратегов, дизайнеров, отраслевых экспертов. Такие исследователи агрегируют очень глубокую экспертизу из разных сфер и создают дополнительную ценность для клиента. (Анастасия Черкашина, Cherkashina Research)

Ключом к успеху станет формирование так называемых Т-компетенций: совмещения глубокой специализации и способности формулировать задачи профессионалам из других отраслей. Во-первых, все более востребованными будут краткосрочные образовательные курсы («микрообучение»), расширяющие кругозор исследователей, а также различные инструменты развития социологического воображения. Во-вторых, потребуется использование специальных фасилитационных технологий групповой работы, которые облегчат взаимопонимание людям из разных базовых областей: консультантам по управлению, разработчикам искусственного интеллекта, филологам, антропологам и т. д.

Социологи и психологи все больше будут востребованы для интерпретации больших данных, чтобы дать возможность корпорациям и государству не только отслеживать, но и прогнозировать поведение людей. Еще одно окно возможностей открывается в связи с форсированным внедрением систем искусственного интеллекта, для обучения которых требуются социологические исследования и концептуализация, позволяющая конструировать модели поведения — не только использовать имеющиеся у клиента «фреймфорки», но и генерировать новые.

С увеличением спроса на работу с Big Data станет востребована новая компетенция — «вдыхатели жизни в большие данные», то есть специалисты интерпретации, социальные аналитики. До тех пор, пока к «кнопкам» не допустят специалиста по социальной аналитике, большие данные будут безжизненны. (Владимир Звоновский, Фонд социальных исследований)

Необходимость вовлечения респондентов в процессы совместной разработки продукта и даже в продвижение результатов исследований будет предъявлять новые требования к автоматизации взаимодействия исследовательских компаний с участниками своих исследований: такие системы, с одной стороны, будут в большей степени похожими на краудсорсинговые платформы, а с другой — будут более персонализированными.

Нарастание дифференциации различных продуктов, товаров, услуг, вероятно, будет вызывать потребность клиента во взаимодействии с очень-очень узкими целевыми аудиториями, практически в индивидуальной работе с клиентом и потребителем. Есть обратный процесс, то есть клиент также будет активно вовлекаться в процесс создания, проведения исследований, создания исследовательского продукта. Мы это знаем через многочисленные брейнштроминги, которые проводят исследователи-качественники, но, скорее всего, этот процесс будет расширяться и вглубь, и вширь. (Владимир Звоновский, Фонд социальных исследований)

Смещение границ отрасли и рост цифровой составляющей исследований приведет к появлению «продукт-менеджеров», которые будут давать задания разным участникам междисциплинарной команды. Они станут своего рода передаточным звеном между сторонами: высокотехнологичной стороной сбора данных и собственно исследовательской, консалтинговой. Наряду с менеджерами продукта могут появиться «исследовательские омбудсмены», обеспечивающие публичность результатов исследований, повышающие их социальную значимость для различных сообществ и возвращающие доверие к исследовательской индустрии. Чтобы отстаивать свою экспертизу, исследователи будут значительно шире использовать сторителлинг, а также решения по хранению и распространению исследовательской информации внутри компаний (поисковики по внутренним базам данных исследований, дашборды и т. п.). Возрастет роль управления знаниями, потребность в технологиях обмена опытом между исследователями.

Для того чтобы подготовиться к этим изменениям, исследовательские компании уже сегодня должны вступать в более тесный контакт с заказчиками, стараться

не только лучше понимать их бизнес-процессы, но и помогать им зарабатывать деньги.

Можно выделить несколько компетенций, которые станут более востребованными в ближайшие годы. Во-первых, умение профессионально самоидентифицироваться, выделить свою экспертизу и четко позиционировать себя для клиентов. Во-вторых, способность сделать исследовательские процессы прозрачными, понятными для клиента. В-третьих, умение не только быть постоянно доступным для клиента, но и глубоко погружаться в его специфику, интегрироваться с его бизнес-процессами. Одним словом, перейти от исследований к непрерывной поддержке принятия решений. Это потребует способности поддерживать тесные коммуникации с клиентом на всех этапах, вместе с ним продвигаться от брифа к совместному формированию гипотез и задач. В-четвертых, все более востребованным будет умение агрегировать глубокую экспертизу из разных сфер, опираясь при этом на принципы дизайн-мышления, непрерывно перепроверяя гипотезы. Для того, чтобы договариваться с игроками, обладающими нужным уровнем и профилем экспертизы, и создавать дополнительную ценность для клиента, исследователям потребуются навыки нетворкинга и фасилитации. В-пятых, универсальной компетенцией станет работа с большими данными, в том числе и «грязной», и «черной» датой. В-шестых, будет востребовано умение работать с кросс-инструментами, комбинировать их, создавая большую ценность для клиента. Наконец, важной компетенцией станет знание фреймворков, моделей из других смежных отраслей, умение их адаптировать для решения задачи клиента. В условиях автоматизации исследовательских процессов способность к концептуализации станет очевидным конкурентным преимуществом.

Необходимо проанализировать степень успешности и уникальности своих продуктов и выбрать какие-то из них как основу будущей глубокой специализации. Уже сейчас нужно вступать в альянсы со своими конкурентами, а также создавать новую ценность для клиентов, обеспечивая им доступ к опыту друг друга, вовлекать их в совместные исследовательские проекты. Выиграют те, кто будет увеличивать свою ценность для заказчика за счет координации обмена ресурсами различных участников складывающейся экосистемы исследовательского рынка.

Список литературы (References)

- Булычева Е. Е., Мальцева Д. В. Выделение актуальных тематик в социологии: взгляд сквозь призму анализа сети цитирований // Мониторинг общественного мнения: экономические и социальные перемены. 2020. № 6. С. 113—140. <https://doi.org/10.14515/monitoring.2020.6.971>.
- Bulycheva E. E., Maltseva D. V. (2020) Highlighting Key Topics in Sociology: A Glance Through the Prism of Citation Network Analysis. *Monitoring of Public Opinion: Economic and Social Changes Journal*. No. 6. P. 113—140. <https://doi.org/10.14515/monitoring.2020.6.971>. (In Russ.)
- Горшков М. К. Будущее социологического знания и вызовы социальных трансформаций (к 90-летию В. А. Ядова) // Гуманитарий Юга России. 2020. Т. 9. № 1. С. 15—21. <https://doi.org/10.19181/2227-8656.2020.1.1>.

Gorshkov M. (2020) The Future of Sociological Knowledge and Challenges of Social Transformations (to the 90th Anniversary of V.A. Yadov). *Humanities of the South of Russia*. Vol. 9. No. 1. P. 15—21. <https://doi.org/10.19181/2227-8656.2020.1.1>. (In Russ.)

Дулина Н. В., Звоновский В. Б., Токарев В. В. Рынок социологических и маркетинговых исследований в России: состояние, динамика, прогноз // Социологические исследования. 2017. № 12. С. 110—123. <https://doi.org/10.7868/s0132162517120121>.

Dulina N. V., Zvonovskiy V. B., Tokarev V. V. (2017) Sociological and Market Research Industry in Russia: Position, Dynamic, Prognosis. *Sociological Studies*. No. 12. P. 110—123. <https://doi.org/10.7868/s0132162517120121>. (In Russ.)

Нестик Т. А. Психологические аспекты корпоративного форсайта // Форсайт. 2018. Т. 12. № 2. С. 78—90. <https://doi.org/10.17323/2500-2597.2018.2.78.90>.

Nestik T. A. (2018) The Psychological Aspects of Corporate Foresight. *Foresight and STI Governance*. Vol. 12. No. 2. P. 78—90. <https://doi.org/10.17323/2500-2597.2018.2.78.90>.

Au A. (2018) Sociology and Science: The Making of a Social Scientific Method. *The American Sociologist*. Vol. 49. P. 98—115. <https://doi.org/10.1007/s12108-017-9348-y>.

Beckert J., Suckert L. (2020) The Future as a Social Fact. The Analysis of Perceptions of the Future in Sociology. *Poetics*. 101499. <https://doi.org/10.1016/j.poetic.2020.101499>.

Bryant R., Knight D. M. *The Anthropology of the Future*. Cambridge: Cambridge University Press, 2019.

Chen J., Wu C., Cai S., Torabi M. (2019) Contemporary Theory and Practice of Survey Sampling: A Celebration of Research Contributions of J. N. K. Rao. *International Statistical Review*. Vol. 87. No. S1. P. S1-S2. <https://doi.org/10.1111/insr.12312>.

Lösch A., Grunwald A., Meister M., Schulz-Schaeffer I. (2019) Introduction: Socio-Technical Futures Shaping the Present. In: Lösch A., Grunwald A., Meister M., Schulz-Schaeffer I. (eds) *Socio-Technical Futures Shaping the Present*. Technikzukünfte, Wissenschaft und Gesellschaft/Futures of Technology, Science and Society. Wiesbaden: Springer VS. P. 1—14. https://doi.org/10.1007/978-3-658-27155-8_1.

Popper R. (2008) Foresight Methodology. In: Georghiou L., Harper J. C., Ceenan M., Miles I., Popper R. *The Handbook of Technology Foresight. Concepts and Practice*. Cheltenham: Edward Elgar Publishing. P. 44—88.

Schwemmer C., Wieczorek O. (2020) The Methodological Divide of Sociology: Evidence from Two Decades of Journal Publications. *Sociology*. Vol. 54. No. 1. P. 3—21. <https://doi.org/10.1177/0038038519853146>.

Slaughter R. A. (2008) What Difference Does “Integral” Make. *Futures*. V. 40. No. 2. P. 120—137. <https://doi.org/10.1016/J.FUTURES.2007.11.015>.

Urry J. (2016) *What is the Future?* Cambridge: Polity.

DOI: [10.14515/monitoring.2021.1.1900](https://doi.org/10.14515/monitoring.2021.1.1900)



А. А. Иванова

ВЗАИМООТНОШЕНИЯ «ЧЕЛОВЕК — АЛГОРИТМ — ЧЕЛОВЕК» КАК СОЦИАЛЬНАЯ ПРОБЛЕМА И ПРОБЛЕМА ДЛЯ СОЦИОЛОГОВ.

**РЕЦ. НА КН.: CATHY O'NEIL. WEAPONS OF MATH DESTRUCTION:
HOW BIG DATA INCREASES INEQUALITY AND THREATENS
DEMOCRACY. NEW YORK: CROWN PUBLISHERS, 2016**

Правильная ссылка на статью:

Иванова А. А. Взаимоотношения «человек — алгоритм — человек» как социальная проблема и проблема для социологов // Мониторинг общественного мнения: экономические и социальные перемены. 2021. № 1. С. 387—397. <https://doi.org/10.14515/monitoring.2021.1.1900>. Рец. на кн.: Cathy O'Neil. Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. New York: Crown Publishers, 2016.

For citation:

Ivanova A. A. (2021) Human-Algorithm-Human Relationship as a Social Problem and a Sociological Problem. *Monitoring of Public Opinion: Economic and Social Changes*. No. 1. P. 387–397. <https://doi.org/10.14515/monitoring.2021.1.1900>. Book Review: Cathy O'Neil “Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy”. New York: Crown Publishers, 2016. (In Russ.)

ВЗАИМООТНОШЕНИЯ «ЧЕЛОВЕК — АЛГОРИТМ — ЧЕЛОВЕК» КАК СОЦИАЛЬНАЯ ПРОБЛЕМА И ПРОБЛЕМА ДЛЯ СОЦИОЛОГОВ. РЕЦ. НА КН.: CATHY O'NEIL. WEAPONS OF MATH DESTRUCTION: HOW BIG DATA INCREASES INEQUALITY AND THREATENS DEMOCRACY. NEW YORK: CROWN PUBLISHERS, 2016

ИВАНОВА Анастасия Андреевна — аспирант факультета социологии, Санкт-Петербургский государственный университет, Санкт-Петербург, Россия
E-MAIL: anaspis100@gmail.com
<https://orcid.org/0000-0003-4296-8530>

Аннотация. Сегодня исследования взаимоотношений «человек — алгоритм — человек» вызывают все возрастающий интерес среди представителей разных дисциплин. Рецензируемая книга вносит вклад в эту область и предлагает критический взгляд на нерефлексивное внедрение алгоритмов в жизнь общества. Предлагая оригинальный термин «оружие математического поражения» (*weapons of math destruction*, ОМП), автор книги Кэти О'Нил формулирует основной тезис: в современном американском обществе именно такой тип моделей оказывает все более пагубное влияние на жизнь людей. Книга посвящена анализу проблем, возникающих с проникновением ОМП в различные сферы: от выбора университета до приема на работу, от оценки здоровья сотрудников до получения кредита, от организации работы полиции до выдачи новостей в социальных сетях. Некоторые из описанных алгоритмов подпадают под определение искусственного интеллекта, некоторые могут потенциально им стать при дальнейшем развитии технологий.

HUMAN-ALGORITHM-HUMAN RELATIONSHIP AS A SOCIAL PROBLEM AND A SOCIOLOGICAL PROBLEM. BOOK REVIEW: CATHY O'NEIL “WEAPONS OF MATH DESTRUCTION: HOW BIG DATA INCREASES INEQUALITY AND THREATENS DEMOCRACY”. NEW YORK: CROWN PUBLISHERS, 2016

Anastasia A. IVANOVA¹ — Post-Graduate Student, Faculty of Sociology
E-MAIL: anaspis100@gmail.com
<https://orcid.org/0000-0003-4296-8530>

¹ Saint Petersburg State University, Saint Petersburg, Russia

Abstract. Today the human-algorithm-human relationship is becoming a focal point for research in numerous disciplines. This book is a contribution to the topic and provides a critical perspective on non-reflexive introduction of algorithms into social life. Using “WMD” (*weapons of math destruction*) as a specific term, Cathy O’Neil sets forward her main thesis stating that this very type of models has an increasingly negative impact on the America society. The book analyses the problems related to the “WMD” penetration into various spheres of life ranging from university choices to employment, from health evaluation to loans, from police performance to news content on social media. Some algorithms fit the definition of artificial intelligence; some may become one with further development of technologies. In the first part of the review the author considers the theses of the book and the arguments proving these ideas. The second part of the review addresses the following issues: Why would Russian readers be interested in this book? Would it be of interest to social scientists? In what way would the review’s author be inter-

В первой части рецензии автор описывает основные тезисы, изложенные в книге, и аргументацию, подкрепляющую эти утверждения. Во второй части содержание книги рассматривается через попытку ответить на три вопроса: почему книга будет интересна российскому читателю? чем она интересна социальному ученому? чем исследования социальных ученых могли бы быть интересны автору рецензируемой монографии? Автор приходит к выводу, что понятие «оружие математического поражения» обладает значительной эвристической ценностью для социальных наук.

Благодарность. Исследование выполнено при финансовой поддержке РФФИ и Министерства по науке и технологиям Тайваня в рамках научного проекта № 21-511-52002.

ested in the social scientists' studies? The author concludes that the weapons of math destruction have considerable heuristic value for social science.

Acknowledgments. The study was supported by RFBR and the Ministry of Science and Technology (Taiwan), project No. 21-511-52002.

Развитие современной медицины предполагает внедрение технологий искусственного интеллекта (ИИ): эта тенденция универсальна [Резаев, Трегубова, 2019]. Страновая специфика заключается в том, что агенты ИИ внедряются с разной скоростью и в разных направлениях. Вместе с тем разработчики ИИ и потенциальные пользователи сталкиваются с многочисленными проблемами, которые ограничивают разработку, внедрение и устойчивый рост применения ИИ в области медицины и здравоохранения.

Много проблем связано с тем, что в литературе называют пристрастностью алгоритмов (*algorithmic bias*). Одно из обоснований внедрения агентов ИИ в медицину состоит в том, что от них ожидают объективности, которая понимается как способность судить в соответствии с универсальными критериями. Другое обоснование заключается в указании на эффективность, на способность достигать цель с минимальными затратами. Однако если алгоритм по некоторым причинам отдает предпочтение одним пациентам перед другими, то его внедрение не только не усилит объективность и эффективность, но и приведет к значительным ухудшениям в работе медицинских организаций. И в сочетании со слепым доверием к алгоритмам такая ситуация может сохраняться достаточно долго.

Рецензируемая книга посвящена анализу именно такого рода проблем, которые возникают с проникновением алгоритмов в повседневную жизнь людей. Книга содержит многочисленные примеры из различных сфер: от образования до здравоохранения, от кредитования до организации работы полиции. Некоторые

из описанных алгоритмов подпадают под определение ИИ, другие могут потенциально им стать при дальнейшем развитии технологий.

Многие социальные ученые опасаются математики. С одной стороны, они выражают беспокойство возможностью просчитывания их действий вездесущими алгоритмами в интересах корпораций, государства или иных лиц. Из этих опасений и из специальной терминологии, возникающей вокруг них, возникают новые темы и новые области исследований. С другой стороны, в социальных науках и гуманитарных дисциплинах «боятся» математики потому, что она зачастую представляется чем-то непонятным и сложным. Мы анализируем социальные условия создания и внедрения алгоритмов, рассматриваем, каким образом работа некоторого устройства зависит от того, как понимают и воспринимают его разные люди, выясняем, как действия одних технологий способствуют или препятствуют действию других. Однако сам алгоритм обычно остается для исследователей «черным ящиком», в который лишь немногие имеют желание (и смелость) заглянуть¹.

Иными словами, опасливое отношение социальных ученых к математике, стоящей за алгоритмами, является и стимулом, и препятствием к изучению этих самых алгоритмов.

Что на это могли бы ответить математики? Книга Кэти О'Нил «Оружие математического поражения: как большие данные усиливают неравенство и угрожают демократии» [O'Neil, 2016]² может быть прочитана как ответ на подобный вопрос, способный усилить одни страхи и развеять другие. Она представляет собой исследование того, как разные виды человеческой глупости в сочетании с ошибками алгоритмов приводят к возникновению новых феноменов, которых мы как люди должны опасаться и при этом бесстрашно анализировать — как ученые.

Автор книги — бывший университетский математик, бывший аналитик в нескольких финансовых организациях, в настоящее время — писатель, автор блога³ и основательница консалтинговой компании по внедрению алгоритмов⁴. В своей монографии Кэти О'Нил рассматривает проблемы, интересующие современных социальных ученых. Однако она выходит на эти проблемы с других позиций — опираясь на свой опыт работы в финансовых корпорациях и стартапах, знание математики, а также на повседневный опыт жизни в американском обществе. Из-за разницы в перспективах книга получается одновременно и скромнее, и точнее, и по-своему амбициознее многоного из того, что сегодня публикуют социальные ученые.

Ключевой тезис книги может быть сформулирован следующим образом: в современном американском обществе все более пагубное влияние на жизнь людей оказывает «оружие математического поражения» / ОМП (*weapons of math*

¹ В современных социальных науках возникает целое новое направление исследования алгоритмов, в рамках которого авторы по-разному формулируют, что именно они собираются исследовать, а что оставляют за рамками анализа. Подробнее см. [Lee, Larsen, 2019].

² Автор отсылает читателя к понятию Weapons of Mass Destruction — «оружие массового поражения». В 2018 г. книга была переведена на русский язык под названием «Убийственные большие данные. Как математика превратилась в оружие массового поражения» (М.: Издательство ACT).

³ Mathbabe. Exploring and Venting About Quantitative Issues. URL: <https://mathbabe.org/> (дата обращения: 22.02.2021).

⁴ O'Neil Risk Consulting & Algorithmic Auditing. URL: <https://orcarisk.com/> (дата обращения: 22.02.2021).

destruction, WMD). Следуя за аргументацией автора, следует рассмотреть а) саму идею ОМП, б) различные примеры работы ОМП (или того, что может им стать), в) оценку влияния ОМП на жизнь людей.

Ключевое словосочетание для понимания того, что такое ОМП,— плохая модель. В современном мире используются многочисленные математические модели самых разных процессов, более или менее точные и более или менее полезные (вредные). В качестве примера хороших моделей автор приводит те, что используют бейсбольные команды для подготовки игроков и определения состава команды. Такие модели основаны на понятных («прозрачных») показателях, которые измеряют именно то, что призваны измерить,— различные аспекты игры. Информация об этих показателях накапливается и постоянно обновляется. Существует обратная связь: прогноз об игре, выданный моделью, сравнивается с действительным исходом игры, что позволяет корректировать модель. Наконец, такие модели приносят пользу и командам, и болельщикам, улучшая качество игры.

ОМП представляют собой противоположность по всем пунктам, кроме одного. Эти модели 1) также основаны на большом количестве данных, но при этом 2) не прозрачны: те, чью активность измеряют, не знают или не понимают, что именно измеряют и почему; 3) основаны на косвенных показателях, редко измеряют то, что призваны измерять; 4) не имеют обратной связи, которая позволяла бы корректировать исходную модель. Как следствие, ОМП наносят вред тем, чью деятельность они прогнозируют, хотя при этом могут приносить прибыль компаниям или поддержку политикам. Таким образом, ОМП — это идеальный тип (в веберовском смысле) математической модели, на основании которой люди или алгоритмы принимают решения: в реальности модели могут обладать лишь некоторыми из четырех свойств, и вред моделей, согласно логике автора, зависит от того, в какой мере они этими свойствами обладают.

В книге рассматривается много интересных кейсов ОМП и социальных механизмов, стоящих за их внедрением, в различных ситуациях: от выбора университета до приема на работу, от получения кредита до вынесения судебного приговора, от организации работы полиции до выдачи новостей в социальных сетях. Приведем лишь один пример, с которого начинает автор,— систему оценки деятельности учителей в школах США.

Программа по оценке работы школьных учителей направлена, в частности, на сокращение сотрудников, набравших меньше всего баллов. Кэти О'Нил рассказывает историю молодой учительницы, которая была уволена в соответствии с результатами оценки, несмотря на то, что и ученики, и другие учителя высоко оценивали ее работу. После своего увольнения учительница, имея хорошие рекомендации, устроилась в гораздо более престижную школу в другом штате. Еще один случай, о котором повествует автор,— история опытного учителя, имеющего постоянный контракт. В один год педагог получил 6 баллов из 100, а в следующий—96. Если бы у этого учителя не было постоянного контракта, администрация школы была бы вынуждена его уволить после первого тестирования.

Чем объясняется очевидная «глупость» алгоритмов, оценивающих учителей? Именно теми свойствами, которые позволяют отнести данную модель к ОМП.

Модель оценки учителей основана на большом объеме данных, но она непрозрачна. Сами учителя не знают точно, что именно измеряется и какие показатели имеют какой вес. При этом модель основана на косвенных показателях — в частности, использует такие упрощенные показатели, как ежегодные тесты успеваемости учеников, на которые влияют самые разные факторы и которые, к тому же, поддаются манипулированию. Что даже более важно — модель не имеет механизмов корректировки. Данные о том, что один и тот же педагог получает кардинально различающиеся баллы в разные годы, или о том, что уволенная в соответствии с рекомендацией модели сотрудница успешно устраивается на новую работу, никак не учитываются моделью. Как следствие, возникает самосбывающееся пророчество: те, кто классифицируется алгоритмами как «негодный», остаются в этой категории; более того — алгоритмы влияют на жизнь людей, они способны значительно осложнить ее, подтверждая тем самым собственный «приговор»⁵.

Следующий вопрос состоит в том, почему столь неэффективная система продолжает работать. Ответ стоит искать в том, для кого именно система неэффективна. Для учителей — безусловно, для учеников, их родителей и школы как сообщества — тоже. Однако для административного аппарата такая система вполне эффективна: она позволяет провести сокращение персонала, опираясь на внешне «объективные» и стандартизованные критерии. Иными словами, ОМП действуют в интересах тех, кто их внедряет, и эти интересы, как правило, формализуются: в числе «оптимизированных» кадров, в доле раскрытых преступлений, в сумме прибыли. Содержательные же характеристики, важные для тех, чью деятельность моделируют ОМП, остаются за рамками рассмотрения тех, кто принимает решения.

Критика ОМП в монографии сосредотачивается (как это видно из названия) на проблеме социального неравенства и на угрозах демократии. Под «угрозой демократии» понимается то обстоятельство, что одни люди принимают решения, а другие сталкиваются с последствиями этих решений. Посредником между первыми и вторыми выступает ОМП, причем такой посредник блокирует обратную связь, которая могла бы повлиять на решения финансовых и административных элит.

Это особенно проблематично, если учитывать второй аспект — усиление неравенства: ОМП по-разному действует в отношении представителей разных классов. Согласно аргументации автора, действия ОМП поддерживают благосостояние высшего класса и усиливают у него иллюзию собственных достижений, при этом консервируют или поддерживают бедность и «размывают» средний класс. Это происходит по нескольким причинам, наиболее важная из которых — соотношение между данными на «ходе» и «выходе» модели. ОМП дают рекомендации об индивидах, исходя из прогнозов, построенных на разделении людей

⁵ Это не так очевидно в примере с учителями, где рекомендации коллег все еще имеют вес. Однако монография содержит многочисленные примеры, где это становится ясно. Более суровый приговор суда, основанный на решении алгоритма (которое, например, может быть основано на наличии преступников среди знакомых обвиняемого), приводит к меньшим возможностям устроиться на хорошую работу после выхода из тюрьмы. Низкооплачиваемая работа приводит к худшим условиям в получении кредита и страховки, исходя из рисков, рассчитываемых математическими моделями. Значительные траты на кредит и страховку при низких доходах вынуждают снимать жилье в непрестижных районах, которые, в соответствии с рекомендациями алгоритмов, усиленно патрулируются полицией — что увеличивает шанс привлечения к суду даже за незначительные правонарушения. Круг замыкается, и это лишь одна из схем, описываемых автором.

на категории. Например, при расчете стоимости страховки автомобиля модель опирается на предположение, что люди с хорошей кредитной историей будут более ответственными водителями. В результате принадлежность к категории «обладатель хорошей кредитной истории» оказывается более значимой при расчете страховки, чем эпизоды вождения в нетрезвом виде в личной истории водителя. Кроме того, важно, что различные виды ОМП усиливают друг друга: расчет кредитного рейтинга учитывается при расчете страховки и при сортировке резюме при приеме на работу, поиск низкооплачиваемой работы в интернете стимулирует выдачу рекламы «хищных» вузов и микрокредитования и т. п. Как следствие, те, кто попал в категорию «бедных» и «неперспективных», с большой вероятностью останутся в ней, каким бы ни было их личное поведение. Модели «видят» их не как отдельных людей, а как представителей определенной категории, в итоге люди и алгоритмы, принимающие решения, обращаются с ними соответствующим образом.

В заключение монографии автор призывает к действию: аналитиков данных — не допускать создания плохих моделей, законодателей — создавать новые законы о защите личной информации, широкую общественность — обсуждать опасности распространения ОМП, оказывая тем самым давление на элиты. Математические модели способны принести столько же пользы, сколько и вреда, а вред этот может оказаться гораздо более значительным, чем сейчас.

Рассуждая о содержании данной книги, поставим три вопроса:

- Почему книга будет интересна российскому читателю?
- Чем книга интересна социальному ученому?
- Чем исследования социальных ученых могли бы быть интересны автору книги?

На первый вопрос ответить просто. Во-первых, хотя автор рассматривает процессы, происходящие в американском обществе, те явления, которые она называет «оружием математического поражения», возникают и распространяются в разных странах мира, и Россия не исключение. Университетские рейтинги, расчет кредитной истории, персонализированная «хищная» реклама в интернете — все это знакомо и российскому читателю.

Во-вторых, в монографии описываются особенности американского общества, которые задают критерии для сравнительного анализа. Как различается нормативное регулирование использования личной информации в системах с разными типами права? Что происходит, когда ОМП используются там, где ведущая роль принадлежит не бизнесу, а государству (например, в случае системы социального кредита в Китае)? Чем воспроизводство расового неравенства отличается от воспроизведения этнических иерархий (как в случае мигрантов на территории бывшего СССР)? Список вопросов может быть продолжен. Иными словами, взгляд на американское общество поможет увидеть общее и отличное в том, какие формы может принимать ОМП в России.

Наконец, исследование «хороших» и «плохих» моделей прогнозирования и принятия решений на их основе, безусловно, актуально в период пандемии COVID-19. Регулирование поведения людей, внедрение различных мер той или иной степени обязательности принимаются в значительной степени на основе моделей, кото-

рые разрабатывают ученые⁶. И здесь знание того, какими свойствами обладает хорошая модель, становится ключевым для понимания ситуации.

На второй вопрос — чем книга будет интересна социальному ученому — также легко ответить. Основное достоинство книги — в сочетании насыщенного эмпирического материала с ключевой идеей ОМП, вокруг которой этот материал организован. Кроме того, особый интерес представляет то, как автор прослеживает возникновение интереса к проблеме ОМП в связи с поворотами в собственной профессиональной карьере на фоне финансового кризиса.

Богатство эмпирических примеров и размышление над биографией автора мы оставляем читателю, а здесь прокомментируем ключевую идею книги. Само понятие «оружие математического поражения» обладает значительной эвристической ценностью для социальных наук. Прежде всего, оно указывает на то, что модели могут быть «хорошими» и «плохими». Даже основанные на сложном математическом моделировании алгоритмы работают неэффективно, если математический аппарат плохо «подсоединен» к реальности. Становится ясно, на какие именно аспекты модели нужно смотреть в первую очередь, какие вопросы задавать. Имеется ли возможность скорректировать модель на новых данных или она превращается в самосбывающееся пророчество? На основании каких — относящихся к кому — данных делается прогноз об индивидуальном поведении? Существует ли механизм отслеживания ошибок при сборе данных? И каковы последствия неверного прогноза для тех, кто моделирует, и для тех, кого моделируют? Наконец, в концепцию ОМП входит положение о том, что модель изменяет поведение людей, которых она моделирует, и это изменение должно быть учтено теми, кто оценивает работу модели⁷.

Еще одно важное достоинство книги заключается в ясной постановке проблемы. В том, как Кэти О'Нил характеризует ОМП, можно различить отголоски и публичной социологии Майкла Буравого, и концепции великой трансформации Карла Поланьи. С одной стороны, это старая идея — оградить гражданское общество от наступления рынка и государства в эпоху капитализма. С другой стороны, речь все-таки идет о принципиально новых феноменах, новых проблемах и новых угрозах. Пример, который приводит автор: в XX веке при приеме на работу алгоритмы исправляли предрассудки людей, а сейчас люди исправляют «предрассудки» алгоритмов. Чем глупость и несправедливость людей отличается от глупости моделей и несправедливости алгоритмов? Как соединить деятельность людей и алгоритмов так, чтобы они компенсировали, а не усиливали ошибки друг друга? Вот вопросы, ключевые для социальной науки XXI века, и рассматриваемая монография намечает пути их исследования.

Книга ставит проблему не только для позитивной науки, но и для этики. Автор показывает, что при построении моделей возникает выбор между анализом всей возможной информации и сохранением приватности, между повышением прибыли за счет категоризации клиентов и готовностью пойти на риск, чтобы дать шанс преуспеть конкретным людям, между частными и общественными интересами.

⁶ При этом результаты прогнозирования, выполненные на основании разных моделей, могут различаться кардинально. См., например: Как СПбГУ и Политехник второй волны высчитывали: почему их прогнозы отличаются в несколько раз // Фонтанка. 2020. 7 декабря. URL: <https://www.fontanka.ru/2020/12/07/69607966/>.

⁷ Здесь можно заметить перекличку с базовыми идеями кибернетики об обратной связи и о роли наблюдателя.

Можно не соглашаться с ответами, которые предлагает монография, но проблема «или-или» поставлена, и игнорировать ее нельзя.

Третий вопрос — чем исследования социальных ученых могут быть интересны автору книги — не столь тривиален. Легко заметить (как мы это уже сделали), что социальная критика автора основана на идее разделения между рынком, государством и гражданским обществом, представленной, в частности, в работах Карла Поланьи, а предложенный курс действий пересекается с программой публичной социологии. Однако эти идеи были сформулированы в прошлом веке и стали чем-то вроде «здравого смысла» для социальных ученых (с которым можно, но утомительно спорить). Более того, они не специфичны для анализа тех конфигураций моделей и алгоритмов, которые рассматривает Кэти О'Нил. Что может предложить социальная наука XXI века? Здесь следует выделить несколько направлений.

Во-первых, это исследования алгоритмов, которые существуют в основном в рамках STS и принимают разные теоретические и методологические установки [Lee, Larsen, 2019; Ziewitz, 2016]. Здесь интересен анализ отдельных кейсов взаимодействия людей и алгоритмов: любой из примеров рассматриваемой монографии мог бы стать предметом для подобного анализа. Иными словами, исследования алгоритмов способны продолжить эмпирический анализ, намеченный в книге.

Во-вторых, это анализ новых форм капитализма, связанных с распространением цифровых технологий. Здесь следует выделить концепции «надзорного капитализма» (surveillance capitalism) Шушаны Зубофф [Zuboff, 2019] и «капитализма платформ» Ника Срничека [Срничек, 2019]. Данные концепции помещают анализ Кэти О'Нил в более широкий социально-экономический контекст, где возникновение и распространение ОМП становится лишь одним из следствий трансформаций современной экономики.

В-третьих, более глубокий анализ изменений социальной структуры в связи с деятельностью сортирующих алгоритмов представлен в исследованиях ординализации Марион Фуркад [Fourkade, 2016; Fourkade, Healy, 2017]. Если анализ трансформаций капитализма позволяет ответить на вопрос, почему и для кого выгодно использовать ОМП, то концепция ординализации отвечает на вопрос, как использование ОМП соотносится со стратификацией современного общества, а также ценностями и конфликтами, лежащими в ее основании.

Наконец, следует выделить литературу по социальной теории, которая могла бы подкрепить подход к анализу деятельности людей и алгоритмов, представленный в рассматриваемой монографии. Кэти О'Нил рассматривает человека не как атомизированного индивида, а как находящегося в отношениях с другими людьми, причем эти отношения опосредуются, определяются и искажаются действием алгоритмов. Сама идея о человеке как совокупности отношений является ключевой для социальной науки, начиная с Маркса и заканчивая представителями «реляционной социологии» [Emirbayer, 1997; Gergen, 2009]. Вместе с тем для анализа того, что происходит с общественными отношениями в эпоху развития математических моделей и алгоритмов, перспективна концепция «искусственной социальности» [Резаев, Стариков, Трегубова, 2020].

В заключение вернемся к тому, с чего начинали,— к проблеме разделения между социальными учеными и математиками, «гуманитариями» и «технарями» или — используя другой понятийный ряд,— «лириками» и «физиками». В 1956 г. Ч. П. Сноу сформулировал проблему «двух культур» — разделения между технической и художественной интеллигенцией [Сноу, 1985]. С тех пор данная проблема многократно обсуждалась, само это различие оспаривалось и уточнялось, и, тем не менее, проблема «двух» культур продолжает возникать в различных формах и контекстах. Сегодня она явно наблюдается в разработках и исследованиях искусственного интеллекта и машинного обучения, в рамках которых и возникает ОМП: разработчики восхищаются своими изобретениями и превозносят их ценность для человечества, в то время как писатели, философы, социальные критики формулируют все новые претензии и страхи в отношении этих изобретений. Однако дискуссии зачастую не возникает: стороны остаются глухи к аргументам друг друга, и не в последнюю очередь потому, что плохо разбираются в том, о чем говорят их оппоненты. В этом отношении монография Кэти О'Нил представляет собой шаг вперед: автор понимает и техническую, и «человеческую» сторону того, о чем пишет, более того — способствует пониманию и приглашает к обсуждению проблемы все заинтересованные стороны. В том числе и социальных ученых.

Подводя итоги, следует безусловно рекомендовать к прочтению книгу «Оружие математического поражения: как большие данные усиливают неравенство и угрожают демократии». Монография обладает несомненными достоинствами: широкий охват, свежий взгляд на проблему, концептуальная ясность и увлекательный стиль, сочетание ясного анализа и горячей социальной критики делают книгу привлекательной для самых разных читателей. Некоторая категоричность автора в обсуждении этических вопросов, спорные интерпретации, а также отсутствие полноты в анализе отдельных кейсов можно рассматривать как приглашение к дискуссии и дальнейшим исследованиям. Эти особенности полностью компенсируются смелостью, с которой автор подводит читателей к рассмотрению новых проблем, касающихся нас и как рядовых граждан, и как исследователей.

Список литературы (References)

- Резаев А. В., Стариков В. С., Трегубова Н. Д. Социология в эпоху «искусственной социальности»: поиск новых оснований // Социологические исследования. 2020. № 2. С. 3—12. <https://doi.org/10.31857/S013216250008489-0>.
- Rezaev A. V., Starikov V. S., Tregubova N. D. (2020) Sociology in the Age of 'Artificial Sociality': Search of New Bases. *Sociological Studies*. No. 2. P. 3—12. <https://doi.org/10.31857/S013216250008489-0>. (In Russ.)
- Резаев А. В., Трегубова Н. Д. Искусственный интеллект и искусственная социальность: новые явления и проблемы для развития медицинских наук // Эпистемология и философия науки. 2019. Т. 56. № 4. С. 183—199. <https://doi.org/10.5840/eps201956475>.
- Rezaev A. V., Tregubova N. D. (2019) Artificial Intelligence and Artificial Sociality New Phenomena and Problems for Medical and Life Sciences Advance. *Epistemology*

& *Philosophy of Science*. Vol. 56. No. 4. P. 183—199. <https://doi.org/10.5840/eps201956475>. (In Russ.)

Сноу Ч. П. Две культуры и научная революция // Портреты и размышления. М.: Прогресс, 1985. С. 195—226.

Snow Ch. P. (1985) The Two Cultures and the Scientific Revolution. In: Snow Ch. P. *Portraits and Reflections*. Moscow: Progress. (In Russ.)

Срничек Н. Капитализм платформ / пер. с англ. и науч. ред. М. Добряковой. М.:Изд. дом Высшей школы экономики, 2019.

Srnicek N. (2019) Platform Capitalism. Moscow: The Higher School of Economics Publishing House. (In Russ.)

Emirbayer M. (1997) Manifesto for a Relational Sociology. *The American Journal of Sociology*. Vol. 103. No. 2. P. 281—317. <https://doi.org/10.1086/231209>.

Fourcade M. (2016) Ordinalization: Lewis A. Coser Memorial Award for Theoretical Agenda Setting 2014. *Sociological Theory*. Vol. 34. No. 3. P.175—195. <https://doi.org/10.1177/0735275116665876>.

Fourcade M., Healy K. (2017) Seeing Like a Market. *Socio-Economic Review*. Vol. 15. No. 1. P. 9—29. <https://doi.org/10.1093/ser/mww033>.

Gergen K. (2009) Relational Being: Beyond Self and Community. Oxford: Oxford University Press.

Lee F., Larsen L.B. (2019) How Should We Theorize Algorithms? Five Ideal Types in Analyzing Algorithmic Normativities. *Big Data & Society*. Vol. 6. No. 2. <https://doi.org/10.1177/2053951719867349>.

O’Neil C. (2016) Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. New York: Crown Publishers.

Ziewitz M. (2016) Governing Algorithms: Myth, Mess, and Methods. *Science, Technology, & Human Values*. Vol. 41. No. 1. P. 3—16. <https://doi.org/10.1177/0162243915608948>.

Zuboff S. (2019) The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power. New York, NY: PublicAffairs.

BOOK REVIEW

DOI: [10.14515/monitoring.2021.1.1907](https://doi.org/10.14515/monitoring.2021.1.1907)



D. M. Zhikharevich

HOW ALGORITHMS BECAME CAPITALIST. BOOK REVIEW OF: ZUBOFF SH. SURVEILLANCE CAPITALISM: THE FIGHT FOR A HUMAN FUTURE AT THE NEW FRONTIER OF POWER. NEW YORK: PUBLICAFFAIRS, 2019

For citation:

Zhikharevich D. M. (2021) How Algorithms Became Capitalist. *Monitoring of Public Opinion: Economic and Social Changes*. No. 1. P. 398–409. <https://doi.org/10.14515/monitoring.2021.1.1907>. Book Review of: Zuboff Sh. Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power. New York: PublicAffairs, 2019.

Правильная ссылка на статью:

Жихаревич Д. М. Как алгоритмы стали капиталистическими // Мониторинг общественно-го мнения: экономические и социальные перемены. 2021. № 1. С. 398—409. <https://doi.org/10.14515/monitoring.2021.1.1907>. Рец. на кн.: Zuboff Sh. Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power. New York: PublicAffairs, 2019. (In Eng.)

HOW ALGORITHMS BECAME CAPITALIST.
BOOK REVIEW OF: ZUBOFF SH. SURVEILLANCE CAPITALISM: THE FIGHT FOR A HUMAN FUTURE AT THE NEW FRONTIER OF POWER. NEW YORK: PUBLICAFFAIRS, 2019

Dmitrii M. ZHIKAREVICH^{1,2} — Research Fellow, Science and Technology Studies Centre (STS Centre); Researcher, International Research Laboratory TANDEM
E-MAIL: dzhikharevich@eu.spb.ru
<https://orcid.org/0000-0002-2518-7858>

¹ European University at Saint Petersburg,
Saint Petersburg, Russia

² Saint Petersburg State University, Saint Petersburg,
Russia

Acknowledgments. The study was supported by Russian Science Foundation: Grant No. 18-78-10049.

КАК АЛГОРИТМЫ СТАЛИ КАПИТАЛИСТИЧЕСКИМИ. РЕЦ. НА КН.: ZUBOFF SH. SURVEILLANCE CAPITALISM: THE FIGHT FOR A HUMAN FUTURE AT THE NEW FRONTIER OF POWER. NEW YORK: PUBLICAFFAIRS, 2019

ЖИХАРЕВИЧ Дмитрий Михайлович — научный сотрудник, центр исследований науки и технологий (центр STS), Европейский университет в Санкт-Петербурге, Санкт-Петербург, Россия; исследователь, международная исследовательская лаборатория ТАНДЕМ, Санкт-Петербургский государственный университет, Санкт-Петербург, Россия
E-MAIL: dzhikharevich@eu.spb.ru
<https://orcid.org/0000-0002-2518-7858>

Благодарность. Работа выполнена при поддержке Российского научного фонда: грант № 18-78-10049.

This book is monumental, and claiming that a short review can do it justice would be laughable. Hence, like an Annual Review author, I must cheat either breadth or depth [Abbott, 1993: 187], and I choose the former. Let me spare the reader the tedious part: I will not retell *Surveillance Capitalism* content chapter by chapter. Instead, I will focus on what I think is its most original aspect, namely, the theory of surveillance capitalism outlined in the first part of Zuboff's edifice.

So, what is surveillance capitalism? Zuboff starts with a dictionary-like definition which serves as an epigraph and lists as many as eight (!) related meanings, including 'a new economic order', 'a parasitic economic logic', 'a rogue mutation of capitalism', 'the foundational framework of a surveillance economy', and 'the origin of a new instrumentarian power', among other things [Zuboff, 2019]. Some of these initial formulations are further refined throughout the book's eighteen chapters. Yet, since all meaning is relational, such dictionary-like definitions are of little use unless adequately contextualised within the author's network of concepts. One of the definitions Zuboff provides halfway through the second chapter is a good place to start: surveillance capitalism is a 'new market form' and a 'unique logic of accumulation in which surveillance is a foundational mechanism in the transformation of investment into profit' [Zuboff, 2019: 55]. Note that surveillance does not enter the equation as merely a 'moment' of capital accumulation, as Marxists would (perhaps) have it; it is much more than an external appendage to the circuit of capital. Instead, it is the very foundation on which capital accumulation, that is, the transformation of investment

into profit, rests. As such, it is new and even unprecedented: according to Zuboff, surveillance capitalism 'is a new actor in history, both original and *sui generis*' [ibid.: 20]. The book documents its emergence within (Part 1), and diffusion beyond (Part 2), the 'networked sphere', as well as analyses the new form of 'instrumentarian' power it gave rise to (Part 3).

What is Surveillance Capitalism?

To be sure, the concept of 'surveillance capitalism' is not of Zuboff's coinage. Marxist political economists John Bellamy Foster and Robert McChesney used the term to describe, focusing on the U.S. case, the most recent amalgamation of monopoly-finance capital, the national security state, and digital technology on the pages of *Monthly Review* [Foster, McChesney, 2014]. Zuboff does address some of these issues, too, but from a different vantage point. Firstly, her preferred unit of analysis is a firm, like Apple or Google, considered simultaneously as the point of origin and the embodiment of new capitalism's economic imperatives. As different from the holistic perspective of political economists, Zuboff locates the essential dynamics at the firm level, where new logics of capital accumulation originate and from where they diffuse. Hardly any reviewer failed to quote at least one of these suggestive lines: 'Google is to surveillance capitalism what the Ford Motor Company and General Motors were to mass-production-based managerial capitalism' [Zuboff, 2019: 65]; 'Google invented and perfected surveillance capitalism in much the same way that a century ago General Motors invented and perfected managerial capitalism' [ibid.: 16]. In Zuboff's view, Google became 'the pioneer, discoverer, elaborator, experimenter, lead practitioner, role model and a diffusion hub' [ibid.: 65] of the new model of capitalism, later adopted by Facebook, Microsoft, and other firms, including those operating beyond the Internet sector.

Secondly, Zuboff's understanding of surveillance is much more specific and less literal than the one suggested by Foster and McChesney. For Zuboff, surveillance has less to do with CCTV¹ cameras on the high street feeding into a government-run (and occasionally leaking) face recognition system than with the cookie files stored in your browser. In her very first academic paper on surveillance capitalism, Zuboff makes clear that by 'surveillance' she means something like 'predict[ing] and modify[ing] human behaviour as a means to produce revenue and market control' [Zuboff, 2015: 75]. This network of concepts unfolds in the book, making their meaning visible: surveillance capitalism is a form of capitalism whose basic operation is the rendition of human experience into behavioural data to manufacture prediction products and sell them on behavioural futures markets. To be sure, companies do capture user data all the time, but, according to Zuboff, this does not amount to surveillance capitalism as long as behavioural value, that is, the insights into user behaviour derived from user data analytics, is reinvested to improve user experience. Zuboff calls this process 'behavioural value reinvestment cycle', whereby 'only behavioural data needed for service improvements are rendered' [Zuboff, 2019: 72]. Capitalism begins with a shift from the behavioural value reinvestment cycle to the hunt for behavioural surplus: 'More behavioural data are rendered than required for service improvements. This

¹ Closed Circuit Television.

surplus feeds machine intelligence — the new means of production — that fabricates predictions of user behaviour. These products are sold to business customers in new behavioural futures markets' [ibid.: 97]. Make no mistake, surveillance capitalists do not care about your UX², even though they usually say otherwise. Instead, it is the 'data exhaust' what they are after: additional data points, unnecessary from the point of view of UX improvement, but extremely valuable for predicting users' future behaviour. Zuboff calls these data points 'behavioural by-products' or 'collateral data', such as 'the number and pattern of search terms, how a query is phrased, spelling, punctuation, dwell times, click patterns, and location' [ibid.: 69]. Capturing, or, in Zuboff's preferred phrasing, 'extracting' this behavioural surplus, is key to making profits on behavioural futures markets, since it enhances the value of surveillance capitalists' prediction products.

Now, this is starting to look familiar. From Marx's general formula of capital Polanyi's embedded markets, we know that production for subsistence plus selling occasional surpluses is quite different from production for the purposes of selling. The discipline of market competition introduces new forms of dependency, or, in Zuboff's terms, new 'economic imperatives', and this changes everything. In the words of the great Marxist historian Ellen Meiksins Wood, the rise of capitalism, or the ultimate 'market society', precisely means this shift 'from opportunity to imperative', whereby market exchange ceases to be a voluntary option and becomes the only way to subsist [Wood, 1994]. Hence, if one drops the assumptions of liberal anthropology, suggesting that markets basically are spontaneous emanations of human nature and will appear anywhere unless blocked by some artificial constraints, it becomes necessary to explain the emergence of the market imperative, rather than merely positing it. Zuboff is up for the task, approaching it with a narrative of Google's discovery of behavioural surplus that prompted a shift from behavioural value reinvestment cycle to behavioural surplus extraction.

Larry Page and Sergey Brin invented the web search algorithm PageRank in 1996, but, despite the brilliance of technology, they struggled to find a way to sustainable revenues. Indeed, Google's early business model was based on providing web services to portals like Yahoo! through licensing, as well as sponsored ads linked to search keywords. Zuboff carefully documents the early years of the fledgling company, noting the founders' disgust towards advertising and their venture capitalist backers' confidence in the company's ability to find some way of monetising the algorithm. But then came the dot-com bubble, and Google's investors suddenly became seriously interested in returns on their money. Analysts sometimes call venture capital 'patient capital', but, as Zuboff shows, this was not the case at Google: scared of the bubble and its consequences, venture capitalists became impatient and declared a state of emergency. It is in this context that Google learned to valorise 'data exhaust' that used to be 'haphazardly stored and operationally ignored' [Zuboff, 2019: 69]. In Zuboff's telling, the company's transition from occasional advertising deals and licenses to surveillance revenues was accomplished when Amit Patel, the data scientist credited with 'the groundbreaking insight into the significance of Google's accidental data caches', met Eric Schmidt, the new CEO hired in 2002. Schmidt turned Patel's insight into a strategy,

² User Experience.

shaping Google's 'specific response to financial emergency', and opening up the way for the 'crucial mutation that ultimately turned... the Internet, and the very nature of information capitalism toward an astonishingly lucrative surveillance project' [ibid.: 77].

What is Wrong with Surveillance Capitalism?

Google's realisation that data exhaust is an untapped source of value is the turning point not only for surveillance capitalism but also for Zuboff's narrative. Suddenly a business school professor talking about selected cases from recent corporate history becomes a social theorist well-versed in the critical sociology playbook. Zuboff shows how competitive pressures push surveillance capitalism beyond the Internet and into the real world, as Google, Facebook, and their likes aggressively acquire startups developing technologies for the Internet of things, thus setting in motion the digitisation of the physical. On the other hand, some businesses that are not digitally native, like Verizon and other telecom companies, are quickly picking up some of the profit-making techniques of the pristine surveillance capitalists. As a result, surveillance is increasingly becoming omnipresent with every new 'smart' device — a phone, a watch, or a home — entering our everyday lives. For Zuboff, this means that surveillance capitalists are now also able to modify human behaviour by prompting, or rather nudging users to engage with their services more intensively, thus producing more data points to be fed into predictive analytics and adding even more value to the prediction products manufactured by the owners of the means of behaviour modification. This point allows Zuboff to argue that human experience itself now joins the Polanyian triumvirate of 'fictitious commodities', becoming 'subjugated to surveillance capitalism's market mechanisms and reborn as "behaviour"' [Zuboff, 2019: 100]. Behaviour is then turned into data from which predictions are fabricated to be exchanged in the new behavioural futures markets.

In this story, Google and Facebook feature as the agents of surveillance capitalism's primitive accumulation, claiming human experience as a free good to be commodified, thus threatening to inflict as much damage to human nature as industrial capitalism did to nature tout court. Zuboff goes on, arguing, with Arendt [2004], that primitive accumulation, in fact, never stops, thus becoming accumulation by 'digital dispossession' [Zuboff, 2019: 98—100]. Here Zuboff draws on David Harvey's [2003] famous concept, arguing that 'surveillance assets', the by-products of users' search actions, are claimed at no cost. Surveillance capitalists rely on the classic set of capitalist tactics to create the 'first-mover advantage' for themselves. They strive to occupy unregulated territories of the digital world as fast as possible, set a precedent, and then push the regulators back with the help of the network effects, expensive lawyers, and revolving doors between private companies and the policy circles. And yet, it is not capitalism per se that worries Zuboff the most. She dismisses the standard liberal critiques of the 'platform economy' as monopolistic, playing the Arendtian tune and appealing to the unprecedented character of surveillance capitalism as a form of power, not just monopoly power: 'Even when surveillance capitalist operations are also monopolistic and a threat to privacy, the existing categories nevertheless fall short in identifying and contesting the most crucial and unprecedented facts of this new regime' [Zuboff, 2019: 20].

So, what is wrong with surveillance capitalism? Basically, it is surveillance. Firstly, Zuboff is worried about privacy, or, more precisely, about the rights to privacy. These rights enable one to choose whether to share something or keep it secret; deprived of privacy rights, one loses an important decision right as well. Surveillance capitalism triggers a massive redistribution of these rights, whereby individuals are increasingly deprived of their choices ‘in the matter of what about their lives remains secret’, and their privacy rights accrue to surveillance capitalists [Zuboff, 2015: 82—83]. Like other capitalists, the latter take what they need without asking by planting surveillance technologies into unregulated domains of everyday life and then mount sophisticated defenses of what has already been taken. Ironically, these defenses often involve appeals by surveillance capitalists to *their* rights to privacy, for example, in the form of legal protections derived from the trade secrets law.

Secondly, unequal distribution of the rights to privacy goes hand in hand with the increasingly unequal distribution of knowledge in society. The commodification of human experience by turning it into ‘behaviour’, as if taken from the pages of B. F. Skinner’s writings [see, e.g., Skinner, 1976], operates through a form of indifference: the products of surveillance capitalism ‘manage to be derived from our behaviour while remaining indifferent to our behaviour. Its products are about predicting us, without actually caring what we do or what is done to us’ [Zuboff, 2019: 71]. This indifference brings Zuboff to her larger concern: the distortion of what she calls the ‘social division of learning’, a concept derived from Durkheim’s *Division of Labour*. Here Zuboff draws on her own earlier studies of workplace automation. Back in the 1980s, she argued that, as different from industrial automation, the advent of ‘computer-mediated work’ did not simply replace human labour but added a new layer of complexity to the entirety of organisational life, rendering ‘events, objects, processes, and people...visible, knowable, and shareable in a new way. The world is reborn as data and the electronic text is universal in scale and scope’ [Zuboff, 2015: 77]. Information technology’s capacity not only to automate but also to ‘informate’, therefore, is pregnant with an emancipatory possibility of making work ‘smart’ and turning learning into a new form of labour [Zuboff, 1988: 395; quoted in Zuboff, 2015: 76]. Yet whether this possibility will become a reality is contingent on the balance of power within and across organisations. Surveillance capitalism realises the bad scenario, the darkening of the digital dream, whereby the massive electronic text produced by computer-mediated social life is turned into an engine for private profits.

Consequently, the privatisation of the division of learning in society, ‘the critical axis of social order in the 21st century’, changes the relationships between knowledge and power in dramatic ways. It is interesting to note that Foucault is barely mentioned on the pages of Zuboff’s book, even though some formulations will ring the relevant bells with the readers who had been through a French theory class. For Zuboff, however, the division of learning in society seems to refer to an unlikely combination of Durkheim’s ‘division of labour in society’ and Hayek’s ‘use of knowledge in society’. Indeed, as surveillance capitalism increasingly becomes what Zuboff calls ‘reality business’, migrating from the Internet into the ‘real’ world, its economic imperatives become salient for the new hybrid of the digital and the physical created by the Internet-of-Things (IoT), Augmented Reality (AR), and other similar technologies. Surveillance

capitalists' ultimate commodity is the certainty of behavioural predictions, and as such, it works against human freedom premised upon the inherent indeterminacy of human action. Here Zuboff draws both on Hayek's ideas about the distributed character of social knowledge and Arendt's theory of human action as defined by its power to create new things. Surveillance capitalism's quest for certainty embodied in predictive algorithms and 'smart' contracts work to eliminate this basic human quality, thereby also threatening to erode the social itself because, in the world of total certainty, there is no place for trust-based relationships and even for most of the legal institutions, starting from the simple contract. By gaining control over the division of learning in society, surveillance capitalists become not only rich but also immensely powerful while remaining completely unaccountable. Zuboff calls their power 'instrumentarian' and reserves the nickname 'Big Other' for the infrastructure of computer mediation that underpins it. By taking our data from us, surveillance capitalists also take our right to a sanctuary: there is no place to hide from the Big Other's decentralised surveillance. As different from much of contemporary writing in the subfield of 'surveillance studies', it is not Foucault or Deleuze (with their respective notions of disciplinary vs. control society), but Orwell and Arendt that Zuboff talks to.

What is Wrong with Zuboff's Surveillance Capitalism?

So far, so good — or, rather, so bad. From the normative point of view, Zuboff's critique calls for the new modes of collective action capable of resisting surveillance capitalism, a Polanyian double movement for the digital age [see Polanyi, 1957]. Surveillance capitalism is undesirable, but there is nothing inevitable about it. Zuboff insists that the quest for 'behavioural surplus appropriation' does not result from some technological necessity driving the development of Big Data; instead, this turn of events results from historical contingency. Surveillance capitalism 'was invented by a specific group of human beings in a specific time and place' [Zuboff, 2019: 85], namely, at Google in the aftermath of the dot-com bubble. However, this invention was made possible by some larger-scale social and economic forces: not just the dot-com crisis per se, but also the spread of neoliberalism and financialisation of the corporate world. Combined, the latter two trends stabilised the idea that the only thing business corporations should care about is shareholder value maximisation, thus setting the stage for surveillance capitalism's distinctly dismissive attitude to accountability and responsibility to anybody but its investors. Therefore, surveillance capitalism is a product of specific historical circumstances that could have been different; it is undesirable but mutable. To drive the point home and accomplish the critical gesture, Zuboff also needs to show the direction of change, a normative project. It is here where she betrays some of the problematic assumptions of her theory.

It turns out that the crucial problem with surveillance capitalism is the lack of fundamental 'reciprocities' between capitalists and their 'populations'. What makes surveillance capitalism a particularly 'rogue' mutation of capitalism is that its users and customers do not overlap. Users who like Facebook pages and search the web with Google do improve the algorithms by allowing the companies to capture their data, but at the end of the day it is the advertisers who buy the 'prediction products' manufactured from these data. In Zuboff's [2019: 17] words,

Surveillance capitalism's products and services are not the objects of a value exchange. They do not establish constructive producer-consumer reciprocities. Instead, they are the 'hooks' that lure users into their extractive operations in which our personal experiences are scraped and packaged as the means to others' ends. We are not surveillance capitalism's 'customers'. Although the saying tells us 'If it's free, then you are the product', that is also incorrect. We are the sources of surveillance capitalism's crucial surplus: the objects of a technologically advanced and increasingly inescapable raw-material-extraction operation. Surveillance capitalism's actual customers are the enterprises that trade in its markets for future behaviour.

Indeed, being a customer and a source of the raw materials supply is not the same. In the former case, what Zuboff calls 'producer-consumer reciprocity' obtains, and the consumers have the levers to hold the firm accountable by exercising one of Albert O. Hirschman's [1970] three options of exit, voice, and loyalty. Conversely, no such reciprocity is possible between the firm and its raw materials; this relationship is not one of 'value exchange', but of extraction pure and simple. Thus, surveillance capitalism deviates from the historical relationship between markets and democracy, threatening to become the digital version of 'a pre-modern absolutist authority' that evades public scrutiny or 'the traditional market pressures of consumer reciprocity and choice' [Zuboff, 2015: 83]. UX improvements are just a way of 'cornering' new sources of raw materials supply, not the means of beating the competition. Instead, surveillance capitalist competition is about better prediction products and advertising revenues flowing from their sales.

But wait, something does not add up here. Should not surveillance capitalists' structural independence from their users work both ways? In other words, what locks the users in the surveillance economy in the first place? If there is no market exchange between surveillance capitalist companies and their users, to paraphrase Wood [1994], there is neither 'opportunity', nor 'imperative'. Since the users' subsistence does not depend on the surveillance economy (defined, somewhat tautologically, as 'a surveillance-based economic order'), what is so tragic about it that Zuboff even feels compelled to invoke the Faustian compact? The answer lies in Zuboff's theoretical assumptions rooted not in critical sociology but, somewhat surprisingly, in modernisation theory. Following Durkheim's argument that modern division of labour does not only unlock productivity but creates new 'conditions of existence', triggering the shift from mechanic to organic solidarity, Zuboff suggests that it is the new 'conditions of existence that create and sustain demand for surveillance capitalism's services' [Zuboff, 2019: 100]. To theorise these conditions, she draws on the theory of reflexive modernisation outlined by Ulrich Beck, Scott Lash, Anthony Giddens, and Ronald Inglehart, explicitly invoking one of its central concepts, the 'second modernity' [see, e.g., Beck, Bonss, Lau, 2003]. For Zuboff, the most important aspect of modernisation is progressive individualisation, whereby the individual becomes 'the locus of moral agency and choice' [Zuboff, 2019: 37]. Having emerged from the mass-production world of the first modernity, individuals now require individualised consumption, unmediated by standardised distribution channels. Ultimately, the driving force of all individual activity, consumption included, is the need to live an 'effective life' under the

given conditions of existence that evolve in modernisation. Very much in tune with the turn-of-the-century mood of reflexive modernisation theory, Zuboff sees consumption as a part of individual lifestyle projects, thus adding an existential dimension to it. Historically changing conditions of existence connect supply and demand and trigger the emergence of new market forms, which Zuboff equates with forms of capitalism. New market forms emerge in specific historical and existential circumstances; some of them die out, while others manage to catch on, get adopted, and eventually institutionalised in the producer-consumer reciprocities that are tragically lacking in surveillance capitalism [Zuboff, 2015: 77].

Put simply, surveillance capitalism is a ‘bad’ way of responding to the new conditions of existence that individuals face on the verge of the third modernity. In his insightful review of Zuboff’s book, Evgeny Morozov [2019] notes that her theory is a strange hybrid of functionalist business history à la Alfred Chandler and somewhat naïve appreciation of the wonders of the digital age characteristic of Italian post-workerist thought, making it a Harvard Business School version of Toni Negri. This description is accurate, as evidenced by the way Zuboff discusses capitalism. Recall that, in her view, surveillance capitalism is a ‘rogue mutation’ of capitalism, implying that there is also a proper, or, as she prefers to call it, ‘rational’ capitalism. Already in the Introduction, Zuboff insists that it is necessary to distinguish between capitalism and surveillance capitalism: ‘When a firm collects behavioural data with permission and solely as a means to product or service improvement, it is committing capitalism but not surveillance capitalism’ [Zuboff, 2019: 28]. Moreover, collecting data to customise service is precisely what firms operating in the new, third-modernity information capitalism are supposed to do. While Google is the ultimate embodiment of surveillance capital, for Zuboff, the opposite model is best exemplified by...Apple, and, more specifically, by iPod, which, together with the iTunes platform, ‘made it possible for listeners to continuously re-configure their songs at will’ [ibid.: 34]. It turns out that the distance between the good, ‘rational’ information capitalism and the ‘rogue’ surveillance capitalism is not that great after all. Through different analytical lenses, the two corporations would, perhaps, be characterised as belonging to a common category of ‘platforms’ [see Srnicek, 2016]. Perhaps behavioural surplus extraction is not that different from behavioural reinvestment cycle to amount to an entirely new economic imperative, rather than a strategy of a few powerful corporations?

Symptomatically, at times, Zuboff’s narrative seems to lose sight of the systemic character of surveillance capitalism, equating it with the practices of individual firms like Google, Facebook, and Microsoft. Effectively, and in line with her functionalist leanings well-spotted by Morozov [Morozov, 2019], Zuboff does away with the structure vs. agency problem by casting it simply as a matter of scale, thus reducing economic imperatives to scalable business strategies, possibly even chosen at will. Yet the question remains, why did the strategy of capturing data exhaust become scalable in the first place? This is significant because Zuboff assumes that, at the bottom, capitalism is just a spontaneous response to the human need to live an ‘effective life’. In the second chapter, she writes: ‘Capitalism evolves in response to the needs of people in a time and place’ [Zuboff, 2019: 36]. Superficially, this sentence seems to be historicising capitalism. In fact, it does precisely the opposite. While showing that surveillance cap-

italism is a rogue mutation of capitalism produced by a series of contingent historical events, she at the same time assumes that capitalism tout court is simply a way of satisfying the human needs, thus sparing it of contingency and historicity.

...And What is Right with It?

Notwithstanding the influences of modernisation theory, the book does resonate very powerfully with some other contemporary critiques of capitalism. For example, Zuboff's point that the products of surveillance capitalism 'are about predicting us, without actually caring what we do or what is done to us' [Zuboff, 2019: 71], sounds strikingly similar to the late anthropologist David Graeber's [2006] argument about the affinity between violence and commerce. Both are extremely simplified forms of interaction that enable communication between strangers but do not require any interpretative labour necessary to understand each other. Violence and market exchanges enable one to 'have relatively predictable effects on other's actions without understanding anything about them', but do not require one to 'know or figure out who they think they are, what they want, find objectionable, etc.', except the 'single thing they want to acquire: gold, or fish, or calicoes' [Graeber, 2006: 76—77]. Interpretative labour is basically an effort of imagination: understanding ultimately requires imagining oneself in someone else's place. In another well-known piece, Graeber [Graeber, 2012] argued that hierarchical relationships tend to distribute the load of 'interpretative labour' very unevenly, producing ignorance at the top because those in control of wealth and power do not have to understand anyone else and are unable to imagine what it is like to be toiling far below the commanding heights. Thus, along with Foucauldian power/knowledge, one could also speak of power/ignorance [ibid.: 111]. Hierarchies can be simultaneously hierarchies of knowledge and ignorance: 'seeing like a state' is not the same as 'seeing like a peasant'. In fact, the former may be the exact inversion of the latter.

There are a few points on which Zuboff's account of surveillance capitalism seems to converge with Graeber's ideas. Zuboff is worried about the unequal distribution of knowledge that is getting concentrated in the hands of surveillance capitalists who use it to predict human behaviours, but what kind of knowledge is this? Does it make the algorithms—or the capitalists who own them—really so 'smart' as they claim? From the perspective outlined by Graeber, algorithms are paradoxical in the sense of being extremely knowledgeable and yet also extremely ignorant at the same time. Machine intelligence amasses so much data that no human being could ever possess, let alone process, but it is hardly capable of performing interpretative labour, even though it may pass the Turing test. The Big Other may be able to predict typical behaviours, but this does not necessarily amount to actual understanding, let alone anticipating what is genuinely new. In this sense, being smart as an algorithm may be practically equivalent to being stupid as a human being: just think about the annoying ads that are being shown to you in a completely inappropriate context, thus being experienced as a tactless intrusion, rather than smooth anticipation of what you really want, here and now.

This reading also sheds light on another issue in Zuboff's analysis—namely, her conspicuous silence on labour. Labour only gets mentioned in the context of its social division or as a fictitious commodity characteristic of the previous phase of capitalist

development, since under surveillance capitalism, its place is taken by human experience. In other words, Zuboff does not treat labour as a counterpart to capital, and for a reason. Ultimately, her problem with surveillance capitalism is that it is not capitalist enough, at least in its relationships with the users who do not enjoy the reciprocities of workers in the mass-production era. They are just sources of raw materials and not a party in a supposedly reciprocal market exchange. But does this mean that labour—in the double sense of work and the working class—has no relevance in surveillance capitalism?

In his review of Zuboff, Morozov [Morozov, 2019] makes a similar point, noting that Google's advertising revenues depend on the work of engineers and computer scientists but also, crucially, on indexing content from other sites that is necessary to link search queries to targeted ads. Indexing production costs are negligible, 'as the indexed content arrives almost for free', performed by 'the usual suspects: bots, hobbyists, academics, teenagers. But also plenty of precarious media professionals' [ibidem]. Thus, surveillance capitalism may have its proletariat, after all. Just think about the numberless crowd of SEO professionals and copywriters who spend much of their working time doing 'interpretative labour', so that a search algorithm will be able to recognise and index some ad published by some online retailer. It is tempting to say that machine learning algorithms indeed are capital, not exactly in the sense implied by Zuboff, but more in line with the Marxian concept of fixed capital, a chunk of congealed dead labour—in this case, 'interpretative labour'. Perhaps the analysis should look beyond surveillance and include the ways in which 'interpretative labour' gets objectified in the form of algorithms and the 'living' labour of the workers who make these algorithms 'smart'? Perhaps, then, unlike the students of business who begin from the firm level, one should follow Marx and start from the elementary cell of surveillance capitalism—namely, its 'prediction products'—and then try to trace their origins to human 'interpretative labour' that gives them their certainty and, thereby, their value? Even if somewhat far-fetched, this direction of analysis seems promising, testifying to the book's richness and provocative character. Shoshana Zuboff's *Surveillance Capitalism* is a wonderful intellectual provocation, a must-read for anyone interested in contemporary mutations of capitalism and information technology, and a timely reminder that something may be rotten in the digital kingdom.

References (Список литературы)

- Abbott A. (1993) The Sociology of Work and Occupations. *Annual Review of Sociology*. Vol. 19. P. 187—209. <https://doi.org/10.1146/annurev.so.19.080193.001155>.
- Arendt H. (2004) *The Origins of Totalitarianism*. New York, NY: Schocken Books.
- Beck U., Bonss W., Lau C. (2003) The Theory of Reflexive Modernization: Problematic, Hypotheses and Research Programme. *Theory, Culture & Society*. Vol. 20. No. 2. 1—33. <https://doi.org/10.1177/0263276403020002001>.
- Foster J. B., McChesney R. W. (2014) Surveillance Capitalism: Monopoly-Finance Capital, the Military-Industrial Complex, and the Digital Age. *Monthly Review*. Vol. 66. No. 3. P. 1—31. https://doi.org/10.14452/MR-066-03-2014-07_1.

Graeber D. (2006) Turning Modes of Production Inside Out: Or, Why Capitalism is a Transformation of Slavery. *Critique of Anthropology*. Vol. 26. No. 1. 61—85. <https://doi.org/10.1177/0308275X06061484>.

Graeber D. (2012) Dead Zones of the Imagination: On Violence, Bureaucracy, and Interpretive Labor. The 2006 Malinowski Memorial Lecture. *HAU: Journal of Ethnographic Theory*. Vol. 2. No. 2. 105—128. URL: <http://eprints.lse.ac.uk/53222/> (accessed: 18.02.2021).

Harvey D. (2003) The New Imperialism. Oxford: Oxford University Press.

Hirschman A. O. (1970) Exit, Voice, and Loyalty: Responses to Decline in Firms, Organizations, and States. Cambridge, Mass: Harvard University Press.

Morozov E. (2019) Capitalism's New Cloth. *The Baffler*. February 4. URL: <https://the-baffler.com/latest/capitalisms-new-clothes-morozov> (accessed: 18.02.2021).

Polanyi K. (1957) The Great Transformation. Boston: Beacon Press.

Skinner B. F. (1976) About Behaviourism. New York, NY: Vintage Books.

Srnicek N. (2016) Platform Capitalism. Oxford: Polity Press.

Wood E. M. (1994). From Opportunity to Imperative: The History of the Market. *Monthly Review*. Vol. 46. No. 3. 14—40. https://doi.org/10.14452/MR-046-03-1994-07_2.

Zuboff S. (2015) Big Other: Surveillance Capitalism and the Prospects of an Information Civilization. *Journal of Information Technology*. Vol. 30. No. 1. P. 75—89. <https://doi.org/10.1057/jit.2015.5>.

Zuboff S. (2019) The Age of Surveillance Capitalism: The Fight for the Human Future at the New Frontier of Power. New York: PublicAffairs.



