

РИСКИ ПРИМЕНЕНИЯ АЛГОРИТМОВ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В СОЦИАЛЬНО-ПОЛИТИЧЕСКОЙ СФЕРЕ: ОБЗОР СОВРЕМЕННЫХ НАУЧНЫХ РАБОТ



Сергей Владимирович Володенков,

Государственный академический университет гуманитарных наук,
Московский государственный университет имени М.В. Ломоносова,
Москва, Россия
s.v.cyber@gmail.com



Сергей Николаевич Федорченко,

Государственный академический университет гуманитарных наук,
Московский государственный университет имени М.В. Ломоносова,
Москва, Россия,
s.n.fedorchenko@mail.ru

Получена 05.10.2023.

Поступила после рецензирования 12.01.2024.

Принята к публикации 28.02.2024.

Для цитирования: Володенков С. В., Федорченко С. Н. Риски применения алгоритмов искусственного интеллекта в социально-политической сфере: обзор современных научных работ // Дискурс-Пи. 2024. Т. 21. № 2. С. 24–48. https://doi.org/10.17506/18179568_2024_21_2_24

© Володенков С. В., Федорченко С. Н., 2024



Аннотация

Целью данной статьи является выявление современных значимых научных проблем, связанных с рисками применения алгоритмов искусственного интеллекта в социально-политической сфере. Актуальность работы определяется текущими технологическими трендами, приводящими к метаморфозам в обществе и государстве и характеризующимися внедрением приемов предиктивной аналитики, рейтингования, пострасовой логики, различных режимов темпоральности и рекурсивности социальных процессов. При этом последствия и системные эффекты от внедрения алгоритмов искусственного интеллекта в актуальную социально-политическую практику различных стран еще недостаточно глубоко изучены. Основной методологической оптикой был выбран критический дискурс-анализ академических исследований. Также в работе использовались принципы сравнительного анализа значимых для общества и государства явлений. Вклад работы видится в теоретическом обобщении существующих исследований по выбранной теме. Проведенный анализ позволил выделить несколько ключевых проблем, обозначенных в современном научном дискурсе и связанных с внедрением алгоритмов искусственного интеллекта в социально-политическую сферу, – алгоритмизацию власти, непрозрачность нейросетевых алгоритмов, потенциал манипуляции массовым сознанием и алгоритмическую дискриминацию. На основе анализа академических исследований в статье сформулированы потенциальные сценарии внедрения алгоритмов искусственного интеллекта в актуальную социально-политическую практику. По итогам проведенного исследования делается ключевой вывод о том, что наибольшую значимость в социально-политическом пространстве имеет проблематика, связанная с алгоритмизацией власти на основе применения технологий искусственного интеллекта и нейросетевых алгоритмов. Происходящее перераспределение власти в сфере политического управления не просто обусловлено распространением алгоритмических практик, но также связано с возрастающей ролью цифровых технологических корпораций, претендующих на собственную политическую субъектность. Наиболее критичным сценарием авторам работы видится зарождение «алгоритмически замкнутого круга» и формирование широкого спектра связанных с таким алгоритмическим кругом эффектов непрозрачности, манипуляции и дискриминации в социально-политической сфере.

Ключевые слова:

искусственный интеллект, социально-политические риски, алгоритмическая дискриминация, непрозрачность алгоритмов, манипуляции массовым сознанием, алгоритмизация власти

Источники финансирования:

исследование выполнено в рамках проекта «FZNF-2023-0011 – Искусственный интеллект и нейросетевые технологии как инструмент современных социально-

политических коммуникаций: риски, угрозы и вызовы» при поддержке Министерства науки и высшего образования Российской Федерации и Экспертного института социальных исследований.

UDC 32.019.5

DOI: 10.17506/18179568_2024_21_2_24

RISKS OF USING ARTIFICIAL INTELLIGENCE ALGORITHMS IN THE SOCIO-POLITICAL SPHERE: A REVIEW OF CONTEMPORARY SCIENTIFIC WORKS

Sergey V. Volodenkov,

State Academic University of Humanities,
Lomonosov Moscow State University,
Moscow, Russia,
s.v.cyber@gmail.com

Sergey N. Fedorchenko,

State Academic University of Humanities,
Lomonosov Moscow State University,
Moscow, Russia,
s.n.fedorchenko@mail.ru

Received 05.10.2023.

Revised 12.01.2024.

Accepted 28.02.2024.

For citation: Volodenkov, S.V., & Fedorchenko, S.N. (2024). Risks of Using Artificial Intelligence Algorithms in the Socio-Political Sphere: a Review of Contemporary Scientific Works. *Discourse-P*, 21(2), 24–48. (In Russ.). https://doi.org/10.17506/18179568_2024_21_2_24

Abstract

This article identifies contemporary significant scientific problems related to the risks of using artificial intelligence algorithms in the socio-political sphere. The relevance of this work is determined by current technological trends, leading to metamorphoses in society and the state, and characterized by the introduction of predictive analytics techniques, ratings, post-racial logic, various modes of temporality, and the recursiveness of social processes. At the same time, the consequences, and systemic effects of the introduction of artificial intelligence algorithms into the current socio-political practice of various countries have not yet been sufficiently studied. Critical discourse analysis of academic research was chosen as the main methodological lens. This study also used the principles

of comparative analysis of phenomena significant for society and the state. The contribution of this work is seen in the theoretical generalization of existing research on the chosen topic. The analysis made it possible to identify several key problems identified in contemporary scientific discourse and associated with the implementation of artificial intelligence algorithms in the socio-political sphere – algorithmization of power, opacity of neural network algorithms, potential for manipulation of mass consciousness, and algorithmic discrimination. Based on an analysis of academic research, this study describes potential scenarios for the introduction of artificial intelligence algorithms into current socio-political practice. Based on the results of the study, a key conclusion is drawn that the greatest significance in the socio-political space is related to the algorithmization of power based on the use of artificial intelligence technologies and neural network algorithms. The ongoing redistribution of power in the sphere of political governance is not simply due to the spread of algorithmic practices but is also associated with the growing role of digital technology corporations claiming their own political subjectivity. The authors of this study see the most critical scenario as the emergence of an “algorithmically closed circle” and the formation of various opacity, manipulation, and discrimination effects associated with such an algorithmic circle in the socio-political sphere.

Keywords:

artificial intelligence, socio-political risks, algorithmic discrimination, opacity of algorithms, manipulation of mass consciousness, algorithmization of power

Funding:

the study was carried out within the framework of the project “FZNF-2023-0011 – Artificial intelligence and neural network technologies as a tool for contemporary socio-political communications: risks, threats and challenges” with the support of the Ministry of Science and Higher Education of the Russian Federation and the Expert Institute of Social Research.

Введение

Возникновение и развитие искусственных интеллектуальных систем заставляет современное политологическое сообщество всерьез задуматься о новых рисках, вызовах, актуальных проблемных полях, влияющих на изменения в важной для человека, общества и государства социально-политической сфере. Комплекс вопросов, рассматриваемых в связи с этим в политической науке, весьма разнообразен и произведен от воздействия технологической среды на традиционные социальные институты и отношения. Нельзя исключать, что растущий запрос на формирование адаптирующейся к технологическим вызовам гибкой научной картины мира, учитывающей междисциплинарный уровень знания, может стать важнейшей интегральной задачей и связующей

функцией политической науки. Такая полноценная картина мира будет отражать не только постоянно обновляемые достижения технических, естественных наук, но и социальных, гуманитарных наук, обладая объяснительным и прогностическим инструментарием.

Большинство научных исследований, затрагивающих поднимаемый нами вопрос, эмпирически опираются на реальную обеспокоенность людей из-за алгоритмизации традиционной системы социальных отношений. Ученые уже фиксируют разные группы таких людей (Kieslich et al., 2022): «human in the loop» (не доверяют искусственному интеллекту, лучше остальных осведомлены о его рисках, против автономии машин, менее образованы), «ethically concerned» (представители данной группы относятся к молодежи, требовательны к высоким этическим стандартам техники, хорошо образованы, имеют доверие и интерес к искусственному интеллекту, характеризуются низким восприятием рисков из-за него), «safety concerned» (среднего возраста, интересуются искусственным интеллектом, в оценках о его рисках расположены между двумя предыдущими группами, обеспокоены безопасностью), «fairness concerned» (представители молодежи, обеспокоены проблемами справедливости, менее всех склонны видеть в искусственном интеллекте риски, больше всех им интересуются), «indifferent» (меньше всех групп интересуются вопросами искусственного интеллекта).

Не касаясь футурологических тем сингулярности, «сверхума», трансгуманизма, «войны машин против человечества» и чрезвычайно отдаленного будущего, в данной работе основной акцент сделан на задаче выявления современных значимых научных проблем, связанных с рисками применения алгоритмов искусственного интеллекта в социально-политической сфере. Актуальность настоящей работы и ее основной вклад видятся в теоретическом обобщении исследований по теме, в обозначении, с одной стороны, базовых технологических новаций, влияющих на структуру современного общества, и, с другой стороны, обнаружении самых значимых последствий таких новаций в социальной сфере и политике.

Чтобы определить степень академической полемики по выбранному вопросу, достаточно обратить внимание на противоречивые оценки, касающиеся базовой терминологии. Ведь рабочий понятийный аппарат – это связующий компонент любой объяснительной модели, теории или сценарного прогноза. Например, исследователи стараются отличать «управление с помощью ИИ» (governance by AI) от «управления ИИ» (governance of AI). Если «управление с помощью ИИ» включает механизмы, технологии автоматизированного принятия решений, используемых органами государственной власти, то «управление ИИ» отсылает к тем политическим структурам и практикам (Erman, Furendal, 2022), которые занимаются регулированием, направлением и внедрением разработок технологий искусственного интеллекта (например, Европейская комиссия). Несмотря на сохраняющееся терминологическое разнообразие, во многих научных работах стали общеупотребимыми такие понятия, как алгоритмы, нейросети, машинное обучение. Тогда как среди понятий, описывающих социально-политические эффекты внедрения искусственного интеллекта, усто-

явшимися являются дискриминация, непрозрачность, «черный ящик» и близкие им по смыслу. Эти тенденции были учтены в настоящей работе.

В качестве основной методологической оптики работы был выбран критический дискурс-анализ научных проблем академических исследований, касающихся темы рисков применения алгоритмов искусственного интеллекта в социально-политической сфере. Поэтому критериями выбора материала исследования стали аспекты дискриминации, неравенства, новых практик доминирования, подчинения в условиях внедрения алгоритмов нейронных сетей и интеллектуальных автоматизированных систем. Наиболее важные предположения, выводы, гипотезы, модели разных ученых по этому поводу были условно сгруппированы по нескольким проблемным направлениям: «алгоритмизация власти», «непрозрачность», «манипуляции», «дискриминация». Рассмотрение таких проблем исходило из их соотношения со значимыми для общества и государства явлениями, к которым можно применять принципы сравнительного анализа. При этом сразу важно уточнить, что данные риски во многом тесно взаимосвязаны между собой.

Краткий обзор научной литературы

Следует подчеркнуть, что исследователи выявляют самые разные проблемные поля, научные проблемы и аспекты, посвященные рискам внедрения цифровых технологий, Интернета вещей, искусственного интеллекта в социально-политическую жизнь. Имеются научные труды, в которых анализируются не отдаленные риски, а те социальные проблемы цифрового общества, которые уже наметились, хотя и не играют фундаментальной роли в настоящее время, но могут обостриться из-за внедрения алгоритмов.

Например, социотехническая реальность может изучаться как новый этап развития современной реальности, где «социальное» уже нельзя отделить от «технического». Г. Ловинк в связи с этим в своих трудах предположил, что активность такой социальной группы, как программисты, привела к тому, что социальное стало обуславливаться программами, техническими системами и больше проявляться в сетевом формате (Ловинк, 2019, с. 25–39). Но социотехническая реальность может рассматриваться и в рамках концепции фиджитал-мира, допускающей тесное переплетение физических (physical) и цифровых (digital) пространств. Такой фиджитал-мир ориентирован на конструирование взаимосвязи между брендом и потребителем корпорации, а его создание сопровождается внедрением приемов сбора и анализа данных пользователей (Švec, Madleňák, 2017).

Важно заметить, что современные исследователи часто не разделяют четко угрозы возникновения новых практик управления массовым сознанием (Ulbricht, 2020) и риски нарушения конфиденциальности информации, сбора данных о гражданской активности, формирования социума круглосуточной тотальной слежки. Встречаются интересные попытки реконцептуализации уже существующей модели новой алгоритмической идентичности Дж. Чейни-Липпольда (идентичности, формируемой с помощью вмешательства алгоритмов в социальную

жизнь (Cheney-Lippold, 2011)) и концепции алгоритмического чрезвычайного положения Д. Маккуиллана (переосмысления теории чрезвычайного положения в цифровых условиях (McQuillan, 2015)). В них внедрение алгоритмов рассматривается как вызов сложившемуся механизму государственных субъектов власти (Kubler, 2017). Но воздействие данных интеллектуальных систем классифицируется как неоднозначное: в одних случаях это мягкие приемы биополитики, а в других – более жесткие варианты государственного принуждения.

Исследовательский интерес представляет и модель концентрации цифровой власти П. Немитца, который усматривает в качестве источника данной формы власти усиление социально-политической роли крупных цифровых корпораций. По Немитцу, новые риски обусловлены доминированием таких корпораций в сфере инвестиций в разработку искусственного интеллекта, их потенциалом влияния на политику и общество (Nemitz, 2018). Критически переосмысливает концепцию концентрации цифровой власти Д. С. Жуков, предложивший дефиницию «политический искусственный интеллект», под которым он понимает интеллектуальные приложения в политической сфере (Жуков, 2020). Жуков, в отличие от Немитца, видит риски в игнорировании проблемы «политического искусственного интеллекта».

Ряд авторов в качестве проблемного поля затрагивают риски некачественных результатов алгоритмических систем, грубого вторжения в частную жизнь человека, нарушения автономии личности, разрушения существующих социальных связей (Талапина, 2021). Обсуждаются риски ошибок алгоритмов, риски военного плана, дегуманизации страны (Романова, 2021), а также риски дегуманизации самого права (Овчинников, 2020).

Безусловно, такой краткий обзор не позволяет охватить все тренды или пробелы в исследовательской литературе. Между тем, важно сконцентрироваться на тех научных проблемах и аспектах, которые отражают наиболее существенные риски внедрения алгоритмов искусственного интеллекта для общества и государства. Далее будут представлены полученные результаты анализа в соответствии с выявленными проблемами в научной литературе.

Алгоритмизация власти

В выявленной проблеме алгоритмизации власти (проникновения в практику функционирования органов власти алгоритмов, алгоритмизированных систем) основные смысловые конструкции относятся к степени рисков изменения власти из-за распространения алгоритмов искусственного интеллекта. Обозначенную научную проблему, тесно связанную со значимыми метаморфозами власти, можно условно свести к двум основным позициям. С одной стороны, некоторые авторы полагают, что алгоритмизация, алгоритмическое управление, нейронные сети в политике – это отражение чрезвычайно важной интеграции интеллектуальных систем в традиционные системы власти, не меняющей их кардинально, но все же адаптирующей к цифровизации социально-политической сферы. С другой стороны, ряд авторов считают, что алгоритмизация власти означает

принципиально новый феномен, не имеющий аналогов до эпохи цифровых трансформаций. Например, появились работы, в которых в качестве результата алгоритмизации власти исследуется аллократия – особая разновидность власти, основанная на алгоритмах, распространении программных кодов, и, как минимум, меняющая существующий рынок труда (Aneesh, 2006). Несмотря на неоднозначность, дискуссионность категории аллократия и встречающиеся параллели с уже изучаемыми феноменами механизации и автоматизации в научной литературе, новизна такого явления видится в факторе машинного обучения, самообучающихся алгоритмических систем, принципиально отличающихся от прежних информационных систем (Фурс, 2022).

Что говорят зарубежные и отечественные политологи об известной природе традиционной власти? С. Льюкс предложил довольно простую формулу, согласно которой *A* осуществляет власть над *B*, так как *A* воздействует на сторону *B* способом, противоречащим интересам последней (Льюкс, 2010, с. 49). Тогда как А. Ф. Филиппов предлагает считать властью «асимметричное стабильное социальное отношение, наблюдаемое как перепад возможностей и шансов эффективных преднамеренных действий его участников» (Филиппов, 2019, с. 141). Филиппов отмечает, что современные трактовки власти могут касаться ее критически, рассматривать ее как универсальное социальное отношение либо специфическое средство, понимать ее как индивидуальное отношение или характеристику отношений групп, раскрывать ее как эффект либо возможность ее достижения (Филиппов, 2019, с. 151–153). В. Г. Ледаев понимает под властью способность субъекта обеспечивать в соответствии со своими намерениями подчинение объекта (Ледаев, 2001, с. 268). На первый взгляд, алгоритмы как раз и представляют собой лишь новые способы в обеспечении власти, не меняя ее принципиально. Между тем, серьезные изменения все же происходят. Так, на сегодняшний день уже встречаются работы о становлении алгоритмически обусловленного рекурсивного общества, в котором действия основаны на предыдущих действиях, циркуляции данных, обработанных алгоритмическими процессами сортировки (Beer, 2022). Рекурсивное общество характеризуется не изолированными циклами, а свертыванием извлекаемых и возвращающихся в нашу социально-политическую жизнь данных (data coiling), трансформирующих отдельные петли обратной связи в закрепленные образы практик и поведения. При таком развитии у владельцев алгоритмических систем возникают большие соблазны по закреплению своего доминирования на рынке или в политике.

Первоначально ставка правительств на алгоритмическое управление аргументировалась необходимостью совершенствования самой власти, лучшего понимания проблем общества. Отсюда во властных структурах стали популярными концепции «радар общества» (society's radar), «доставки» обществу результатов (delivery). Если концепция «радар общества», продвигаемая премьер-министром Канады П. Трюдо во второй половине прошлого века, была сосредоточена на превращении федеральной власти в активный планирующий орган посредством получения правильной информации (идея возникла не без влияния М. Маклюэна), то, согласно подходу «доставки», в Канаде стали рассматривать государственную

службу наподобие брокера данных, как дискретную систему, где политический и управленческий контроль мог осуществляться путем передачи на аутсорсинг технологическим корпорациям правительственных функций сбора, обработки информации, количественной оценки социальной жизни с помощью искусственного интеллекта. Однако такая модель рационального правительства вызывает критику из-за чинимых ею препятствий распространению публичной информации (Lepage-Richer, McKelvey, 2022). Машинное обучение способно внести в политическое управление новый метауровень формирования правил, так как алгоритмы могут автоматически корректироваться и подстраиваться под текущую ситуацию для повышения точности по причине получения новых данных (Plantinga, 2022).

Алгоритмическое управление как признак новой социотехнической реальности, где социальные отношения попадают в прямую зависимость от технологической среды, меняет саму темпоральность, формируя непрерывное взаимодействие людей и алгоритмов, опосредующих их активность (при этом нельзя исключать формирования отличающихся режимов темпоральности для разных социальных категорий – более быстрого темпа времени для большинства населения и более размеренного образа жизни для представителей элиты, обладающих абсолютным правом на цифровой детокс). Так, цифровые платформы, использующие различные виды рейтингования, репутационного контроля, наблюдения, автоматизации, сокрытия информации, усиливают власть менеджеров над работниками. К примеру, алгоритмы способны увеличить темп рабочего процесса (посредством оповещающих портативных устройств), создавая риски повышения стресса, травмирования среди работников. Кроме того, автоматизация уменьшает власть менеджеров среднего звена в пользу самих алгоритмов (Jarrahi et al., 2021). Научные труды других авторов доказывают, что похожие процессы перераспределения власти происходят и в политической сфере. П. Плантинга на примере анализа алгоритмизации в африканских странах предположил, что алгоритмы искусственного интеллекта приведут к усилению централизации власти в африканских государствах, лишению полномочий у местных чиновников, одновременно увеличивая нагрузку на рядовых сотрудников и зависимость от зарубежных технических специалистов (Plantinga, 2022). В свою очередь, на примере исследования внедрения алгоритмизированных систем прогнозирования в немецкой полиции также получены выводы о централизации власти посредством аллократии, автоматизирующей и рационализирующей традиционную бюрократию в веберовском понимании (Lorenz et al., 2021). В США алгоритмизированная система COMPAS используется судебными властями с целью оценки потенциальных преступлений со стороны граждан (Borgesius, 2020). А. Аниш исследовал схожее явление на примере индийских программистов, проживающих в разных городах (Нойда, Гургаон (Индия) и Нью-Джерси (США)) и работающих в американской сфере программного обеспечения. Работа Аниша позволяет уточнить, что феномен аллократии в сочетании с фактором программного кода отражает рост влияния корпораций на социальные процессы (Aneesh, 2006).

Однако следует признать, что, несмотря на саму фиксацию социально-политических рисков от внедрения алгоритмов, эффекты от их внедрения в разных странах еще недостаточно изучены. В настоящее время имеется определенный пробел в понимании того, почему в одних странах алгоритмизация власти больше связана с усилением надзорных возможностей и полномочий центральных органов политической власти, а в других – с ростом влияния крупных корпораций и корпоративных органов власти. Тем не менее, ученые отмечают, что сегодня зарождается алгоритмический метакapитал – специфическая символическая власть, действующая посредством алгоритмов, устанавливающих новые правила видимости/невидимости, а также заметности, приоритетной видимости для людей, данных, знаний (Lundahl, 2022).

Непрозрачность нейросетевых алгоритмов

Проблема алгоритмической непрозрачности подразумевает рассмотрение нейронных сетей и интеллектуальных систем как некоего опасного «черного ящика», чья непредсказуемость и неоднозначность не дают полностью просчитать риски внедрения алгоритмов для человека, общества и государства.

Действительно, случаи эффектов непрозрачности уже обсуждаются учеными, а также экспертами в особых группах ООН. Например, до сих пор остается до конца неясным, действовал ли самостоятельно дрон при выслеживании и атаке боевиков в Ливии в марте 2021 г. (Reed, 2022). С этой точки зрения интересно обратить внимание на попытки анализа источника эффекта «черного ящика» – феномена рациональной безотчетности (*rationalized unaccountability*) как новой идеологии осуществления власти и контроля алгоритмическим способом, через программы интеллектуальных агентов (Vesa, Tienari, 2022). Эта идеология разветвляется на утопические темы (от трансгуманизма до сверхинтеллекта), оправдывая дисциплинарные практики как самоочевидные, рассматривая интеллектуальных агентов как совершенных рациональных агентов и отвергая любую критику как иррациональную, старомодную, романтическую. М. Весса и Я. Тиенари из Школы экономики Ханкен считают, что такая идеология – замаскированный вариант «патрицианского популизма», который формируется элитой для сохранения своей власти и борьбы с набирающим силу «плебейским популизмом» (Vesa, Tienari, 2022).

Непрозрачность нейросетевых алгоритмов уже начала приводить к драматическим последствиям: в 2018 г. в Китае система распознавания правонарушений оштрафовала Д. Минчжу, ошибочно приняв ее изображение в рекламе за пешехода, в том же году в США беспилотная машина Uber сбила велосипедистку, не распознав в ней пешехода, в 2020 г. в России система распознавания лиц в магазине «Ашан» ошибочно приняла покупателя за преступника (охранники передали человека полицейским, которых в незаконности задержания смог переубедить только адвокат). Подобные риски усугубляются тем фактом, что люди часто доверяют готовым решениям от нейросетей. Эксперименты, проводившиеся в полицейских управлениях Йокогамы и Ванкувера, свидетельствуют,

что ответственные лица, вовлеченные в механизм принятия решений с использованием искусственного интеллекта, соглашались с предлагаемыми системой решениями в 98 % случаях (Миронова, 2021).

В некоторых работах эффект «черного ящика» сравнивают с алхимией: если раньше алхимики обещали превратить свинец в золото, то современные предприниматели точно также обещают благодаря алгоритмам искусственного интеллекта достичь впечатляющих результатов и прогнозов, хотя механика интеллектуальных систем может наподобие алхимической лаборатории по-прежнему оставаться непроницаемой для взглядов непосвященных. Например, в одной из работ изучается кейс оборонной корпорации Anduril, основанной сторонником Д. Трампа, американским бизнесменом П. Лаки в 2017 г. Anduril позволяет создавать целые «виртуальные стены» на государственной границе, состоящие из наблюдательных дронов, сторожевых вышек, обрабатывающих их данные интеллектуальных систем, а также системы предупреждения пограничников. Другой стартап Brinc рекламировал дроны, способные допрашивать и наносить электрошоки мигрантам на границе Мексики и США (Milivojevic, 2022). По сути, такие эксперименты частных компаний снижают ответственность государств за гибель человека на границе на фоне абсолютного отсутствия общественного контроля. Непрозрачность становится идеологией, прикрывающейся коммерческой тайной и сложным экспертным дискурсом.

Д. Маккуиллан из Лондонского университета подчеркивает, что в большинстве случаев, кроме наиболее простых, почти нереально понять, как именно алгоритмы машинного обучения при всей сложности выведенной для сопоставления выходных и входных данных функции и огромном объеме используемых переменных обрабатывают данные. При этом непрозрачность накладывается на отмеченную выше рекурсивность – алгоритмы связывают прогноз с прошлыми наблюдениями, подменяя корреляцией выявление причинно-следственных связей. Из-за эффекта рекурсии машинные процессы основываются сами на себе (McQuillan, 2016), усиливая способность к определению образов, явлений (инцепционизм), но не исключают проблему необратимости и ошибок, находя закономерности там, где их нет. Маккуиллан называет этот эффект «алгоритмической паранойей». При этом следует отметить, что в настоящее время также встречаются критические оценки рисков от «черного ящика», больше понимающие алгоритмы как элемент коллективных человеческих практик, институтов. Другими словами, по мнению ученых, не только алгоритмы могут влиять, например, на культуру, но и культура воздействует на алгоритмы (Seaver, 2017).

Манипуляции массовым сознанием

Риски, относящиеся к потенциальным негативным социально-политическим эффектам из-за воздействия алгоритмов на массовое сознание, можно в целом отнести к проблеме манипуляции. Такие риски, по сути, могут расцениваться как последствия цифровых метаморфоз власти и ее сохраняющей-

ся непрозрачности, поэтому данная область также связана с соответствующими исследовательскими аспектами. Как правило, манипуляции массовым сознанием становятся инструментом для различного рода дискриминации.

У современных специалистов определенные опасения вызывает алгоритмическое конструирование информационной среды, основанное на предпочтениях людей, а также их прогнозируемых поведенческих реакциях. Манипулятивные риски усиливаются тем фактом, что искусственный интеллект перенастраивает привычные способы доступа общества к политической информации и новостям. Корпоративные техногиганты наподобие Google и Microsoft, проводя эксперименты с большими языковыми моделями (LLM) для генерации автоматически формируемого контента по запросу пользователей, снижают возможности монетизации для средних и малых корпораций (Jungherr, 2023). Такой процесс может привести к ограниченному, одностороннему освещению политических событий, уменьшению количества новостных агентств. Одновременно возрастают риски того, что при развитии капитализма платформ и сетевых эффектов, описанных Н. Срничеком, цифровые платформы станут распространять тот контент, который повышает вовлеченность пользователей в социальные сети, и ликвидировать контент, препятствующий достижению этой цели (Elkin-Koren, 2020). Тем самым негативные сетевые эффекты потенциально могут угрожать конфиденциальности, свободе слова и политическому плюрализму (регулярный мониторинг пользовательского контента осуществляют проекты Amazon Project Zero, Content ID YouTube и др.).

Некоторые исследования показывают, что манипуляции могут снизить доверие к политическому лидеру с помощью негативно подающих его дипфейков (технологий искусственного интеллекта, синтезирующих реалистичные фото-, аудио- и видеоматериалы, в том числе, для целенаправленного ввода в заблуждение). Помимо этого, дипфейки могут исказить дебаты, снижать доверие к новостному контенту в социальных сетях, переводя микротаргетированную политическую пропаганду для задач манипулятора на более совершенный уровень (Ó Fathaigh et al., 2021). Справедливости ради стоит отметить наличие научных работ, критически оценивающих потенциалы дипфейков в процессах политической манипуляции и определяющих пределы генерации фальшивой реальности, но даже в них подтверждается способность дезинформации влиять на оценку политиков среди пользователей (Hameleers et al., 2022).

Распространению дипфейков поспособствовали удобные и легкодоступные инструменты наподобие FaceApp и DeepFaceLab. В Китае, где дипфейки получили название «хуанлянь», эту роль сыграло приложение ZAO от социальной сети Momo. На этом фоне авторы пишут о рисках преследования, нанесения вреда, потери прав на изображение из-за метрик (особенностей) алгоритмов (Seta, 2021). Манипуляции, вызванные дипфейками, могут помешать получить истинное понимание ситуации, если, к примеру, других свидетельств, кроме видео, аудио и фото, просто не существует. Эпистемические угрозы дипфейков возрастают и из-за того, что они снижают объем информации, которая может транслироваться зрителям (Fallis, 2021).

Недавний эксперимент, проведенный с алгоритмами Netflix, свидетельствует в пользу возможности конструирования информационных капсул в социально-политическом сегменте потребления контента. Активность трех стилистически разных пользовательских аккаунтов на платформе данного стримингового сервиса сериалов и фильмов («Культурный сноб», «Заядлый спортивный фанат» и «Безнадежный романтик») к концу эксперимента показала, что в отношении каждого из них произошла сильная персонализация предпочтений. Ввиду этого возникают опасения, что «алгоритмическое самоподтверждающееся эстетическое потребление» с петлями обратной связи еще больше усиливает имеющиеся у пользователя предпочтения, снижая его интерес к альтернативным культурным предложениям, гомогенизирует культуру телевидения и кино, а также создает риски влияния отдельных петель обратной связи пользователей платформ на формирование глобальных циклов, навязывающих рекомендации другим пользователям по всему миру (Pajkovic, 2022). Эти наблюдения подтверждают идею Ж. Эллюля об ортопраксии: эффективная пропаганда не просто убеждает человека в чем-то, но и вовлекает его в определенную систему действий, способствующих закреплению необходимой идеи. И в данном случае возникает важный вопрос: кто и для каких целей будет использовать подобные механизмы создания популярных информационных капсул?

Алгоритмическая дискриминация

Проблеме алгоритмической дискриминации посвящено много научных работ, выделяющих различные ее проявления по расовому, гендерному, политическому, социальному и другим признакам. Если касаться темы дискриминации цифровых прав, то в это направление можно включить приемы цифрового контроля и слежки за человеком. К проблемному контуру дискриминации также логично отнести риски замены искусственным интеллектом рабочих мест и целых профессий.

На рост рисков дискриминации со стороны алгоритмов повлияли военно-политические стратегии, например, американская стратегия глобальной войны с терроризмом (Global War on Terror, WoT), которая стала реализовываться США после терактов 11 сентября 2001 г. Стратегия WoT стимулировала инвестиции в корпорации, занимающиеся обеспечением безопасности. Одним из пионеров в этой сфере стала PredPol (затем переименованная в Geolitics), соучредитель которой профессор антропологии Дж. Брантингем сделал ставку на выявление наиболее подверженных рискам городских геолокаций. Брантингем перефилировал алгоритмы, ранее прогнозировавшие землетрясения, для обнаружения «горячих точек» города. Существует позиция, согласно которой запрос на алгоритмы безопасности приводит к рискам «пострасовой логики» – сохранения под риторикой расовой беспристрастности факта маркировки групп, вызывающих подозрение (например, мусульман), соотнесения с белой расой разных категорий населения. Под пострасовой логикой некоторые авторы подразумевают возобновление условий расового угнетения из-за внедрения практик

алгоритмического профилирования в рамках борьбы с терроризмом (Sharma, Nijjar, 2023). Пострасовая логика алгоритмов включает и принцип «цифровой красной линии», означающей отличающееся отношение тех же банков к клиентам разных рас (Jouce et al., 2021).

Авторы отмечают появившиеся случаи дискриминации афроамериканцев в США со стороны алгоритма, внедренного для распределения ресурсов в системе здравоохранения. В результате из-за алгоритмической оценки афроамериканцев стали гораздо реже отправлять на лечение, чем представителей белой расы, хотя они были одинаково больны (Bartlett, 2023). Внедрение прогнозной аналитики в полицию Лос-Анджелеса усилило неравенство, так как способствовало ужесточению слежки за считавшимися подозрительными горожанами и более пристальной оценке рисков со стороны цветных кварталов с низкими доходами (Jouce et al., 2021). Еще в 2015 г. возник расовый скандал из-за того, что Google Photos ошибочно соотносил фото афроамериканской пары с тегом «гориллы». Фиксировалась и дискриминация женщин со стороны программы проверки резюме корпорации Amazon, которая за основу для анализа использовала уже имевшиеся в базе резюме более успешных мужчин. Дискриминация может быть также связана со стереотипами, предвзятостями, ценностями сотрудников таких корпораций как Google, стремление которых к максимизации прибыли заставляет их отдавать предпочтение довольно специфическим системам поиска и ранжирования. Помимо этого, новые формы дискриминации могут возникать из-за неверных алгоритмических прогнозов или ранжирования. К примеру, услуги страховых кампаний могут стать для некоторых граждан недоступными, так как алгоритмические системы будут связывать низкий доход жителей конкретного района города с высокими рисками краж со взломом и повышать цены за страховку (Borgesius, 2020). Однако при этом Л. Рафанелли отмечает, что считать подобные случаи доказательством того, что происходит замена «власти человека» на «власть искусственного интеллекта», неверно, так как искусственный интеллект – лишь инструмент, благодаря которому людьми осуществляется власть (Rafanelli, 2022).

На фоне тотальной роботизации и построения «социумов роботов» также прогнозируются риски фундаментальных трансформаций социальной структуры – от трудовой деградации и социального расслоения до превращения среднего класса в прекариат с непостоянной занятостью (Лексин, 2021, с. 232–233). Вместе с тем встречаются работы, авторы которых сомневаются в стремительной замене искусственным интеллектом существующих профессий и эффектах тотальной социально-профессиональной дискриминации. По мнению таких авторов, ситуация намного сложнее, так как последовательная эволюция искусственного интеллекта (механического, аналитического, интуитивного и эмпатического) приводит к замене рабочих мест не на уровне должностей, а на уровне задач (Huang, Rust, 2018). Тем не менее, по итогам анализа научных источников в качестве одного из сценариев, связанных с внедрением алгоритмов искусственного интеллекта в ключевые сферы жизнедеятельности государства и общества, нам видится зарождение «алгоритмически замкнутого круга», характеризующегося

различными социально-политическими эффектами алгоритмизации власти, непрозрачности, манипуляции и дискриминации.

Сценарии внедрения алгоритмов искусственного интеллекта в актуальную социально-политическую практику

Алгоритмизация власти видна в постепенном дополнении традиционных социальных отношений «человек – человек» интерфейсом «человек – AI» и практиками «AI – AI», что, в свою очередь, приводит к конвергенции традиционных и цифровых форм пропаганды. Какие могут быть сценарии рисков такой конвергенции для цифровой и ценностной составляющей общества, государства и его суверенитета? Полностью исключать сценарий сохранения традиционных форм информационно-коммуникационного социально-политического взаимодействия «человек – человек», конечно, нельзя (это может быть обусловлено уровнем экономики, цифровизации, степенью развития информационно-коммуникационных систем в конкретной стране). Как показывают исследования на примере США, Германии, африканских и других государств, алгоритмизация способна поддерживать архаизацию, централизацию власти, псевдоцифровизацию, зависимость страны от внешних специалистов или же усиливать власть крупных корпораций.

Наиболее сложно просчитываемый сценарий развития социально-политической сферы, власти и пропаганды – распространение процессов, зависящих от практик «AI – AI», когда резко усилятся эффекты «черного ящика», вызовы предвзятости. Сама власть может попасть в созданную ею же информационную капсулу. Алгоритмизация традиционных субъект-объектных отношений, по нашему мнению, становится наиболее опасной угрозой для человека, так как она имеет «сквозной», «зонтичный характер» и дает импульс для непрозрачности, манипуляции и дискриминации нового типа. Поэтому теории эхо-камер, пузырей фильтров, информационных капсул могут быть не только вариантами концептуализации описываемых явлений, но и интеллектуальными провокациями-попытками приоткрыть «черный ящик» непрозрачности алгоритмов.

Безусловно, существует целый корпус аргументов (подробный обзор соответствующих исследований провел А. Брунс, ссылающийся на З. Борхесиуса, Р. Гарретта, С. Майнека), критикующих феномен эхо-камер, лежащий в основе генерации информационных капсул. В работе Брунса приводятся разные аргументы (Брунс, 2023): обосновывается тезис о том, что влияние алгоритмов на людей излишне переоценено; пользователи по-прежнему могут выбирать источники для потребления информации; эхо-камеры и пузыри фильтров не могут быть полностью герметически закрытыми от внешнего информационного воздействия; самые активные идеологи не способны запирается в данных пузырях фильтров и эхо-камерах; представить соблюдение постоянной коммуникационной дисциплины из-за алгоритмов невозможно; в Интернете не может быть политической поляризации и т.п.

Однако можно выдвинуть и корпус контраргументов. Критика алгоритмических эффектов часто не учитывает регистрируемые процессы складывания социотехнической реальности (фиджитал-мира), пострасовой логики сокрытия новых типов дискриминации, появления алгоритмически обусловленного рекурсивного общества, в котором извлекаемые, обрабатываемые интеллектуальными системами цифровые следы людей возвращаются в нашу социально-политическую жизнь, трансформируя отдельные петли обратной связи в закрепленные образы социальных практик и политического поведения. Тем самым феномен ортопраксии обретает новое алгоритмическое воплощение. Дополнительным подтверждением этого являются сетевые эффекты цифровых платформ (Н. Срничек) и особенности капитализма слежения (Ш. Зубофф), формирующие благоприятные условия для информационных капсул в виде гиперперсонализации предпочтений (Срничек, 2020; Зубофф, 2022).

Кроме того, критика алгоритмизации не учитывает трансформацию роли гейткиперов сообществ (модераторов правил, цензоров дискурса), в качестве которых могут выступать уже не только традиционные лидеры общественного мнения, политики, влогеры, блогеры, но и сами техногиганты, вовлеченные в процессы государственно-корпоративной гибридизации, генерации отношений «человек – AI». В роли таких гейткиперов, по сути, выступают программисты и их заказчики-корпорации, которые способны задать искусственным интеллектуальным системам определенную систему ценностных предпочтений и базовых элементов мировоззрения (при этом принципиально ничто не мешает искусственным интеллектуальным системам взять на себя часть функций гейткиперов). Потенциально такие гейткиперы могут преследовать задачи пропаганды новых ценностей в обществе. Риски алгоритмической таргетированной пропаганды гейткиперов гипотетически могут предполагать дискриминацию через цифровую стигматизацию тех групп, которые поддерживают власть, и продвижение ценностей, политического месседжа тех групп, которые хотят эту власть свергнуть. Политические манипуляции, ориентированные на цифровую дискриминацию целых категорий населения, способны спровоцировать кризис общества и запустить процесс его распада, сопоставимый с крахом Советского Союза, чтобы привести к власти новую элиту.

Конвергенция же современной пропаганды может предполагать формирование так называемых цифровых кентавров – автономных социотехнических систем, состоящих из людей-специалистов экстра-класса и полностью адаптированных под них нейросетевые алгоритмы (сценарий развития интерфейса «человек – AI»). Данная конвергенция в своей основе подразумевает процесс сближения (лат. *convergo* – «сближаю»), объединения традиционных и цифровых технологий пропаганды, нацеленных на решение общих задач в физическом и виртуальном мирах. Фактически сегодня появляются гибридные формы пропагандисткой деятельности, включающие автоматизированные системы и их мониторинг, выполняемый оператором-человеком. «Цифровые кентавры» могут не только открывать позитивные перспективы для функционирования разных видов цифровой дипломатии, парламентаризма, политических партий

и государственных ведомств, но и предполагают новые риски возникновения частных военных корпораций, частных информационно-пропагандистских компаний нового типа (работающих в режиме 24/7), цифровых сект, неуловимых экстремистских и террористических сетевых групп, автономного кибероружия. На деле те же гейткиперы имеют все шансы стать такого рода «цифровыми кентаврами».

Угрозы для цифрового и ценностного суверенитета также могут провоцировать разные государства на конструирование независимых от транснациональных корпораций государственных цифровых экосистем, в рамках которых могут реализовываться процессы социально-политической коммуникации. Необходимо учитывать и контрглобализационные процессы, которые можно условно назвать «цифровым отгораживанием» друг от друга политических режимов, государств. Основной сутью таких процессов является то, что государства начинают связывать геополитическую деятельность технологически развитых держав с активностью лояльных им цифровых корпораций, запускающих петли обратной связи алгоритмически обусловленного рекурсивного общества для неоглобализации, проводимой в пользу лишь ограниченного числа политических акторов. Примечательно, что в книге Г. Киссинджера, Э. Шмидта и Д. Хаттенлокера «Искусственный разум и новая эра человечества» даже предложена категория «AI-держав», к которым авторы причисляют США, Китай и Россию, гипотетически способных для защиты своих интересов применять новейшие типы оружия на основе искусственного интеллекта (Киссинджер и др., 2022, с. 147–148).

Формирование нового многополярного мироустройства происходит не просто через создание государствами внутренних цифровых экосистем, но и активную генерацию собственных «цифровых ойкумен» – сложных систем, включающих медийные агентства, порталы, интернет-ресурсы, сетевые сообщества и продвигающих ценностную и политическую повестку данных государств в других странах посредством сложных практик цифровой дипломатии. Вероятно, техники цифровой дипломатии той или иной страны будут зависеть от особенностей гибридизации форм «мягкой силы», «умной силы» и «острой силы». Другим значимым направлением становится анализ потенциальных рисков и угроз, связанных с деятельностью «виртуальных дипломатов» (интеллектуальных систем, ориентированных на дипломатические задачи), которые могут спровоцировать новые конфликты (в этом случае отдельные приемы цифровой дипломатии, например, генерация дипфейков, становятся неотличимыми от изоциренных видов кибероружия). Фактически цифровая ойкумена государства является его дополнительной формой геополитического влияния в новых условиях. Конвергенция пропаганды, объединяющей традиционные и новые алгоритмические приемы политического влияния на массовое сознание, делает такие методы цифровой дипломатии малозаметными и всепроникающими.

Заклучение

Подводя итоги, важно на основании ключевых особенностей выявленных современных проблем в научной литературе, связанных с активным внедрением в актуальную практику алгоритмов искусственного интеллекта, сделать некоторые выводы. Необходимо признать, что среди обозначенных проблем наибольшую значимость представляет проблема алгоритмизации власти. Это ключевой тренд, влияющий на остальные – непрозрачность алгоритмов, манипуляции массовым сознанием и алгоритмическую дискриминацию. Если меняется природа власти, то такая метаморфоза затрагивает все другие социально-политические, культурные и экономические сферы. Внедрение инструментов предиктивной аналитики в сферу государственной политики, перераспределение власти на примере групп менеджеров и сферы политического управления являются доказательством такого феномена.

Вместе с тем важно признать, что несмотря на наличие существенных социально-политических рисков от внедрения алгоритмов, определяемых в современной научной литературе, ситуация не является однозначной. Последствия, эффекты от внедрения алгоритмов искусственного интеллекта в разных странах еще недостаточно изучены. По сути, сохраняется определенный пробел в понимании того, почему в одних странах алгоритмизация власти больше связана с усилением надзорных возможностей и полномочий центральных органов политической власти, а в других такая алгоритмизация больше связана с ростом влияния крупных корпораций и корпоративных органов власти. Возможно, восполнению этого пробела послужат новые исследования, позволяющие не просто выявить социально-политические риски алгоритмов, но и сравнить их специфику в различных странах с учетом существующих там социальных, экономических, мировоззренческих и политико-культурных особенностей.

Список литературы

1. Брунс, А. (2023). *Реальна ли стена фильтров?* Москва: Издат. дом Высш. шк. экономики.
2. Жуков, Д. С. (2020). Искусственный интеллект для общественно-государственного организма: будущее уже стартовало в Китае. *Журнал политических исследований*, 4(2), 70–79. <https://doi.org/10.12737/2587-6295-2020-4-2-70-79>
3. Зубофф, Ш. (2022). *Эпоха надзорного капитализма. Битва за человеческое будущее на новых рубежах власти.* Москва: Изд. Ин-та Гайдара.
4. Ледяев, В. Г. (2001). *Власть: концептуальный анализ.* Москва: РОССПЭН.
5. Лексин, В. Н. (2021). *Искусственный интеллект в экономике, политике и частной жизни: Опыт системной диагностики.* Москва: Ленанд.

6. Ловинк, Г. (2019). *Критическая теория Интернета*. Москва: Ad Marginem: Музей современ. искусства «Гараж».
7. Льюкс, С. (2010). *Власть: Радикальный взгляд*. Москва: Издат. дом Высш. шк. экономики.
8. Киссинджер, Г., Шмидт, Э., Хаттенлокер, Д. (2022). *Искусственный разум и новая эра человечества*. Москва: Альпина ПРО.
9. Миронова Н. Г. (2021). Философское осмысление социальных рисков интеллектуальной автоматизации социального управления. *Цифровой ученый: лаборатория философа*, 4(2), 125–144. <https://doi.org/10.32326/2618-9267-2021-4-2-125-144>
10. Овчинников, А. И. (2020). Безопасность личности и государства в цифровую эпоху: политико-правовой аспект. *Журнал российского права*, (6), 5–21. <https://doi.org/10.12737/jrl.2020.064>
11. Романова, И. Н. (2021). Внедрение технологий искусственного интеллекта: анализ вероятных рисков и возможных угроз. *Материалы Ивановских чтений*, (4), 15–18.
12. Срничек, Н. (2020). *Капитализм платформ*. Москва: Издат. дом Высш. шк. экономики.
13. Талапина, Э. В. (2021). Использование искусственного интеллекта в государственном управлении. *Информационное общество*, (3), 16–22. https://doi.org/10.52605/16059921_2021_03_16
14. Филиппов, А. Ф. (2019). Власть. *Философские науки*, 62(10), 139–159. <https://doi.org/10.30727/0235-1188-2019-62-10-139-159>
15. Фурс, С. П. (2022). Анализируя новый социальный феномен – алгократия. Культура и безопасность, (1), 5–9. <https://doi.org/10.25257/КВ.2022.1.5-9>
16. Aneesh, A. (2006). *Virtual Migration: The Programming of Globalization*. Durham, NC.; London: Duke Univ. Press.
17. Bartlett, B. (2023). The possibility of AI-induced medical manslaughter: Unexplainable decisions, epistemic vices, and a new dimension of moral luck. *Medical Law International*, 23(3), 241–270. <https://doi.org/10.1177/09685332231193944>
18. Beer, D. (2022). The problem of researching a recursive society: Algorithms, data coils and the looping of the social. *Big Data & Society*, 9(2). <https://doi.org/10.1177/20539517221104997>
19. Borgesius, F. J. Z. (2020). Strengthening legal protection against discrimination by algorithms and artificial intelligence. *The International Journal of Human Rights*, 24(10), 1572–1593. <https://doi.org/10.1080/13642987.2020.1743976>
20. Cheney-Lippold, J. (2011). A New Algorithmic Identity: Soft Biopolitics and the Modulation of Control. *Theory Culture & Society*, 28(6), 164–181. <https://doi.org/10.1177/0263276411424420>
21. Elkin-Koren, N. (2020). Contesting algorithms: Restoring the public interest in content filtering by artificial intelligence. *Big Data & Society*, 7(2). <https://doi.org/10.1177/2053951720932296>

22. Erman, E., & Furendal, M. (2022). Artificial Intelligence and the Political Legitimacy of Global Governance. *Political Studies*, 72(2), 421–441. <https://doi.org/10.1177/00323217221126665>
23. Fallis, D. (2021). The Epistemic Threat of Deepfakes. *Philosophy & Technology*, 34, 623–643. <https://doi.org/10.1007/s13347-020-00419-2>
24. Hameleers, M., Meer, T. G. L. A. van der, & Dobber, T. (2022). You Won't Believe What They Just Said! The Effects of Political Deepfakes Embedded as Vox Populi on Social Media. *Social Media + Society*, 8(3). <https://doi.org/10.1177/20563051221116346>
25. Huang, M.-H., & Rust, R. T. (2018). Artificial Intelligence in Service. *Journal of Service Research*, 21(2), 155–172. <https://doi.org/10.1177/1094670517752459>
26. Jarrahi, M. H., Newlands, G., Lee, M. K., Wolf, C. T., Kinder, E., & Sutherland, W. (2021). Algorithmic management in a work context. *Big Data & Society*, 8(2). <https://doi.org/10.1177/20539517211020332>
27. Joyce, K., Smith-Doerr, L., Alegria, S., Bell, S., Cruz, T., Hoffman, S. G., Noble, S. U., & Shestakofsky, B. (2021). Toward a Sociology of Artificial Intelligence: A Call for Research on Inequalities and Structural Change. *Socius Sociological Research for a Dynamic World*, 7(4). <https://doi.org/10.1177/2378023121999581>
28. Jungherr, A. (2023). Artificial Intelligence and Democracy: A Conceptual Framework. *Social Media + Society*, 9(3). <https://doi.org/10.1177/20563051231186353>
29. Kieslich, K., Keller, B., & Starke, C. (2022). Artificial intelligence ethics by design. Evaluating public perception on the importance of ethical design principles of artificial intelligence. *Big Data & Society*, 9(1). <https://doi.org/10.1177/20539517221092956>
30. Kubler, K. (2017). State of urgency: Surveillance, power, and algorithms in France's state of emergency. *Big Data & Society*, 4(2). <https://doi.org/10.1177/2053951717736338>
31. Lepage-Richer, T., & McKelvey, F. (2022). States of computing: On government organization and artificial intelligence in Canada. *Big Data & Society*, 9(2). <https://doi.org/10.1177/20539517221123304>
32. Lorenz, L., Meijer, A., & Schuppan, T. (2021). The Algocracy as a New Ideal Type for Government Organizations: Predictive Policing in Berlin as an Empirical Case. *Information Polity*, 26(1), 71–86. <https://doi.org/10.3233/IP-200279>
33. Lundahl, O. (2022). Algorithmic meta-capital: Bourdieusian analysis of social power through algorithms in media consumption. *Information, Communication & Society*, 25(10), 1440–1455. <https://doi.org/10.1080/1369118X.2020.1864006>
34. McQuillan, D. (2015). Algorithmic states of exception. *European Journal of Cultural Studies*, 18(4–5), 564–576. <https://doi.org/10.1177/1367549415577389>
35. McQuillan, D. (2016). Algorithmic paranoia and the convivial alternative. *Big Data & Society*, 3(2). <https://doi.org/10.1177/2053951716671340>
36. Milivojevic, S. (2022). Artificial intelligence, illegalised mobility and lucrative alchemy of border utopia. *Criminology & Criminal Justice*. <https://doi.org/10.1177/17488958221123855>

37. Nemitz, P. (2018). Constitutional democracy and technology in the age of artificial intelligence. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 376(2133). <https://doi.org/10.1098/rsta.2018.0089>
38. Ó Fathaigh, R., Dobber, T., Borgesius, F. J. Z., & Shires, J. (2021). Microtargeted propaganda by foreign actors: An interdisciplinary exploration. *Maastricht Journal of European and Comparative Law*, 28(6), 856–877. <https://doi.org/10.1177/1023263X211042471>
39. Pajkovic, N. (2022). Algorithms and taste-making: Exposing the Netflix Recommender System's operational logics. *Convergence*, 28(1), 214–235. <https://doi.org/10.1177/13548565211014464>
40. Plantinga, P. (2022). Digital discretion and public administration in Africa: Implications for the use of artificial intelligence. *Information Development*, 40(2), 332–352. <https://doi.org/10.1177/02666669221117526>
41. Rafanelli, L. M. (2022). Justice, injustice, and artificial intelligence: Lessons from political theory and philosophy. *Big Data & Society*, 9(1). <https://doi.org/10.1177/20539517221080676>
42. Reed, E. D. (2022). Truth, Lies and New Weapons Technologies: Prospects for Jus in Silico? *Studies in Christian Ethics*, 35(1), 68–86. <https://doi.org/10.1177/09539468211051240>
43. Seaver, N. (2017). Algorithms as culture: Some tactics for the ethnography of algorithmic systems. *Big Data & Society*, 4(2). <https://doi.org/10.1177/2053951717738104>
44. Seta, G. de (2021). Huanlian, or changing faces: Deepfakes on Chinese digital media platforms. *Convergence*, 27(4), 935–953. <https://doi.org/10.1177/13548565211030185>
45. Sharma, S., & Nijjar, J. S. (2023). Post-racial politics, pre-emption and in/security. *European Journal of Cultural Studies*, 27(2), 275–294. <https://doi.org/10.1177/13675494231168177>
46. Švec, M., & Madleňák, A. (2017). Legal Frameworks for the Phygital Concept. *European Journal of Science and Theology*, 13(6), 209–217.
47. Ulbricht, L. (2020). Scraping the demos. Digitalization, web scraping and the democratic project. *Democratization*, 27(3), 426–442. <https://doi.org/10.1080/13510347.2020.1714595>
48. Vesa, M., & Tienari, J. (2022). Artificial intelligence and rationalized unaccountability: Ideology of the elites? *Organization*, 29(6), 1133–1145. <https://doi.org/10.1177/1350508420963872>

References

1. Aneesh, A. (2006). *Virtual Migration: The Programming of Globalization*. Durham, NC.; London: Duke Univ. Press.

2. Bartlett, B. (2023). The possibility of AI-induced medical manslaughter: Unexplainable decisions, epistemic vices, and a new dimension of moral luck. *Medical Law International*, 23(3), 241–270. <https://doi.org/10.1177/09685332231193944>
3. Beer, D. (2022). The problem of researching a recursive society: Algorithms, data coils and the looping of the social. *Big Data & Society*, 9(2). <https://doi.org/10.1177/20539517221104997>
4. Borgesius, F. J. Z. (2020). Strengthening legal protection against discrimination by algorithms and artificial intelligence. *The International Journal of Human Rights*, 24(10), 1572–1593. <https://doi.org/10.1080/13642987.2020.1743976>
5. Bruns, A. (2023). *Real'na li stena fil'trov?* [Are Filter Bubbles Real?]. Moscow: Izdat. dom Vyssh. shk. ekonomiki.
6. Cheney-Lippold, J. (2011). A New Algorithmic Identity: Soft Biopolitics and the Modulation of Control. *Theory Culture & Society*, 28(6), 164–181. <https://doi.org/10.1177/0263276411424420>
7. Elkin-Koren, N. (2020). Contesting algorithms: Restoring the public interest in content filtering by artificial intelligence. *Big Data & Society*, 7(2). <https://doi.org/10.1177/2053951720932296>
8. Erman, E., & Furendal, M. (2022). Artificial Intelligence and the Political Legitimacy of Global Governance. *Political Studies*, 72(2), 421–441. <https://doi.org/10.1177/00323217221126665>
9. Fallis, D. (2021). The Epistemic Threat of Deepfakes. *Philosophy & Technology*, 34, 623–643. <https://doi.org/10.1007/s13347-020-00419-2>
10. Filippov, A. F. (2019). Vlast' [Power]. *Filosofskie nauki*, 62(10), 139–159. <https://doi.org/10.30727/0235-1188-2019-62-10-139-159>
11. Furs, S. P. (2022). Analiziruya novyy sotsial'nyy fenomen – algokratiya [Analysing new social phenomenon – algocracy]. *Kul'tura i bezopasnost'*, (1), 5–9. <https://doi.org/10.25257/KB.2022.1.5-9>
12. Hameleers, M., Meer, T. G. L. A. van der, & Dobber, T. (2022). You Won't Believe What They Just Said! The Effects of Political Deepfakes Embedded as Vox Populi on Social Media. *Social Media + Society*, 8(3). <https://doi.org/10.1177/20563051221116346>
13. Huang, M.-H., & Rust, R. T. (2018). Artificial Intelligence in Service. *Journal of Service Research*, 21(2), 155–172. <https://doi.org/10.1177/1094670517752459>
14. Jarrahi, M. H., Newlands, G., Lee, M. K., Wolf, C. T., Kinder, E., & Sutherland, W. (2021). Algorithmic management in a work context. *Big Data & Society*, 8(2). <https://doi.org/10.1177/20539517211020332>
15. Joyce, K., Smith-Doerr, L., Alegria, S., Bell, S., Cruz, T., Hoffman, S. G., Noble, S. U., & Shestakofsky, B. (2021). Toward a Sociology of Artificial Intelligence: A Call for Research on Inequalities and Structural Change. *Socius Sociological Research for a Dynamic World*, 7(4). <https://doi.org/10.1177/2378023121999581>
16. Jungherr, A. (2023). Artificial Intelligence and Democracy: A Conceptual Framework. *Social Media + Society*, 9(3). <https://doi.org/10.1177/20563051231186353>

17. Kieslich, K., Keller, B., & Starke, C. (2022). Artificial intelligence ethics by design. Evaluating public perception on the importance of ethical design principles of artificial intelligence. *Big Data & Society*, 9(1). <https://doi.org/10.1177/20539517221092956>
18. Kissinger, H. A., Schmidt, E. & Huttenlocher D. (2022). *Iskusstvennyy razum i novaya era chelovechestva* [The Age of AI: And Our Human Future]. Moscow: Al'pina PRO.
19. Kubler, K. (2017). State of urgency: Surveillance, power, and algorithms in France's state of emergency. *Big Data & Society*, 4(2). <https://doi.org/10.1177/2053951717736338>
20. Ledyayev, V.G. (2001). *Vlast': kontseptual'nyy analiz* [Power: a conceptual analysis]. Moscow: ROSSPEN.
21. Leksin, V.N. (2021). *Iskusstvennyy intellekt v ekonomike, politike i chastnoy zhizni: Opyt sistemnoy diagnostiki* [Artificial intelligence in economics, politics and private life: Experimental diagnostic systems]. Moscow: Lenand.
22. Lepage-Richer, T., & McKelvey, F. (2022). States of computing: On government organization and artificial intelligence in Canada. *Big Data & Society*, 9(2). <https://doi.org/10.1177/20539517221123304>
23. Lorenz, L., Meijer, A., & Schuppan, T. (2021). The Algocracy as a New Ideal Type for Government Organizations: Predictive Policing in Berlin as an Empirical Case. *Information Polity*, 26(1), 71–86. <https://doi.org/10.3233/IP-200279>
24. Lovink, G. (2019). *Kriticheskaya teoriya Interneta* [Critical Internet Theory]. Moscow: Ad Marginem: Muzei sovremen. iskusstva «Garazh».
25. Lukes, St. (2010). *Vlast': Radikal'nyy vzglyad* [Power. A Radical View]. Moscow: Izdat. dom Vyssh. shk. ekonomiki.
26. Lundahl, O. (2022). Algorithmic meta-capital: Bourdieusian analysis of social power through algorithms in media consumption. *Information, Communication & Society*, 25(10), 1440–1455. <https://doi.org/10.1080/1369118X.2020.1864006>
27. McQuillan, D. (2015). Algorithmic states of exception. *European Journal of Cultural Studies*, 18(4–5), 564–576. <https://doi.org/10.1177/1367549415577389>
28. McQuillan, D. (2016). Algorithmic paranoia and the convivial alternative. *Big Data & Society*, 3(2). <https://doi.org/10.1177/2053951716671340>
29. Milivojevic, S. (2022). Artificial intelligence, illegalised mobility and lucrative alchemy of border utopia. *Criminology & Criminal Justice*. <https://doi.org/10.1177/17488958221123855>
30. Mironova, N.G. (2021). *Filosofskoe osmyslenie sotsial'nykh riskov intellektual'noy avtomatizatsii sotsial'nogo upravleniya* [Philosophical consideration of social risks of intellectual automation of social management]. *Tsifrovoy uchenyy: laboratoriya filosafo*, 4(2), 125–144. <https://doi.org/10.32326/2618-9267-2021-4-2-125-144>
31. Nemitz, P. (2018). Constitutional democracy and technology in the age of artificial intelligence. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 376(2133). <https://doi.org/10.1098/rsta.2018.0089>

32. Ó Fathaigh, R., Dobber, T., Borgesius, F. J. Z., & Shires, J. (2021). Microtargeted propaganda by foreign actors: An interdisciplinary exploration. *Maastricht Journal of European and Comparative Law*, 28(6), 856–877. <https://doi.org/10.1177/1023263X211042471>
33. Ovchinnikov, A. I. (2020). Bezopasnost' lichnosti i gosudarstva v tsifrovuyu epokhu: politiko-pravovoy aspect [Individual and state security in the digital age: political and legal aspect]. *Zhurnal rossiyskogo prava*, (6), 5–21. <https://doi.org/10.12737/jrl.2020.064>
34. Pajkovic, N. (2022). Algorithms and taste-making: Exposing the Netflix Recommender System's operational logics. *Convergence*, 28(1), 214–235. <https://doi.org/10.1177/13548565211014464>
35. Plantinga, P. (2022). Digital discretion and public administration in Africa: Implications for the use of artificial intelligence. *Information Development*, 40(2), 332–352. <https://doi.org/10.1177/02666669221117526>
36. Rafanelli, L. M. (2022). Justice, injustice, and artificial intelligence: Lessons from political theory and philosophy. *Big Data & Society*, 9(1). <https://doi.org/10.1177/205395172211080676>
37. Reed, E. D. (2022). Truth, Lies and New Weapons Technologies: Prospects for Jus in Silico? *Studies in Christian Ethics*, 35(1), 68–86. <https://doi.org/10.1177/09539468211051240>
38. Romanova, I. N. (2021). Vnedrenie tekhnologiy iskusstvennogo intellekta: analiz veroyatnykh riskov i vozmozhnykh ugroz [Introduction of artificial intelligence technologies: analysis of probable risks and possible threats]. *Materialy Ivanovskikh chteniy*, (4), 15–18.
39. Seaver, N. (2017). Algorithms as culture: Some tactics for the ethnography of algorithmic systems. *Big Data & Society*, 4(2). <https://doi.org/10.1177/2053951717738104>
40. Seta, G. de (2021). Huanlian, or changing faces: Deepfakes on Chinese digital media platforms. *Convergence*, 27(4), 935–953. <https://doi.org/10.1177/13548565211030185>
41. Sharma, S., & Nijjar, J. S. (2023). Post-racial politics, pre-emption and in/security. *European Journal of Cultural Studies*, 27(2), 275–294. <https://doi.org/10.1177/13675494231168177>
42. Srnicek, N. (2020). *Kapitalizm platform* [Platform Capitalism]. Moscow: Izdat. dom Vyssh. shk. ekonomiki.
43. Švec, M., & Madleňák, A. (2017). Legal Frameworks for the Phygital Concept. *European Journal of Science and Theology*, 13(6), 209–217.
44. Talapina, E. V. (2021). Ispol'zovanie iskusstvennogo intellekta v gosudarstvennom upravlenii [The use of artificial intelligence in public administration]. *Informatsionnoe obshchestvo*, (3), 16–22. https://doi.org/10.52605/16059921_2021_03_16
45. Ulbricht, L. (2020). Scraping the demos. Digitalization, web scraping and the democratic project. *Democratization*, 27(3), 426–442. <https://doi.org/10.1080/13510347.2020.1714595>

46. Vesa, M., & Tienari, J. (2022). Artificial intelligence and rationalized unaccountability: Ideology of the elites? *Organization*, 29(6), 1133–1145. <https://doi.org/10.1177/1350508420963872>

47. Zhukov, D. S. (2020). Iskusstvennyy intellekt dlya obshchestvenno-gosudarstvennogo organizma: budushchee uzhe startovalo v Kitae [Artificial intelligence for society and government: the future has already begun in China]. *Zhurnal politicheskikh issledovaniy*, 4(2), 70–79. <https://doi.org/10.12737/2587-6295-2020-4-2-70-79>

48. Zuboff, Sh. (2022). *Epokha nadzornogo kapitalizma. Bitva za chelovecheskoe budushchee na novykh rubezhakh vlasti* [The Age of Surveillance Capitalism. The Fight for a Human Future at the New Frontier of Power]. Moscow: Izd. In-ta Gaydara.

Информация об авторах

Сергей Владимирович Володенков, доктор политических наук, главный научный сотрудник научно-проектного отдела Научно-инновационного управления Государственного академического университета гуманитарных наук, профессор кафедры государственной политики факультета политологии Московского государственного университета имени М.В. Ломоносова, Москва, Россия, ORCID: <https://orcid.org/0000-0003-2928-6068>, e-mail: s.v.cyber@gmail.com

Сергей Николаевич Федорченко, доктор политических наук, главный научный сотрудник научно-проектного отдела Научно-инновационного управления Государственного академического университета гуманитарных наук, доцент кафедры истории и теории политики факультета политологии Московского государственного университета имени М.В. Ломоносова, Москва, Россия, ORCID: <https://orcid.org/0000-0001-6563-044X>, e-mail: s.n.fedorchenko@mail.ru

Information about the authors

Sergey Vladimirovich Volodenkov, Doctor of Sciences (Political Sciences), Chief Researcher for the Research and Design Department of the Scientific and Innovation Department, State Academic University for the Humanities, Professor of the Department of Public Policy, Faculty of Political Science, Lomonosov Moscow State University, Moscow, Russia, ORCID: <https://orcid.org/0000-0003-2928-6068>, e-mail: s.v.cyber@gmail.com

Sergey Nikolaevich Fedorchenko, Doctor of Sciences (Political Sciences,) Chief Researcher for the Research and Design Department of the Scientific and Innovation Department, State Academic University for the Humanities, Associate Professor of the Department of Department of History and Theory of Politics, Faculty of Political Science, Lomonosov Moscow State University, Moscow, Russia, ORCID: <https://orcid.org/0000-0001-6563-044X>, e-mail: s.n.fedorchenko@mail.ru