

---

## КОНТЕКСТ

---

**В.С. АВДОНИН, В.Л. СИЛАЕВА\***

### **НЕЙРОСЕТИ НОВОГО ПОКОЛЕНИЯ В КОНТЕКСТЕ ТЕХНОЛОГИЙ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА, ФИЛОСОФИИ И СОЦИАЛЬНО-ПОЛИТИЧЕСКИХ НАУК**

*Аннотация.* В статье процессы влияния технологических изменений на политическую науку исследуются на примере технологии нейросетей нового поколения (чат GPT и др.), вызвавших большой резонанс в медиасфере на рубеже 2022–2023 гг. Анализу собственно этого влияния авторы предпосылают краткий историко-научный обзор возникновения и развития нейросетевых технологий и тех их особенностей и возможностей, которые нашли воплощение в системах нового поколения. Само появление этих систем и демонстрация их возможностей в различных областях и сферах деятельности рассматривается не только как важный этап в развитии компьютерных технологий, прокладывающий путь к созданию «сильного» искусственного интеллекта (ИИ), но и как событие в мировом технологическом развитии, способное влиять на самые разные области жизни и деятельности человечества, что показывает беспрецедентная реакция на него в мировой политике (вплоть до мирового саммита и СБ ООН). В статье также освещается философский контекст, в котором предпочтительно вести осмысление этих новых технологий и их возможностей. Таковым может быть область философии информации и ее трактовка в трудах философа Лючано Флориди, обла-

---

\* **Авдонин Владимир Сергеевич**, доктор политических наук, ведущий научный сотрудник Отдела политической науки, Институт научной информации по общественным наукам (ИНИОН) РАН (Москва, Россия), e-mail: avdoninvla@mail.ru; **Силаева Виктория Леонидовна**, кандидат философских наук, научный сотрудник Отдела политической науки, Институт научной информации по общественным наукам (ИНИОН) РАН (Москва, Россия), e-mail: seva@starlink.ru

дающая рядом преимуществ перед другими подходами. В анализе влияния этих новых технологий на социально-политические науки авторы исходят из концепции гибридных междисциплинарных научных областей, которые возникают на стыках наук и обладают особенностями, отличающими их от традиционных полей этих наук. В статье рассматриваются гибридные области вычислительных социальных наук, развивающиеся на стыке компьютерных и социальных наук. На примере политической науки авторы предлагают рассмотреть два сценария («адаптационный» и «ревизионистский») влияния на нее появившихся технологий ИИ нового поколения, каждый из которых уже находит некоторое проявление в ряде новейших публикаций политологов.

*Ключевые слова:* нейронные сети; искусственный интеллект; большие языковые модели; философия ИИ; философия информации; Лючано Флориди; инфосфера; социальные науки; гибридные области; вычислительные социальные науки; традиционная политология; реформа политических теорий.

*Для цитирования:* Авдонин В.С., Силаева В.Л. Нейросети нового поколения в контексте технологий искусственного интеллекта, философии и социально-политических наук // Политическая наука. – 2023. – № 4. – С. 127–154. – DOI: <http://www.doi.org/10.31249/poln/2023.04.05>

## Введение

Исследования и практическое применение искусственного интеллекта (ИИ) получили новый мощный импульс в наступившем третьем десятилетии XXI в. Пока трудно оценить в полной мере все его масштабы, аспекты и тем более последствия для дальнейшего развития различных областей науки и практики. Но среди лежащих на поверхности причин и условий можно обратить внимание на появление и быстрое распространение в широком публичном доступе новых сервисов ИИ, связанных с нейросетями нового поколения. Оно вызвало в последний год буквально взрыв разного рода оценок, прогнозов и действий как в кругах специалистов разного профиля (не только связанных с ИИ), так и среди широкой публики и даже в мировой политике. Возможно, что мы здесь имеем дело с событием, способным повлиять на ход мирового технологического развития и затрагивающим многие стороны жизни, в том числе и область социально-политических наук.

В нашем подходе мы исключаем жесткое противопоставление социально-политических и технико-технологических аспектов этой проблематики, а, напротив, основываемся на принципе их взаимодействия и взаимовлияния. Исходя из этого, методологиче-

скому анализу потенциала влияния нового поколения нейросетевого интеллекта на социально-политические науки должно предшествовать его хотя бы краткое технологическое рассмотрение, позволяющее лучше понять и проблемы развития этих новых технологий и характер, и потенциал их влияния на познавательные, социальные и политические процессы. Мы полагаем сначала кратко остановиться на истории появления и развития нейросетевых технологий, включая условия их перехода к новому современному этапу. Затем рассмотрим сам этот этап сквозь призму открывающихся возможностей, сфер применения, перспектив развития, а также рисков этих технологий, вызывающих столь бурную общественную реакцию. Далее мы полагаем поместить их в контекст современной философской проблематики, сопоставляя оптику и особенности двух смежных контекстных подходов – философии искусственного интеллекта и философии информации. И, наконец, перейти к оценке влияния этих технологий в области социально-политических наук, используя науковедческую концепцию междисциплинарных трансферов и междисциплинарных (гибридных) научных областей. И уже в этом контексте проанализировать их влияние на политическую науку.

Метод исследования, используемый в данной статье, преимущественно эмпирико-описательный, историко-технологический, историко-философский, историко-научный, дополненный в ряде случаев элементами науковедческого и теоретико-категориального анализа.

### **Возникновение и развитие технологий нейронных сетей**

Большинство новых сервисов ИИ, вызвавших столь бурную реакцию в последнее время, в основном базируется на технологиях искусственных нейронных сетей, разработка которых началась еще в середине прошлого века. В истории компьютерных технологий появление нейронных сетей и нейрокомпьютеров обычно связывается с применением к области разработки компьютерной техники тех знаний, идей и подходов, которые были получены при изучении естественных нейронов и нейронных сетей, обнаруженных в мозгу человека и животных. Уподобление работы компьютеров работе нейросетей мозга ведет в этом случае к переворачи-

ванию известной «компьютерной» метафоры, сравнивающей работу мозга с работой компьютеров. На этой почве возникли и определенные междисциплинарные связи между биологической нейронаукой и исследованиями искусственного интеллекта как отрасли компьютерных наук (область нейроинформатики), составившими часть еще более обширного поля так называемых когнитивных наук, сформировавшегося несколько позднее.

Как часть компьютерных наук, исследование нейросетей за почти восемь десятилетий прошло значительный путь развития. Сегодня мы не только значительно лучше знаем, что происходит в накапливающих, сохраняющих, обрабатывающих и преобразующих информацию искусственных нейросетях, но и можем их моделировать, обучать, совершенствовать, использовать и управлять ими. Ключевыми вехами на этом пути считаются: создание математической модели искусственного нейрона и создание первых алгоритмов обучения искусственной нейронной сети (1940-е годы); создание первого нейрокомпьютера (перцептрона), моделирующего работу нейронной сети мозга (1950-е годы); создание многослойных нейронных сетей и алгоритмов обучения по правилам «градиентного спуска» (1960-е годы)<sup>1</sup>; разработка нейронных сетей, обладающих памятью и обучающихся по правилам «обратного распространения (ошибки)» (1970-е годы); теоретическое обоснование работы и создание нейросетей с различной архитектурой и различными типами алгоритмов обучения (1980-е годы); рост вычислительной мощности компьютеров и создание сетей и алгоритмов «глубокого обучения», работающих с большими объемами данных (1990–2000-е годы); разработка и внедрение технологии «больших языковых моделей» и разнообразных интернет-сервисов и программ, работающих на основе искусственных нейронных сетей (2010-е годы). [Russel, Norvig, 2021]

Технология нейросетей является сегодня важной и наиболее быстро развивающейся составной частью области технологий ис-

---

<sup>1</sup> Важным моментом, повлиявшим на развитие нейросетей в этот период, стали работы [Minsky, Pappert, 1969], указавшие на трудности корректного выполнения в этих сетях операции логической дизъюнкции (операции XOR – «исключающего ИЛИ»). Считается, что это тормозило развитие нейросетей в 1970-е годы, но было преодолено к концу 1980-х годов, когда были найдены средства корректного обеспечения этой операции в нейросетях с помощью статистических методов.

кусственного интеллекта. В самой этой области обычно условно выделяются две большие группы подходов или вычислительных формализмов: логико-символические и сетевые (коннекционистские). Первые представляют собой подходы, ориентирующиеся на сложные логические формализмы и создание ИИ «высокого уровня», способного решать, прежде всего, сложные интеллектуальные задачи (проводить сложные вычисления, решать трудные головоломки, доказывать сложные теоремы и т.д.), и уже во вторую очередь пытаться заниматься решением более простых задач. Вторая группа, напротив, предлагает при строительстве ИИ использовать самые простые, элементарные вычислительные формализмы, доступные для клеточных автоматов, объединяемые в многослойные сети, что позволяет переходить от простых решений к сложным, опираясь на обучение и моделирование естественного интеллекта (естественные языки, биологические нейросети и др.). История разработки технологий ИИ с середины прошлого века позволяет проследить своего рода конкуренцию этих подходов [Charniak, McDermott, 1985; Bringsjord, Govindarajulu, 2022].

На первых порах определенным приоритетом пользовались подходы первой группы, которые в основном ассоциировались с так называемым символическим моделированием ИИ, базирующемся на высокоуровневом символическом представлении задач, символической логике, семантических базах знаний и алгоритмах поиска. Символическая (семиотическая) парадигма в технологиях ИИ считалась доминирующей с 50-х до середины 90-х годов XX в. Она демонстрировала как важные достижения, так и определенные проблемы, связанные в основном с «жесткостью» и «неподатливостью» воплощенных в ней логик. Исследователи отмечают, что в развитии этой парадигмы технологий ИИ возникали как периоды подъема, так и «зимы», т.е. периоды, когда успехи и надежды сменялись разочарованием и стагнацией [Bringsjord, Govindarajulu, 2022].

На этом фоне развитие технологии нейросетей, которые входили в группу технологий, «восходящих» от простых операций к сложным через обучение, выглядело маргинальным. Оно тоже прошло периоды расцвета и стагнации, но в целом по концентрации финансовых и интеллектуальных ресурсов заметно уступало символической или классической парадигме технологий ИИ. Это положение изменилось в середине 2000-х годов, когда в технологиях развития нейросетей начался новый подъем. Современные

исследователи истории ИИ, ретроспективно оценивая причины этого явления, отмечают, что новому взлету нейросетей в этот период способствовало сочетание нескольких факторов [Russel, Norvig, 2021].

Во-первых, важным обстоятельством стало наращивание вычислительной мощности компьютеров за счет широкого внедрения новой аппаратной базы (микропроцессоры, графические процессоры, микропроцессорные системы на кристалле и др.), что повлекло как заметное снижение стоимости и энергозатрат на единицу вычислений, так и миниатюризацию аппаратного обеспечения, а также рост доступности всей этой техники. Во-вторых, растущая компьютеризация многих областей жизни и создание глобальных компьютерных сетей вело к накоплению огромных массивов информации («больших данных»), доступных для компьютерной обработки. Наконец, в-третьих, с учетом этих новых возможностей были предложены разработки и усовершенствования программ обработки информации, многие из которых были связаны с нейросетевой парадигмой развития ИИ.

В отличие от символической парадигмы ИИ, основанной на методах логически строгого и корректного оперирования символами формальных систем, нейросетевая парадигма нашла возможность усиления своей вычислительной мощности с помощью статистических вычислений. Именно включение в нейросети методов статистической обработки большого массива поступающих данных создало для них определенные преимущества в условиях быстрого роста производительности компьютеров и накопления баз «больших данных». Ориентированные на нее компьютерные технологии могли лучше справиться с информационно-вычислительной обработкой процессов повышенной сложности и неопределенности, которые труднее поддавались обработке с помощью символических вычислений. В одном из исследований, объясняя преимущества сетевой парадигмы, авторы выразились так: «Общая тенденция в вычислительном оборудовании отдает предпочтение алгоритмам, способным выполнять большое количество простых операций, которые не так уж зависят друг от друга, по сравнению с меньшим количеством, но сложных и запутанных операций» [Bringsjord, Govindarajulu, 2022].

Помимо освоения статистических вычислений, среди которых особенно важное место занимали методы байесовской стати-

стики, развитие нейросетевой парадигмы ИИ базировалось на интегрировании знаний таких логико-математических направлений, как нечеткая логика (fuzzy logic), эволюционные, генетические и эвристические алгоритмы, естественные вычисления, роевой интеллект, исследования естественных языков и др. На ставшей популярной в современных исследованиях ИИ карте знаний, технологий и парадигм ИИ нейросети занимают место в центре этого поля – между парадигмами символических и субсимволических вычислений, основываясь на вероятностно-статистической вычислительной парадигме машинного обучения [Corea, 2019].

В современных технических устройствах парадигма сетевого ИИ, как правило, интегрирована в той или иной мере с системами логико-символических вычислений. И в этом смысле большинство этих устройств с точки зрения воплощения в них вычислительных формализмов являются гибридными. Но ядром и фундаментом устройств современного сетевого ИИ является, разумеется, нейросеть, осуществляющая вычисления с помощью вероятностно-статистических методов.

### **Появление нейронных сетей нового поколения**

Нейросети нового поколения относятся к классу систем генеративного искусственного интеллекта, которые способны производить (генерировать) понятные для пользователей тексты, изображения, программные коды и другие продукты, а также уточнять (улучшать) их по определенным обучающим подсказкам. В отличие от других моделей генеративного ИИ, ограниченных производством продуктов по специализированным задачам, новые нейросети представляют собой модели, способные генерировать значительно более качественный и разнообразный контент по широкому кругу вопросов, областей и тематик. Скачок в повышении качества и разнообразия контента происходит в этих сетях потому, что они реализуют так называемые «Большие языковые модели» (LLM – Large language model) или модели, состоящие из огромного числа (миллиардов) параметров (или «весов» нейронов в слоях сети) и обученные в режиме «глубокого обучения» и «обучения без учителя» на огромном (миллиарды слов / токенов) корпусе

текстов<sup>1</sup>. Осуществлять огромный объем вычислительных операций за короткое время сетям позволяет сетевая архитектура «трансформер», выполняющая множество вычислений параллельно. Аббревиатура самого известного семейства этих сетей – GPT – так и расшифровывается – Generative Pre-trained Transformer (генеративный предварительно обученный преобразователь).

Большие языковые модели стали воплощаться в технологии искусственных нейронных сетей с конца 2010-х годов и достигать уровня широкого коммерческого применения и выхода на рынки в 2021–2023 гг. Особенно шумным стал выход на рынок нейросетей ChatGPT 3 и ChatGPT 4 в конце 2022 – начале 2023 г., произведенных компанией Open AI. Он вызвал беспрецедентный взрывной рост числа пользователей этим сервисом по всему миру (до 100 млн за 2 месяца)<sup>2</sup>.

О возможностях этих сетей и областях применения уже имеется много публикаций как рекламного, так и аналитического характера, которые продолжают пополняться все новыми «лайфхаками» (полезными советами по применению). Кратко можно сказать, что они используются для написания сравнительно небольших текстов, а также реферирования (краткого изложения) и редактирования более объемных текстов, для перевода текстов с языка на язык, для поиска и анализа разнообразной информации, для написания и улучшения программных кодов, для обсуждения в диалоговом режиме различных вопросов, в том числе и в целях оптимизации решений. В новой модели (ChatGPT 4) также заложена «мульти-modalность», т.е. возможность преобразовывать тексты в изображения, создавать, редактировать и оптимизировать визуальные изображения. Об областях применения глава компании Open AI Сэм Альтман сказал следующее. На первое место он поставил сферу науки и образования, где возможности новых сетей проявятся наиболее заметно и окажут наибольшее влияние. Также он указал на область компьютерного программирования, область аналитики и прогнозирования (экономической, социальной, медицинской, коммерческой), сферу СМИ, некоторых твор-

---

<sup>1</sup> Best language models and their implications // OpenAI. – Mode of access: <https://openai.com/research/better-language-models> (accessed: 20.07.2023).

<sup>2</sup> Altman S. The Future of AI and ChatGPT: Interview with Sam Altman, CEO of OpenAI // Youtube. – 13.07.2023. – Mode of access: <https://www.youtube.com/watch?v=aUrQ7CcNwM> (accessed: 20.07.2023).

ческих профессий и сферу досуга и безопасности. Делая обобщение социально-политического характера, он решился отметить, что зримое влияние ИИ на человечество начинается с довольно высоких по статусу областей и профессий, оставляя менее затронутыми области с относительно низким социальным статусом и престижем, что может создать определенные проблемы и риски<sup>1</sup>.

Относительно того, как новые нейросети могут применяться и влиять на сферу политики, мы решили обратиться к самой нейросети. Вот, что она (ChatGPT, общедоступная версия) ответила авторам (ответ был сгенерирован менее чем за 1 минуту).

«1. Анализ общественного мнения: нейросети могут использоваться для анализа общественного мнения по поводу того или иного кандидата или той или иной повестки. Это может помочь политическим кампаниям понять, какие идеи и сообщения более эффективны для привлечения избирателей.

2. Прогнозирование результатов выборов: нейросети могут использоваться для прогнозирования результатов выборов на основе данных о предыдущих выборах, опросах общественного мнения и других факторах.

3. Создание персонализированных сообщений: нейросети могут использоваться для создания персонализированных сообщений для избирателей на основе их профиля и интересов. Это может помочь политическим кампаниям увеличить эффективность своих сообщений и привлечь больше поддержки.

4. Обнаружение фейковых новостей: нейросети могут использоваться для обнаружения фейковых новостей и дезинформации, которые могут повлиять на выборы. Это может помочь избирателям получать более точную информацию и сделать более осознанный выбор.

5. Прогнозирование рисков: нейросети могут использоваться для прогнозирования рисков в политической сфере, таких как экономические кризисы, конфликты и террористические угрозы. Это может помочь правительствам и политическим лидерам прини-

---

<sup>1</sup> Altman S. The Future of AI and ChatGPT: Interview with Sam Altman, CEO of OpenAI // Youtube. – 13.07.2023. – Mode of access: <https://www.youtube.com/watch?v=aUrQ7CcCnwM> (accessed: 20.07.2023).

мать более обоснованные решения и уменьшить риски для общества<sup>1</sup>».

Приведенный текст – результат работы далеко не самой продвинутой версии сети ChatGPT. К тому же этот текст – лишь один пример и не отражает всего спектра возможностей работы с этим чатом (возможности задавать новые, уточняющие, переформулированные вопросы<sup>2</sup>). Между тем с марта 2023 г. уже работает новая сеть ChatGPT 4, обладающая более широкими возможностями. А в уже упомянутом интервью Альтмана сообщается о подготовке его компанией трех новых, еще более совершенных моделей – GPT 5, GPT 6 и GPT 7. Есть данные и о подготовке к выпуску усовершенствованных нейросетей от других крупных компаний. И то, что их возможности существенно возрастут, уже не вызывает сомнений.

Здесь важно отметить, что в ходе изучения развития нейросетей была установлена четкая корреляция между способностями (возможностями) новых нейросетей и технико-технологическими параметрами реализуемых в них схемах / моделях функционирования. Сама по себе эта связь фиксировалась и раньше, и было хорошо известно, что возможности сети растут с ростом ее параметрической мощности. Но то, что будет происходить с сетью при возрастании ее параметров на порядки, оставалось неясным. Нейросети нового поколения определенно разрешают этот вопрос – многократный рост параметров сети ведет к качественным скачкам в спектре ее возможностей. В ставшей знаменитой анимированной диаграмме с блога Google<sup>3</sup> хорошо показано, как возрастают возможности сети в зависимости от роста ее параметров. Если при 10 миллиардах параметров сеть способна отвечать лишь на примитивные вопросы, выполнять простые арифметические дей-

---

<sup>1</sup> Текст, сгенерированный в чате с ChatGPT, общедоступная версия.

<sup>2</sup> Мы также задавали вопросы, предполагающие связь политики с другими областями, и получали ответы. Например, об оптимизации бюджета политической партии (связная область – нейросети в финансовой аналитике), проверка личности избирателей (связная область – технология распознавания лиц), определение проблемных областей в политике здравоохранения (связная область – аналитика статистики здравоохранения) и т.д.

<sup>3</sup> Pathways Language Model (PaLM): Scaling to 540 Billion Parameters for Breakthrough Performance. – 04.2022. – Mode of access: <https://ai.googleblog.com/2022/04/pathways-language-model-palm-scaling-to.html> (accessed: 20.07.2023).

ствия и находить грубый эквивалент при переводе, то при 200 млрд (это примерный уровень GPT 3) она не только значительно качественнее реализует все уже названные способности, но и приобретает ряд новых – составление программных кодов, рассуждения на уровне логики и здравого смысла, семантический анализ языка, «понимание» пословиц и поговорок, формирование общих знаний о мире и др. При 560 млрд параметров (на этом заканчивается анимация Google) эти способности качественно совершенствуются и намечаются «зародыши» и «побеги» все новых способностей.

По мнению таких экспертов как Билл Гейтс, Сэм Альтман, Рэй Курцвэйл и др., все это может означать, что путь к созданию «сильного» ИИ фактически открыт и в ближайшие годы он будет непременно создан просто посредством сравнительно небольшого (в 3–5 раз) наращивания параметрической мощности уже существующих нейронных сетей. Технологический комплекс для создания такого ИИ, включающий программное, аппаратное, архитектурное обеспечение функционирования и обучения нейросетей, уже сложился и активно развивается. Среди его ключевых компонентов, помимо дальнейшего развития «больших языковых моделей», называют совершенствование архитектуры сетей-трансформеров, технологий «глубокого обучения», параллельных вычислений, новые виды графических процессоров. Наблюдатели компьютерной отрасли прогнозируют в ближайшее десятилетие (2022–2032) существенный рост (более чем в 30 раз) доходов во всех ее сегментах, связанных с генеративным (нейросетевым) ИИ<sup>1</sup>.

Не удивительно, что фактически происходящий на наших глазах поворот к созданию «сильного» ИИ вызывал вспышку озабоченности по поводу недостаточной контролируемости этого процесса. И дополнительным фактором здесь явился беспрецедентный ажиотаж вокруг взрывной популярности первых же пороговых моделей «сильного» ИИ в лице новых сервисов от Open AI. Для многих стало ясно, что пришествие «сильного» ИИ не только началось, но и встречает высокий спрос и готовность и в обществе,

---

<sup>1</sup> Рынок генеративного ИИ на горизонте 10 лет вырастет в 30 раз Bloomberg // RB.RU. – 06.06.2023. – Режим доступа: <https://rb.ru/news/generative-ai-market/> (дата посещения: 12.05.2023).

и в самой компьютерной отрасли к дальнейшему технологическому рывку в этом направлении.

В авангарде групп, забивших тревогу, парадоксальным образом оказались видные функционеры мировой компьютерной отрасли и даже часть самих разработчиков ИИ, в основном сотрудничающих с организацией Future of Life («Будущее жизни»)¹. Именно с их письма в марте 2023 г. с призывом о принятии мер контроля за развитием ИИ (и приостановки до этого практических работ по данному направлению) начался информационный шум в мировых СМИ. На первых порах его могли связывать с конкурентной борьбой самих компаний. Но вскоре стало ясно, что озабоченности присутствуют не только в среде профильных специалистов и ученых, но и во властно-политических и в широких общественных кругах многих развитых стран. Весной 2023 г. тема контроля ИИ затрагивалась на саммите «Большой семерки» в Хиросиме, а в июле, по инициативе Великобритании, прошло специальное заседание Совета Безопасности ООН с выступлением генерального секретаря ООН по проблемам ИИ.<sup>2</sup> До конца этого года намечено проведение мирового саммита в Лондоне по проблемам ИИ, подготовка которого поручена Великобритании.

Некоторые эксперты полагают, что регулирование технологий нейросетей может развиваться по модели, сходной с регулированием интернета, т.е. путем применения трех видов законов и правил: а) национальных законов, созданных до интернета, но ставших использоваться и по отношению к деятельности в интернете; б) законов, созданных и относящиеся к непосредственной деятельности в интернете; в) правил коммуникации в интернет-сообществах, введенных их модераторами или создателями. Сюда же можно добавить и международные конвенции в области интернета, выработанные международным сообществом. И движение в этом направлении, кажется, уже начинается. В некоторых странах уже принимаются или обсуждаются решения о блокировке сети ChatGPT, появляются иски к создателям сетей в связи с незаконным

---

<sup>1</sup> См. открытое письмо на сайте организации Future of Life. – Режим доступа: <https://futureoflife.org/open-letter/pause-giant-ai-experiments/> (дата посещения: 12.05.2023).

<sup>2</sup> Генсек ООН – об искусственном интеллекте: «Это еще только начало...» // Новости ООН. – 18.07.2023. – Режим доступа: <https://news.un.org/ru/story/2023/07/1442977> (дата посещения: 20.07.2023).

использованием в их обучении авторского контента, ряд ограничительных правил на работу сетей вводят и сами их создатели и владельцы, а также некоторые ИТ-компании. Принятие международной конвенции также уже, видимо, не за горами.

В то же время полагать, что регулирование генеративных сетей будет полностью совпадать с нынешней моделью регулирования интернета, возможно, не стоит. Даже специалисты, не говоря уже о футурологах и фантастах, прогнозируют появление в связи с новыми нейросетями множества ситуаций и практик, для регулирования которых опыт контроля интернета будет вряд ли применим. Некоторые из них описаны в фантастике (роман «Нейромант», 1984, фильмы – «Она», 2013, «Бегущий по лезвию», 1982, «Бегущий по лезвию 2049», 2017 и др.), а некоторые – уже существуют в реальности (например, скандальная попытка фракции ЛДПР принять в партию и передать мандат депутата Госдумы нейросети «Жириновский»<sup>1</sup>).

### **Нейросети нового поколения в контексте философии**

*Философия искусственного интеллекта.* Область, в которой ИИ находит свое философское осмысление, обычно называют «философией искусственного интеллекта». В основном она считается отраслью философии и сферой работы профессиональных философов, но заметное место в ней также занимают концепции и суждения специалистов компьютерных наук, разработчиков ИИ, математиков, логиков, а также историков и социологов науки, специалистов из областей психологии, лингвистики, нейробиологии и даже некоторые идеи и образы писателей-фантастов и популяризаторов науки. В философском плане философия ИИ считается наиболее близкой к областям эпистемологии и философии разума, в рамках которых предполагается осмысливать и философские аспекты ИИ, которые концентрируются вокруг вопросов о характере и природе его «разумности» и «субъектности», о его сравнении с человеческим разумом, о возможностях достичь уровня интеллекта

---

<sup>1</sup> Нейросеть «Жириновский» может стать членом ЛДПР // Газета.ру. – 13.07.2023. – Режим доступа: <https://www.gazeta.ru/tech/news/2023/07/13/20862908.shtml> (дата посещения: 14.07.2023).

человека или даже превзойти его и т.д. [Ringle, 1979; Bolter, 1984; Boden, 1990].

В частности, в философии ИИ были предложены аргументы против «разумности» ИИ (аргумент «китайская комната» Сёрля, «аргумент Гёделя» и др.), ставящие под сомнение возможность достижения «сильного» (или универсального) ИИ; обсуждались проблемы «компьютерной метафоры», допустимость уподобления работы мозга работе компьютера и «обратной метафоры» уподобления работы компьютера работе мозга; «тестов Тьюринга», дискутировалась проблематика «вычислительного интеллекта» в природе, человеке, обществе, технологиях и др. [Burkholder, 1992; Nilsson, 1998].

Одним из аспектов философского осмысления ИИ является также проблема экзистенциальных рисков, возникающих для человечества ввиду появления и развития «сильного» ИИ. В этой связи обсуждаются вопросы регулирования и контроля ИИ, возможности ограничения этих технологий, а также этические проблемы взаимоотношений людей с технологиями ИИ. К этому примыкает и проблематика стратегий сближения и «воспитания» дружественного ИИ, предотвращения его враждебности или отчуждения.

Несомненно, что появление нейросетей нового поколения приведет, и уже привело, к активизации дискуссий и обсуждений в сфере проблематики философии ИИ и к появлению в ней новых тем и акцентов. Во-первых, этот технологический скачок означает заметное продвижение на пути к созданию «сильного» ИИ, что ставит под вопрос некоторые представленные в философии ИИ аргументы сомнения, требуя их пересмотра или переосмысления. Во-вторых, он продемонстрировал неожиданные, взрывные темпы и масштабы процесса, указывающие на его скрытый и не вполне осознаваемый потенциал, что требует от философии ИИ разработки концепций более углубленного и всестороннего осмысления этого потенциала, новых интерпретаций его явных и выявления скрытых компонентов. В-третьих, непосредственно возникают и новые направления анализа.

К числу последних, например, можно отнести проблематику «контроля ИИ над языком и другими кодами человеческой цивилизации», которая явно навеяна возможностями новых нейросетей генерировать большие осмысленные тексты и программные коды, основанные на больших языковых моделях. Ее, в частности, озву-

чивает известный историк и философ Ной Харари, также подписавший призыв о приостановке тренинга новых моделей ГPT. В центре его анализа находится способность овладения языком / многими языками и культурными кодами людей, которую уже демонстрируют новые нейросети. И в этом может заключаться опасность<sup>1</sup>.

*Философия информации.* Импульс возможного изменения и обновления проблемного поля философии ИИ может повлечь за собой и усиление ее контактов и обменов и с другими, особенно с близкими и смежными философскими областями. И в этой связи мы хотели бы привлечь внимание к области *философии информации*. Она также, как и философия ИИ, относится в основном к философии, причем к значительно более широкому полю философских направлений (от онтологии и эпистемологии до этики, эстетики и социальной философии), и также инкорпорирует знания компьютерных и информационных наук. Но по отношению к последним она имеет особый статус, выполняя метатеоретические и методологические функции.

В целом, по сравнению с философией ИИ, философия информации является значительно более широкой и всеобъемлющей областью. Автор самого термина «философия информации» и наиболее известный исследователь этого направления Лючано Флориди считает, что философия ИИ исторически служила предшественницей философии информации [Floridi, 2002], пытаясь объединять эпистемологические вопросы информационно-вычислительных технологий с вопросами этики и социального развития. Но она не смогла выйти на достаточно широкий уровень философских обобщений, позволяющих более глубоко и всесторонне осмысливать эти вопросы. Выйти на такой уровень позволяет философия информации, которую Флориди позиционирует как первофилософию (*греч.* πρῶτῆ φιλοσοφία, *лат.* philosophia prima) современного общества, отсылая, в том числе, и к тому, что это общество само все настойчивее идентифицирует себя как информационное [Floridi, 2002].

Одно из ключевых понятий философии информации – понятие *инфосферы* (информационной сферы, среды) как окружающей среды жизни современного человечества. Оно было выработано по

---

<sup>1</sup> Harari Y.N. Will the future be human? – 25.01.2018. – Mode of access: <https://www.youtube.com/watch?v=hL9uk4hKyq4> (accessed: 20.07.2023).

аналогии с понятиями биосферы, техносферы, ноосферы и т.д., подчеркивающими средовой, пространственный и одновременно глобальный характер соответствующих феноменов. Оно также отсылает к вопросам о ее происхождении, составе, процессуальных изменениях и (по аналогии с биосферой) о ее обитателях и отношениях с другими средами. Флориди раскрывает эти вопросы в своей работе «Четвертая революция: как инфосфера меняет человеческую реальность» [Floridi, 2014].

В самом общем виде инфосфера состоит из информации, информационных процессов. При этом важной ее чертой является семантическая природа, раскрываемая Флориди в его семантической теории информации, обосновывающей наличие в информации элементов истинности, правильности, эквивалентности, располагающихся на шкалах от полной истинности до полной ложности. В этом смысле его теория информации отличается от ее количественных версий (теория информации Шеннона и др.), где информация измеряется через обратное соотношение с энтропией (неопределенностью) [Floridi, 2002].

Инфосфера сопровождает человечество с древности, появление языков, письменности, книгопечатания, электронных СМИ были вехами в ее развитии. Но особый, грандиозный размах она приобретает в условиях современной информационной и компьютерной революции. Инфосфера многократно увеличивается, ее влияние и проникновение в жизнь людей становится более глубоким, устойчивым и всесторонним, в том числе и благодаря развитию информационных технологий. Адекватное понимание места людей в инфосфере Флориди называет «четвертым коперниканским поворотом» в истории науки, связывая его с именем Алана Тьюрига. И этот поворот, как и первые три, связанные, соответственно, с именами Коперника, Дарвина и Фрейда в своих областях, также порывает в познании инфосферы с принципом антропоцентризма [Floridi, 2014].

В инфосфере, по Флориди, можно выделить три вида сущностей: документы (данные), информационные агенты (сущности, способные оперировать с данными) и операции. «Документы» – это любой тип данных, информации и знаний, кодифицированных и реализованных в любом семиотическом формате»; «агенты» – это любые системы, способные взаимодействовать и оперировать с документами; «операции» – это любой тип действия, взаимодей-

ствия и преобразования, которые могут быть выполнены агентом и которые могут быть представлены в документе [Floridi, 2014].

Флориди обращает внимание на два процесса, происходящих сегодня в инфосфере и обуславливающие ее изменения. Во-первых, ее цифровизация, экспоненциальный рост ее цифрового сегмента, что обеспечивает ее однородность, «беспрепятственность» / «бесшовность» (*frictionless / seamless*), многократно повышая ее пропускную способность, мощность и операбельность, резко увеличивая масштабы и скорости происходящих в ней процессов. Во-вторых, быстрое совершенствование и рост инфраструктуры интерфейсов – устройств быстрого и удобного взаимодействия с цифровым сегментом инфосферы людей, техники и других объектов нецифрового мира, что ведет к росту прозрачности границ между цифровым и нецифровым мирами и появлению симбиотических форм и модусов существования и деятельности [Floridi, 2014].

Именно эти тенденции (прозрачность границ, симбиотичность и рост мощности цифровой инфосферы) должны учитываться в философском анализе инфосферы и в выработке этики отношений и действий в мире с новыми технологиями ИИ. При этом, разумеется, нельзя игнорировать и проблемы, возникающие в этой связи – злоупотребления и преступления в информационно-цифровой среде, проблемы цифрового неравенства, проблемы утраты людьми ряда традиционных навыков и др. [Floridi, 2021].

Флориди также считал важным для философии информации определить онтологический статус инфосферы и ее цифровых сегментов. Как и в каком смысле они существуют? И как их онтологический статус соотносится с традициями имеющих философовских онтологий? Свой подход в этом вопросе философ определяет как многокомпонентный. С одной стороны, он говорит о его близости к платоновскому реализму (в смысле реализма мира идей), с другой – указывает на связь с реляционной онтологической традицией Аристотеля (в смысле отношений материя – форма). Он также признает важность для понимания бытия инфосферы учения о «третьем мире» Поппера, а для ее семантического измерения – семантической теории истины Тарского. Наконец, важными для ее онтологического понимания могут считаться и основные парадигмы, выработанные в современной философии языка [Floridi, 2002].

Возвращаясь к вопросу о философском осмыслении нового этапа развития нейросетевых технологий ИИ, хотелось бы отметить преимущества контекста философии информации по сравнению со сложившимся контекстом философии ИИ. Во-первых, он заметно шире и содержит более широкую и разностороннюю философскую проработку, претендующую на статус упомянутой выше *philosophia prima*. Во-вторых, он содержит метарефлексию над значительно более широким полем компьютерных и информационных наук, что позволяет шире и перспективнее видеть явный и скрытый потенциал разработок и практик ИИ в общем контексте разработок и практик ИКТ. В-третьих, сама широта подхода нивелирует тяготение к крайним позициям (безудержного оптимизма либо алармизма в отношении ИИ) и способствует выработке более всестороннего, сбалансированного и в то же время продуктивного для научного анализа взгляда на соответствующие проблемы.

В отношении нейросетей нового поколения, основанных на больших языковых моделях, Флориди высказался достаточно осторожно. Он признал их выдающиеся способности в обработке больших текстовых массивов и получении очень похожих на человеческий интеллект выводов. В то же время он пока не считает эти способности интеллект в полном смысле. Они пока не способны на сложные рассуждения и выработку сложных решений. Кроме того, они еще недостаточно надежны технически. Оптимальным он считает объединение усилий искусственного и человеческого интеллекта [Floridi, 2023].

### **Новое поколение нейронных сетей в контексте социальных наук**

*Компьютеризация науки и гибридные области социальных наук.* То, что появление нейросетей нового поколения с их новыми мощностями и возможностями будет оказывать все возрастающее влияние на различные области научных исследований и получение новых знаний, вполне очевидно. Наука как способ познания, наиболее насыщенный интеллектом, особенно заинтересована в использовании ИИ и чувствительна к его новациям. На приоритетное продвижение этих новых технологий в научные исследования не раз указывали и упоминавшиеся выше эксперты. Во многих

областях современной науки давно существуют целые направления и дисциплины, в которых компьютерные компоненты и возможности ИИ нашли широчайшее применение и стали их неотъемлемой частью.

Один из ближайших примеров, который приходилось рассматривать специально – становление вычислительной биологии, тесно интегрировавшей знания биологии и компьютерных наук [Авдонин, Спилов, Еремеев, 2020]. В этой дисциплине вычисления применяются не просто как вспомогательные средства биологических исследований, но и проникают в теоретический аппарат биологической науки, в теоретическое осмысление биологических и биохимических процессов. Наше исследование показало формирование гибридного языка, гибридных понятий этой науки и активный взаимный трансфер знаний. Был рассмотрен, в частности, пример области «иммунных вычислений» в компьютерных науках (применяется в разработке средств борьбы с компьютерными вирусами), влияние на становление которой оказал трансфер знаний из биологической иммунологии, а последняя активно использует иммунные вычисления (иммунные алгоритмы) для исследования средств повышения биологического иммунитета [Спилов, 2022]. Рассматривалась также концепция междисциплинарных трансферов в когнитивной науке, которая является еще более сложной и многосторонней [Авдонин, 2018; 2019].

Примерно то же можно сказать и о перспективе появления новых технологий в социально-политических науках, хотя следует учесть более сложный характер их дисциплинарных парадигм. Все они так или иначе ориентируются на научный метод и включают элементы сциентистской объяснительной парадигмы, но в то же время в них велика роль и так называемых интуитивно-понимающих подходов и нормативных теорий, также входящих в их дисциплинарные парадигмы. В отличие от первой половины и середины прошлого века, когда в социальных науках был заметен тренд на конфликт этих подходов, для современного обществознания характерно их сближение и постепенная интеграция. В сциентистских подходах обнаруживаются все новые интуитивные основания, а в нормативные теории включаются эмпирические объяснения.

Приход компьютерных технологий в социальные науки способствует укреплению и расширению в них сциентистских методов и теорий, но он сталкивается также и с проблематикой

понимания и нормативности, присутствующими в их парадигмах. С этим связано образование на границах пересечения компьютерных и социальных наук гибридных (междисциплинарных) областей *компьютерных (вычислительных) социальных наук* (вычислительной социологии, вычислительной политологии, вычислительной антропологии и др.) [Gilbert, Troitzsch, 2005; Epstein, Axtell, 1996]. Они не только насыщены компьютерными методами анализа и обработки данных, но и характеризуются симбиозом идей, концепций, категориального аппарата. Их особенность – в относительной проницаемости границ и интенсивных обменах (трансферах) знаний как между компьютерными науками, так и между традиционными социальными науками (социологией, политологией и др.).

Тем не менее некоторая граница между высокоинтегрированной с вычислительными / компьютерными науками областью вычислительных социальных наук и традиционной социальной наукой ощутима, так как последняя вовлечена в эту интеграцию заметно меньше, сохраняя определенную теоретическую дистанцию и отводя вычислительным методам лишь вспомогательное место. На этой границе возникает то, что можно назвать различием во взгляде традиционных и вычислительных социальных наук на информационно-вычислительные процессы и технологии, включая ИИ, в современных обществах. Если первые сохраняют в целом внешний, отстраненно-объектный взгляд на эти явления и на проблемы их влияния на человека и общество, то вторые видят в них становление особого типа агентности и более глубокий характер их взаимодействия с людьми. В литературе, например, представлена впечатляющая история вычислительной социологии на протяжении нескольких последних десятилетий, включая развитие ее социальных теорий (от «динамических систем» в прошлом до «агентных сетей» сегодня), происходившее во многом под влиянием трансфера идей из области вычислительных наук [Macy, Willer, 2002].

Можно указать и примеры обратного влияния и трансфера знаний из социально-политических наук в компьютерные. Хорошо известно, что у истоков разработок в области ИИ были не только программисты и вычислительные математики, но и такие специалисты, как психолог и бизнес-аналитик Аллен Ньюэлл и политолог Герберт Саймон. Последний в 1970-е годы стал лауреатом Нобе-

левской премии за разработку в социальных науках (экономика) теории принятия ограниченно рациональных решений и одновременно премии Тьюринга за ее адаптацию для компьютерных наук [Forest, Mehier, 2001].

*Два сценария для политической науки.* Какое же значение для социально-политических наук будет иметь впечатляющий приход нового поколения нейросетей и продвижение к «сильному» ИИ? В общем плане можно сказать, что это, как и для других наук, даст импульс к дальнейшему сближению «объясняющих» и «понимающих» парадигм, которое в силу большей полипарадигмальности социальных наук будет для них, вероятно, более сложным. Но оно, несомненно, будет ускоряться, так как работа с большими данными и построение сложных мультимодальных моделей, ставшие визитными карточками возможностей нового сетевого ИИ, будут все шире охватывать область социальных наук.

Если же говорить более конкретно, взяв в качестве примера политическую науку, то в ближайшей перспективе можно ожидать следующего. Сценарии событий при этом, видимо, будут различаться в тех двух ее сегментах, о которых было сказано выше. В сегменте традиционной политологии можно предположить *адаптационный сценарий*, в рамках которого новые возможности нейросетей будут восприняты как новые средства и инструменты исследований, открывающие новые перспективы, а также как новые объекты изучения, влияющие на сферу политики в обществе. При этом сама эта сфера будет продолжать пониматься сквозь призму теорий и измерений, релевантных самой политической науке (можно также отметить, что сама по себе эта сеть достаточно сложна и разнообразна и открывает разные возможности и направления познания политики, в том числе тех ее явлений и процессов, которые будут связаны с новыми возможностями ИИ). То, что может происходить в рамках «адаптационного» сценария, показывают различные публикации по темам «цифровизация и политика», «ИИ и политика» и т.д.

В качестве иллюстрации можно обратиться к книге «Политика в цифровом обществе», изданной группой авторитетных политологов в 2019 г., еще до широкого объявления о приходе новых нейросетевых технологий [Hofmann et al., 2019]. В ней вполне определенно заявлены приоритетные для политологии направления исследований, связанных с цифровой трансформацией: изменения

в медиасреде современных обществ, изменения в сфере политической конкуренции и в политическом публичном пространстве, изменения в международной политике и др. Специально выделяются темы исследования политического участия в новых условиях, включая исследование изменений участия в политических партиях, общественных движениях, исследование социальных сетей. Также специально выделяются темы управления, электронного правительства и использования в управлении «больших данных». Не касаясь в данном случае содержания этой вполне достойной и интересной работы, можно отметить, что изучение материала в ней проводится на основе базовых теорий и концепций современной политической науки (традиционной политологии), что делает ее примером «адаптационного» сценария.

Примечательно, что и в публикациях, появившихся уже после появления в 2022–2023 гг. новых сетевых технологий ИИ, этот «сценарный» вариант продолжает сохраняться. Например, в последней работе Андреаса Юнгера «Искусственный интеллект и демократия: концептуальная рамка» исследование, уже учитывая появление новых нейросетей, ведется в концептуальной рамке традиционной политической науки, в частности принятых в ней теорий демократии. Темы анализа в этой работе вполне традиционны: влияние новых нейросетей на политическое участие, на политическую конкуренцию, на политические коммуникации и медиасреду современных демократий, на сферу управления. Он также специально обращает внимание на темы влияния этих технологий на международную политику: а) в плане усиления мировой конкуренции государств и крупных компаний, производящих и развивающих эти технологии и б) в плане усиления в мировой политике автократий, использующих новые технологии для контроля и укрепления авторитарной власти [Jungherr, 2023].

Второй сценарий, который мы бы назвали *ревизионистским*, предполагает существенную «ревизию», перестройку и обновление всего теоретико-методологического арсенала политической науки, постановку принципиально новых исследовательских вопросов, новых познавательных задач и разработку новых способов их решения. Он может реализовываться, прежде всего, в том случае, если новые технологии ИИ существенно затронут уже существующий гибридный сегмент вычислительной политической науки. Его наполнение новыми методами и идеями из смежных компью-

терных областей может привести к эффекту их «переплескивания» в поле традиционной политической науки, вызвав там радикальные изменения. Одновременно и в самом этом поле должны быть обнаружены «реперные точки» для освоения в нем новых идей и подходов.

Эти рассуждения о втором сценарии могли бы быть чисто гипотетическими, если бы не появление ряда публикаций, которые можно интерпретировать в этом ключе. Примером здесь может быть книга гарвардского профессора Матиаса Риссе «Политическая теория цифровой эпохи: куда нас может привести искусственный интеллект»<sup>1</sup>, вышедшая в 2023 г. [Risse, 2023].

Не останавливаясь подробно на ее содержании и отвлекаясь от ее неоднозначных оценок, в частности от упреков в излишне представленных в ней суждениях спекулятивной футурологии, нам важно то, что автор указывает в ней возможные «реперные точки» грядущего поворота в теориях политической науки, вызываемого новыми технологиями ИИ. Одной из таких точек может быть широко принятая и признанная теория справедливости Джона Ролза. В свое время она сыграла существенную роль в становлении теоретической парадигмы современной политической науки, введя в договорную теорию политики ряд важных теоретических новаций («исходное состояние», «рефлексивное выравнивание», «завеса незнания» и др.) [Ролз, 2010]. В известной мере она позволила нащупать важные узлы равновесия между свободой, равенством и порядком и пути продвижения к ним. Расширение и применение теории Ролза к эпохе ИИ, по мысли Риссе, позволило бы продуктивно искать аналогичные узлы для цифровой эпохи.

Другой реперной точкой поворота в теориях политики, о которой говорит Риссе, может быть переосмысление техники и технологий в духе «технологического детерминизма» Маркса, представленного в его концепции о влиянии «производительных сил» на «производственные отношения» [Маркс, 1959]. Разумеется, она должна быть переосмыслена с учетом реалий цифрового мира, но сам принцип неизбежного влияния технологий на общество должен так или иначе применяться при разработке политических теорий.

Наконец, еще одним, третьим элементом теоретических преобразований в политической науке в условиях роста влияния

---

<sup>1</sup> См. рецензию А. Долгова в этом номере.

новых технологий ИИ, который не упоминается у Риссе, и его мы добавляем от себя, может быть философия информации и ее разработка Лючано Флориди, о которой говорилось выше.

Таким образом, и соображения Риссе, и другие публикации, и наши собственные наблюдения за полем политических теорий позволяют говорить о некоторой возможности ревизионистского сценария, предполагающего, что рывок нейросетевых технологий ИИ может вызвать глобальную перестройку политической науки. Хотя пока более вероятен, разумеется, адаптационный сценарий.

### **Заключение**

Итак, анализ появления нейросетей нового поколения в трех предложенных контекстах (технологическом, философском и социально-научном) позволяет сформулировать следующие соображения.

Прежде всего, становится очевидно, что этот анализ требует широты подхода, сопоставления и синтеза многих знаний из разных областей для формирования некоторого общего видения проблем, связанных с появлением технологического феномена качественно нового уровня. В то же время этот широкий подход, чтобы не потерять ориентиры, предполагает и учет специфики самих контекстов, понимания способов формирования и решения проблем в соответствующих областях.

В самом общем виде можно заключить, что мы имеем дело с технологическим феноменом, способным изменять разные контексты. И общим моментом его появления является дисбаланс, нарушение сложившегося равновесия в этих контекстах. Наш анализ показал, что в каждом из них этот дисбаланс, в особой форме, но происходит. Инновационная технология такого уровня способна вносить изменения в разные контексты – и технологические, и социально-экономические, и политические, и философские, и научные. Нарушение равновесия в них происходит по-разному, но общей чертой становится рост озабоченностей, тревог, ощущение нестабильности, активизация поисков возвращения к равновесию.

Еще одной общей чертой, как показывает анализ, является формирование проактивных и комплиментарных для новой технологии зон, выделяющихся на общем фоне более традиционного и инертного к ней или охваченного озабоченностями пространства.

Важной проблемой в этих условиях становится проблема трансфера или «диффузии инноваций» между проактивными зонами и относительно инертными или «встревоженными» пространствами. Инновационные домены должны быть более открытыми и привлекательными для более традиционных.

Эти соображения повлияли на наш подход к анализу технологии новых нейросетей в философском и социально-научном контекстах. В философии это побудило нас отдать предпочтение подходу философии информации по сравнению с более традиционным контекстом философии искусственного интеллекта. Прежде всего, из-за большей широты и фундаментальности философии информации, особенно в том теоретико-методологическом формате, который придал ей Лючано Флориди. На этом фоне контекст философии ИИ, сосредоточенной на темах разума, сознания, возможностях создания их искусственных аналогов, возникающих в связи с этим коллизиях и проблемах, представляется недостаточно широкой и фундаментальной рамкой. К тому же широта подхода способна сдерживать тягу к крайним позициям, что может способствовать укреплению в поле разгорающихся вокруг новой технологии и принимающих глобальный политический масштаб дебатов тенденции к более адекватному и согласованному пониманию проблем.

Что касается контекста социально-политических наук и политической науки как его части, то здесь мы предпочитали исходить из опыта его науковедческого, историко-научного и философско-методологического анализа. В современной науке он выявляет тенденцию к междисциплинарности, к формированию все новых гибридных междисциплинарных областей, возникающих на границах различных наук. И появление новой технологии будет, несомненно, вносить вклад в этот процесс. Тем более что наука считается одной из самых приоритетных областей ее применения. И здесь мы говорим о двух возможных сценариях влияния этой технологии нового уровня на политическую науку – «адаптационном» и «ревизионистском». На первый взгляд, наиболее вероятен первый сценарий, когда политическая наука просто адаптирует новую технологию к своему традиционному арсеналу средств, но, помня о «необычности» этой технологии и ее способности менять разные контексты, нельзя исключать и более радикального сценария, ведущего к глубокой перестройке этой науки.

**V.S. Avdonin, V.L. Silaeva\***  
**New generation neural networks in the context of artificial intelligence technologies, philosophy and socio-political sciences**

*Abstract.* In the article, the processes of the impact of technological changes on political science are investigated on the example of new generation neural network technology (GPT chat, etc.), which caused a great resonance in the media at the turn of 2022–2023. The authors preface the analysis of the influence with a brief historical and scientific review of the emergence and development of neural network technologies and their features and capabilities that have been embodied in new generation systems. The very appearance of these systems and demonstration of their capabilities in various fields and spheres of activity is considered not only as an important stage in the development of computer technology, paving the way to the creation of «strong» artificial intelligence (AI), but also as an event in global technological development, capable of influencing various areas of life and activities of mankind, which is shown by the unprecedented reaction to it in world politics (up to the world summit and the UN Security Council). The article also highlights the philosophical context in which it is preferable to think about these new technologies and their capabilities. This may be the area of philosophy of information and its interpretation in the works of the philosopher Luciano Floridi, which has a number of advantages over other approaches. Analyzing the impact of these new technologies on the socio-political sciences, the authors proceed from the concept of hybrid interdisciplinary scientific fields that arise at the intersection of sciences and have features that distinguish them from the traditional fields of these sciences. The article discusses hybrid fields of computational social sciences developing at the intersection of computer and social sciences. Using the example of political science, the authors propose two scenarios («adaptive» and «revisionist») of the influence of emerging technologies of new generation neural networks on it, each of which is already manifested in a number of the latest publications of political scientists.

*Keywords:* neural networks; artificial intelligence; large language models; deep learning; transformer architecture; information philosophy; Luciano Floridi; infosphere; information theories; information agents; social sciences; hybrid fields; computational social sciences; traditional political science; political theory reform.

*For citation:* Avdonin V.S., Silaeva V.L. New generation neural networks in the context of artificial intelligence technologies, philosophy and socio-political sciences. *Political science (RU)*. 2023, N 4, P. 127–154. DOI: <http://www.doi.org/10.31249/poln/2023.04.05>

## References

Avdonin V.S. On the conditions and means of knowledge transfer in interdisciplinary research. *Sociological journal*. 2019, Vol. 25, N 3, P. 99–116. (In Russ.)

---

\* **Avdonin Vladimir**, INION (Moscow, Russia), e-mail: [avdoninvla@mail.ru](mailto:avdoninvla@mail.ru);  
**Silaeva Victoria**, INION (Moscow, Russia), e-mail: [seva@starlink.ru](mailto:seva@starlink.ru)

- Avdonin V.S. Syntheses in evolutionary biology and scenarios of interaction with social sciences. In: *Method: Moscow quarterly of works from social science disciplines*. Moscow: INION RAS, 2018, N 8, P. 111–129. (In Russ.)
- Avdonin V.S., Spirov A.V., Ereemeev A.V. Interdisciplinary knowledge transfer as metaphorical transfers: evolutionary biology, evolutionary computing and computational evolutionary biology as areas of interdisciplinary transfers. *Sociology of science and technology*. 2020, Vol. 11, N 4, P. 111–139. (In Russ.)
- Boden A. (ed.). *The philosophy of artificial intelligence* (Oxford readings in philosophy). Oxford: Oxford university press, USA, 1990, 452 p.
- Bolter J.D. *Turing's man: western culture in the computer age*. Chapel Hill: University of North Carolina press, 1984, 280 p.
- Bringsjord S., Govindarajulu S. Artificial Intelligence. In: Zalta E.N., Nodelman U. (eds.) *Stanford Encyclopedia of Philosophy* (2022 Edition) / Mode of access: <https://plato.stanford.edu/entries/artificial-intelligence/> (accessed: 20.07.23)
- Burkholder L. (ed.). *Philosophy and the computer*. Boulder, San Francisco, Oxford: Westview press, 1992, 268 p.
- Charniak E., McDermott D. *Introduction to artificial intelligence*. Reading, MA: Addison Wesley, 1985, 701 p.
- Corea F. AI knowledge map: how to classify AI technologies. In: Corea F. (ed.). *An introduction to data: everything you need to know about AI, big data and data science*. Cham: Springer, 2019, P. 25–29. DOI: [https://doi.org/10.1007/978-3-030-04468-8\\_4](https://doi.org/10.1007/978-3-030-04468-8_4)
- Epstein J.M., Axtell R. *Growing artificial societies: social science from the bottom up*. Washington DC: Brookings Institution Press, 1996, 224 p.
- Floridi L. AI as agency without intelligence: on ChatGPT, large language models, and other generative models. *Philosophy and technology*. 2023, Vol. 36, N 1. DOI: <https://doi.org/10.1007/s13347-023-00621-y>
- Floridi L. *Ethics, governance, and policies in artificial intelligence*. Cham: Springer, 2021, 394 p. DOI: <https://doi.org/10.1007/978-3-030-81907-1>
- Floridi L. *The fourth revolution: how the infosphere is reshaping human reality*. Oxford: Oxford university press, 2014, 245 p.
- Floridi L. What is the philosophy of information? *Metaphilosophy*. 2002, Vol. 33, N (1/2), P. 123–145.
- Forest J., Mehier C. John R. Commons and Herbert A. Simon on the concept of rationality. *Journal of economic issues*. 2001, Vol. 35, N 3, P. 591–605. DOI: <https://doi.org/10.1080/00213624.2001.11506392>
- Gilbert N., Troitzsch K. *Simulation for social scientists*. New York: McGraw-Hill, 2005, 295 p.
- Hofmann J., Kersting N., Ritzi C., Schünemann W.J. (Hg.) *Politik in der digitalen Gesellschaft. Zentrale Problemfelder und Forschungsperspektiven*. Bielefeld transcript Verlag, 2019, 329 s. DOI: <https://doi.org/10.14361/9783839448649>
- Jungherr A. Artificial intelligence and democracy: a conceptual framework. *Social Media + Society*. 2023, Vol. 9, N 3. DOI: <https://doi.org/10.1177/20563051231186353>

- Macy M.W., Willer R. From factors to actors: computational sociology and agent-based modeling. *Annual review of sociology*. 2002, Vol. 28, N 1, P. 143–166. DOI: <https://doi.org/10.1146/annurev.soc.28.110601.141117>
- Marx K. Toward a critique of political economy preface. In: Marx K., Engels F. (eds). *Works*. 2nd ed., Vol. 13, Moscow: State publishing house of political literature, 1959, P. 5–9. (In Russ.)
- Minsky M., Pappert S. *Perceptrons: an introduction to computational geometry*. Cambridge, MA: MIT Press, 1969, 258 p.
- Nilsson N. *Artificial intelligence: a new synthesis*. San Francisco, CA: Morgan Kaufmann, 1998, 513 p.
- Ringle M. (ed.). *Philosophical perspectives in artificial intelligence*. Atlantic Highlands, NJ: Humanities press, 1979, 244 p.
- Risse M. *Political theory of the digital age: where artificial intelligence might take us*. Cambridge: Cambridge university press, 2023, xxvi, 304 p.
- Rolls J. *Theory of justice*. Moscow: Publishing house LKI, 2010, 536 p. (In Russ.)
- Russell S.J., Norvig P. *Artificial intelligence: a modern approach* (4th ed.). Hoboken: Pearson, 2021, 1136 p.
- Spirov A. Immune computing in computer science and models of immune memory of higher organisms: prospects for mutual methodological enrichment. *Method: Moscow quarterly of works from social science disciplines*. Moscow: INION RAS, 2022, Vol. 2, N 3. (In Russ.)

### Литература на русском языке

- Авдонин В.С. Об условиях и средствах трансфера знаний в междисциплинарных исследованиях // Социологический журнал. – 2019. – Т. 25, № 3. – С. 99–116.
- Авдонин В.С. Синтезы в эволюционной биологии и сценарии взаимодействия с социальными науками // МЕТОД: Московский ежегодник трудов из обществоведческих дисциплин: сб. науч. тр. / РАН. ИНИОН. – М.: ИНИОН РАН, 2018. – Вып. 8. – С. 111–129.
- Авдонин В.С., Спиров А.В., Еремеев А.В. Междисциплинарный трансфер знаний как метафорические переносы: эволюционная биология, эволюционные вычисления и вычислительная эволюционная биология как области междисциплинарных трансферов // Социология науки и технологий. – 2020. – Т. 11, № 4. – С. 111–139.
- Маркс К. К критике политической экономии. Предисловие // Маркс К., Энгельс Ф. Сочинения. – 2-е изд. – Т. 13. – М.: Государственное издательство политической литературы, 1959. – С. 5–9.
- Ролз Дж. Теория справедливости / пер. с англ. В.В. Целищева. – М.: Издательство ЛКИ, 2010. – 536 с.
- Спиров А. Иммунные вычисления в компьютерных науках и модели иммунной памяти высших организмов: перспективы взаимного методологического обогащения // МЕТОД: Московский ежеквартальник трудов из обществоведческих дисциплин: ежекв. науч. изд. – М.: ИНИОН РАН, 2022. – Т. 2, № 3. (В печати).