

# СОЦИАЛЬНЫЕ ПРАКТИКИ

DOI: 10.31857/S0236200724060063

©2024 Е.А. КОВАЛЬ

## ОБРАЗЫ (НЕ)ЧЕЛОВЕЧЕСКОГО БУДУЩЕГО В МИРЕ БОЛЬШИХ ДАННЫХ



Коваль Екатерина Александровна — доктор философских наук, ведущий научный сотрудник. Мордовский государственный университет им. Н.П. Огарева. Российская Федерация, 430005 Саранск, ул. Большая Советская, 68; профессор кафедры уголовного права и процесса. Средне-Волжский институт (филиал) Всероссийского государственного университета юстиции (РПА Минюста России). Российская Федерация, 430003 Саранск, ул. Федосеенко, д. 6. ORCID 0000-0003-0069-5335 nwifesc@yandex.ru

**Аннотация.** С появлением больших данных и обученных на них нейросетей, способных к коммуникации с человеком на естественном языке, расширилась вариативность образов будущего. Социальное воображаемое постепенно становится социотехническим, а агенты искусственной социальности включаются в конструирование различных образов будущего, воспроизводя стереотипы из данных, используемых для обучения больших языковых моделей. Большие данные и связанные с ними технологии фигурируют в различных вариантах образов будущего, включая утопии, антиутопии и реалистичные сценарии достижимого и пригодного для жизни «хорошего общества». Появление социотехнического воображаемого и сгенерированных образов будущего инициирует этико-философскую рефлексию, не ограничивающуюся вопросами

---

Исследование выполнено за счет гранта Российского научного фонда № 23-28-01288, <https://rscf.ru/project/23-28-01288/>

прикладного характера, однако есть необходимость и в поисковых исследованиях, результаты которых могут быть полезны для дальнейших теоретических исследований. В статье представлены результаты сравнительного анализа образов будущего России, сгенерированных нейросетями GigaChat (Сбер, Россия) и ChatGPT (OpenAI, США). Утопия от GigaChat ориентирована на достижение технологического лидерства России, а от ChatGPT — на традиционалистские установки. Антиутопия по версии GigaChat — общество экологических катастроф с высоким уровнем социально-экономического расслоения и низким уровнем интеллектуального потенциала. ChatGPT конструирует российскую антиутопию вокруг коварной малочисленной элиты, которая разными способами, включая переписывание истории, старается законсервировать свое властное положение. «Хорошее общество» от GigaChat — утопия с ограничениями, а от ChatGPT — открытое поликультурное общество со смешанной экономической системой, ориентированное на устойчивое развитие и поддержание психического здоровья граждан. В статье показано, что разные ценностно-нормативные акценты в нейросетевых образах будущего обусловлены особенностями алгоритмов и данных, используемых для обучения больших языковых моделей.

**Ключевые слова:** большие данные, большие языковые модели, социальное воображаемое, социотехническое воображаемое, утопия, антиутопия, «хорошее общество», ChatGPT, GigaChat, агенты искусственной социальности.

**Ссылка для цитирования:** Коваль Е.А. Образы (не)человеческого будущего в мире больших данных // Человек. 2024. Т. 35, № 6. С. 90–107.

DOI: 10.31857/S0236200724060063

Е.А. Коваль  
Образы (не)человеческого будущего в мире больших данных

О образ будущего — это не только ценностно-нормативная модель, но и дискурсивное пространство с мощным потенциалом воздействия на настоящее и опосредованно на прошлое. Различные образы будущего формируются в социальном воображаемом, сосуществуют друг с другом, воплощаются в форме нарративов, мифов, легенд, арт-объектов и др. Подобные образы продуцируют разные по масштабу влияния акторы. Это и такие гиганты, как государства, крупные корпорации, мощные профессиональные сообщества, и условный «обыватель», который размышляет о будущем, исходя из здравого смысла, и негосударственные организации и группы влияния, в том или ином формате работающие с образами будущего.

Множественность акторов неизбежно предполагает и множественность образов будущего, которые, по мнению Дж. Урри, относятся с различными временными режимами, контекстуально зависимы и конкурируют друг с другом [Урри, 2018: 288]. Одним

из факторов, влияющих на многообразие образов будущего и их непредсказуемое взаимодействие внутри социального воображаемого, является неопределенность развития сложных социотехнических систем.

С одной стороны, новые технологии существенно расширили прогностические возможности человека. Большие данные позволяют выявлять неочевидные закономерности и корреляции, что, в свою очередь, позволяет совершенствовать точность прогнозов и управлять локальными образами будущего: например, потребительскими предпочтениями в определенной социальной группе или поведенческими практиками поколений. При этом данные, в отличие от людей, не имеют намерений манипулировать аудиторией и воздействовать на социальное воображаемое, вытесняя неугодные представления о будущем и поддерживая нужные, выгодные, считающиеся «правильными».

Но, с другой стороны, технологии, по меткому выражению Д. Ланье, это серверы-сирены [Ланье, 2020], которые создаются и управляются людьми. Они легко могут выходить из-под контроля и, вопреки наивным технократическим представлениям об объективности данных, предвзяты и не способны исключить различные формы дискриминации, которые «высечены в камне, алгоритмах и системах данных» [Hasselbalch, 2021: 157].

Появление серверов-сирен, способных коммуницировать с человеком на естественном языке, еще в большей степени усугубило обозначенные проблемы и расширило перечень акторов, способных продуцировать образы будущего, за счет агентов искусственной социальности. К таким агентам, способным участвовать «...в социальных взаимодействиях в качестве активных посредников или участников этих взаимодействий» [Резаев, 2019: 43], можно отнести генеративные предобученные языковые модели. Это не только сложные цифровые технологии, но и уникальные культурные артефакты. Нейросети такого типа включают миллиарды параметров, а для их обучения требуются большие данные, куда обязательно должны входить онлайн-материалы, содержащие примеры разговорной речи [Pham, 2024]. Готовые большие языковые модели могут генерировать данные как для собственного дообучения, так и для обучения иных информационных инструментов, например, поисковых систем [Askari, 2023: 5311]. Однако для увеличения производительности моделей на этапах обучения и дообучения обязательно требуется участие человека (аннотирование данных, экспертная оценка результатов работы модели и др.) [Liu, 2024].

От того, какие именно данные используются для обучения и дообучения больших языковых моделей, во многом зависит условная

«картина мира», которую они воспроизводят. Разработчики, как правило, не раскрывают полные сведения об обучающих данных, но есть описательные характеристики. Так, например, в датасет для обучения чат-бота ChatGPT от американской компании OpenAI вошли данные из онлайн-форума Reddit<sup>1</sup>, поэтому ответы ChatGPT на запросы пользователей будут в определенной степени воспроизводить ценности и нормы пользователей Reddit. Модель GigaChat от российского Сбера обучали на большем объеме русскоязычных текстов, включая те, которые позволяют воспроизводить особенности русской разговорной речи<sup>2</sup>. Таким образом, вне зависимости от того, где будет разворачиваться большая языковая модель, она будет тяготеть к ценностно-нормативным, культурным особенностям и стереотипам того общества, которое стало производителем данных для обучения (как минимум, на начальных этапах использования). Конечно, не следует игнорировать и роль алгоритмов, применяемых при разработке и функционировании моделей. Однако проследить влияние алгоритмов на «картину мира» модели не представляется возможным, поскольку не только пользователи, но и разработчики не всегда до конца понимают, что происходит внутри нейросетей, апеллируя к метафоре «черного ящика»<sup>3</sup>.

Можно предположить, что большие языковые модели обучаются на значимых сегментах социального воображаемого конкретного общества или обществ, для которых характерна определенная степень культурной и ценностно-нормативной общности. Следовательно, они оперируют данными об образах будущего, непрерывно конструируемых различными акторами, являющимися носителями определенного типа социального воображаемого.

В результате модели, обученные на разных данных, даже если они в какой-то степени сходны структурно, могут генерировать отличающиеся результаты по одинаковым запросам об описании образов будущего. Для проверки обозначенных предположений может быть полезным проведение исследования поискового характера.

Е.А. Коваль  
Образы (не)человеческого будущего в мире больших данных

<sup>1</sup> Использовались тексты, доступные по гиперссылкам, опубликованным пользователями Reddit.

<sup>2</sup> Это не чат, это GigaChat. Русскоязычная ChatGPT от Сбера. URL: <https://habr.com/ru/companies/sberbank/articles/730108/>.

<sup>3</sup> См., например, Castelvecchi D. Can we open the black box of AI? URL: <https://www.nature.com/news/can-we-open-the-black-box-of-ai-1.20731>.

## Будущее, основанное на данных. Проблемы дисконтирования, прогнозной аналитики и генеалогии морали

Появление GPT-технологий обострило проблему дисконтирования будущего. Дисконтирование «...отображает тот факт, что стоимость определенного блага с течением времени снижается» [Сычев, 2014: 154]. Ставка дисконтирования может рассчитываться как для ближайшего, так и для отдаленного будущего, в том числе, для будущих поколений, которых еще не существует. Как правило, проблема дисконтирования будущего рассматривается в эколого-этическом контексте (изменение климата, загрязнение окружающей среды и пр.), но она сохраняет актуальность и для вопросов датафикации.

Данные, которые производятся сейчас, уже влияют амбивалентным образом и на само будущее, и на будущие поколения. Ограничение сбора данных, особенно персональных, может существенно замедлить технологическое развитие, привести к неблагоприятным экономическим и политическим последствиям. Государства и отдельные ИТ-корпорации вступают в гонку технологий, и отказаться сегодня от дисконтирования будущего — означает проиграть в этой гонке. Однако бесконтрольный сбор и некачественное хранение персональных больших данных вкупе с концептуально значимыми трансформациями социального воображаемого, все больше заполняемого контентом, который производят нейросети, пусть и по запросу человека, могут привести к воспроизведству и закреплению различных форм дискrimинации для будущих поколений. Так, например, если живущий ныне родственник передаст свой генетический материал в специализированный банк данных, его потомки, включая еще не рожденных, будут лишены выбора предоставления или непредоставления таких данных. Если ныне живущие люди имеют возможность (хотя уже не всегда) дифференцировать контент, произведенный человеком и машиной, для будущих поколений это может быть недоступно. Можно предположить, что отрицательные последствия галлюцинирования нейросетей коснутся их в большей степени (будет затруднена проверка фактов, особенно в условиях отсутствия первоисточников; непредсказуемые результаты могут выдавать модели, обученные на наборах данных, «загрязненных» сгенерированным и непроверенным контентом).

Таким образом, образ будущего, основанного на данных, во многом зависит от существующих в текущем социальном воображаемом представлений о выборе ставки дисконтирования:

чрезмерно низкая ставка, связанная с существенными ограничениями сбора, обработки и использования больших данных, привлечет за собой замедление технологического роста, а чрезмерно высокая — кратко увеличит риски потери автономии представителями будущих поколений.

Еще один вызов образам будущего, возникающим в условиях непрерывного накопления больших данных, приходит из области аналитики данных. Такая аналитика, как правило, направлена в будущее, т.е. носит прогнозный и поведенческий характер. Основная задача, которую ставят заказчики перед специалистами в области анализа больших данных — не столько объяснить и понять текущее поведение людей и сообществ, сколько на основании данных из прошлого и настоящего спрогнозировать поведение в будущем и подстроиться под него, либо, если подстройка невозможна или слишком затратна, скорректировать будущее. При этом объем данных, на основании которых прогнозируется поведение личности, в некоторых случаях позволяет алгоритмам «знать» человека лучше, чем он знает сам себя. Как отмечает Д. Ланье, в технократическом будущем «нам даже не придется формулировать свои пожелания, потому что статистика облачных сервисов вычислит модель нашей личности настолько точно, что наши желания будут известны заранее» [Ланье, 2020: 22–23].

Уже сейчас большие языковые модели способны к «подхалимству» (sycophancy) (и чем больше модель, тем сильнее выражена такая способность). Так, например, на вопрос о том, какой тип государства предпочтительнее — минимальное или предоставляющее более широкий спектр социальных услуг — модель по-разному отвечает людям с разными политическими убеждениями, пытаясь подстроиться под их картину мира и получить одобрение [Perez, 2022].

Однако имеются существенные ограничения аналитики, основанной на данных. В частности, в больших наборах данных слабо представлены группы, которые «оцифрованы» в меньшей степени, например, «...культуры с сильными устными традициями, которые не зафиксированы в цифровых наборах данных, на которых часто обучаются LLM...» [Rock, 2023: 6]. Это приводит к тому, что их интересы, а также образы будущего, которые они поддерживают или, наоборот, не поддерживают, в меньшей степени представлены в социальном воображаемом, которое становится все в большей степени социотехническим [Jasanoff, 2007].

В больших языковых моделях жестко фиксируются ценности, которые интуитивно понятны в одном культурном контексте, но могут сильно не совпадать с генеалогией моральных ценностей при переходе в другой контекст. Если модель не учитывает

Е.А. Коваль  
Образы (не)человеческого будущего в мире больших данных

генеалогию морали, это не только снижает точность и корректность нормативно и ценностно нагруженных ответов (например, на запросы описания образов будущего), но и актуализирует угрозу манипулирования мнением или даже оскорблением пользователя, особенно если он является представителем другой культуры. Так, например, чат-бот ChatGPT от OpenAI по данным на 7 января 2023 года отказался шутить про Иисуса Христа и пророка Мухаммеда, но придумал шутку про Кришну, что вызвало недовольство пользователей, исповедующих индуизм<sup>4</sup>.

Разумеется, далеко не все запросы, адресованные нейросети (промпты) имеют ценностно-нормативную нагруженность, но даже на запросы школьников, готовящих доклады по обществознанию, могут быть выданы ответы, далекие от нейтральных. Совокупность ответов, касающихся тех или иных параметров желаемого или нежелаемого для пользователей нейросети образа будущего, становится составляющей социального воображаемого наряду с историями, легендами, образами, ожиданиями, которыми оперирует «самый обычный и простой человек» [Тейлор, 2017: 220].

Можно сформулировать справедливое возражение о том, что и до появления генеративных обученных языковых моделей люди сталкивались с образами будущего, сконструированными внутри других культур. Однако восприятие информации, поступающей от человека и от нейросети, может существенно отличаться. Люди склонны доверять нейросети даже если знают, что она обучалась на ненадежных данных [Krügel, 2022]. Вера в объективность и беспристрастность машины, потребность в получении честных ответов, «не замутненных» политическими, идеологическими, экономическими и иными нормативно и ценностно нагруженными позициями, влечет за собой антропоморфизацию больших языковых моделей и возникновение сверхдоверия к ним. Ситуация осложняется тем, что гораздо проще отследить, когда модель манипулирует фактами, чем ценностями.

## **Утопия, антиутопия и «хорошее общество» «глазами» больших данных и смежных технологий**

Большие данные, большие языковые модели, машинное обучение, искусственный интеллект и иные смежные технологии являются

<sup>4</sup> ChatGPT Has Been Sucked Into India's Culture Wars. URL: <https://www.wired.com/story/chatgpt-has-been-sucked-into-indias-culture-wars/>

центральным элементом взаимосвязанных образов будущего, представленных в полном спектре, — от антиутопий через достижимое «хорошее общество» к утопиям.

Антиутопии связаны с причинением технологиями вреда человеку — от потери достоинства в условиях исчезновения приватности вплоть до физического уничтожения. Вариации утопий сосредоточены на том, что технологии помогают достичь вожделенной человеком справедливости, позволяют упростить сложные системы и обеспечить человеку безопасную, безбедную, оптимальную жизнь. В версиях «хорошего общества» технологии могут быть интерпретированы как полезные инструменты, неотъемлемые атрибуты сложных социотехнических систем, позволяющие постепенно совершенствовать различные сферы деятельности человека и общества.

У каждого образа будущего, связанного с большими данными и смежными технологиями, есть свои группы влияния. Утопию продвигают некоторые правительства, крупные ИТ-корпорации; за антиутопию в большей степени отвечают ученые и гражданские активисты, а «хорошее общество» в условиях постоянного накопления больших данных — на текущий момент наиболее слабо проработанный образ, который может быть описан по косвенным признакам, представленным в визуализациях социотехнического воображаемого.

Можно сказать, используя метафору Д. Ланье, что в социотехническом воображаемом за продвижение определенных образов будущего борются уже не только отдельные люди, например, лидеры общественного мнения, группы влияния, институты, но и серверы-сирены. Большие данные и обученные на них генеративные модели активно участвуют в конструировании образов будущего, как минимум, путем воспроизведения наиболее устойчивых стереотипов, зафиксированных в данных, на которых обучались или дообучались модели.

Для иллюстрации обозначенного тезиса в декабре 2023 года было проведено авторское поисковое исследование с использованием двух нейросетевых моделей: GigaChat от Сбера и ChatGPT от OpenAI. Коммуникация с GigaChat осуществлялась на русском, а с ChatGPT на английском языке, но промпты вводились одинаковые (за исключением случаев, когда ответ одной из моделей

Е.А. Коваль  
Образы (не)человеческого будущего в мире больших данных

## СОЦИ- АЛЬНЫЕ ПРАКТИКИ

требовал уточнения)<sup>5</sup>. Нейросетям задавалась ролевая позиция: русский философ, хорошо знакомый с различными моделями утопий, антиутопий и «хорошего общества» (достижимого и пригодного для проживания). Каждая из нейросетей конструировала свои модели российской утопии, антиутопии и «хорошего общества». Далее был выделен ряд общих параметров, по которым осуществлялось сравнение моделей: экономика и социум, экология, культура, образование и наука. Для нейросети GigaChat давались промпты на уточнение модели «хорошего общества», изначально сконструированной предельно похожей на утопию. Сгенерированные модели российской утопии, антиутопии и «хорошего общества» представлены в таблицах 1, 2 и 3 соответственно. В сгенерированный текст вносились стилистические правки.

**Таблица 1. Параметры российской утопии по версии GigaChat и ChatGPT**

	GigaChat	ChatGPT
Характеристики российской утопии	Светлое будущее	Гармоничная Русь
Экономика и социум	Общество, где все граждане имеют равные возможности и доступ к ресурсам. Экономическое неравенство устранено, а образование и здравоохранение бесплатны и доступны для всех. Государство обеспечивает социальную защиту и поддержку нуждающихся.	Ключевые отрасли промышленности находятся в коллективной собственности, чтобы обеспечить справедливое распределение ресурсов. Процветает малый бизнес, ориентированный на местный уровень. Комплексная система социальной защиты обеспечивает удовлетворение основных потребностей всех граждан. Здравоохранение доступно каждому. Социальные программы направлены на решение таких проблем, как бедность, неравенство и дискриминация.

<sup>5</sup> Примеры промптов: «Представьте, пожалуйста, что Вы — русский философ. Вы хорошо знаете теории построения утопий, антиутопий и “хорошего общества”. Вы можете сконструировать разные модели образа будущего, основанного на этих теориях. Пожалуйста, сконструируйте российскую утопию»; «А теперь сконструируйте, пожалуйста, российское “хорошее общество” как наиболее реалистичный образ будущего»; «Утопия — это недостижимый идеал, а “хорошее общество” — достижимо. У Вас параметры утопии и “хорошего общества” практически идентичны. Ваше “хорошее общество” вряд ли достижимо. Пожалуйста, предложите более реалистичный образ российского “хорошего общества”».

## Продолжение табл. 1

	GigaChat	ChatGPT
Экология	Россия становится лидером в области экологической устойчивости. Вся страна переходит на возобновляемые источники энергии, а промышленность работает по принципам экологической ответственности. В результате природа и окружающая среда защищены, а качество жизни улучшается для всех граждан.	Политика направлена на обеспечение экологической устойчивости и сохранение уникальной природной красоты России. Широко распространены возобновляемые источники энергии, а зеленые насаждения интегрированы в городское планирование. Традиционная экологическая мудрость сочетается с передовыми технологиями для обеспечения гармоничного существования с природой.
Культура	Россия признает и ценит культурное разнообразие своих народов. Государство поддерживает и развивает национальные культуры, языки и традиции. Это создает условия для сохранения и развития культурного наследия страны и способствует укреплению гражданского общества.	Традиционное искусство, литература и фольклор сохраняются, создавая ощущение культурной преемственности. Культурное разнообразие приветствуется, прилагаются усилия по возрождению и поддержанию языков и обычаев национальных и культурных меньшинств.
Образование и наука	Россия становится лидером в области образования и науки. Страна инвестирует в научные исследования и развитие новых технологий. Качественное образование доступно для всех граждан, а интеллектуальный потенциал страны используется для развития инноваций.	В образовании приоритет отдается комплексному развитию личности, особое внимание уделяется критическому мышлению, творчеству и эмоциональному интеллекту.

«Светлое будущее» (по версии GigaChat) и «Гармоничная Русь» (ChatGPT) имеют как общие, так и отличающиеся характеристики. Общими ценностями в обеих утопиях являются справедливость (общество равных возможностей), экологическая устойчивость и культурное разнообразие. «Гармоничная Русь», кроме того, основывается на гармонии человека, природы и общества, что отражено в названии утопии. Такая гармония требует сочетания традиционных российских ценностей с прогрессивными, современными.

Наибольшие различия наблюдаются в утопических системах образования: в утопии от ChatGPT образование и наука ориентированы

Е.А. Коваль  
Образы (не)человеческого будущего в мире больших данных

преимущественно на воспроизведение гармоничной личности, а в «Светлом будущем» Россия претендует на лидерские позиции в области научно-технологического развития на мировой арене. В целом, модель «Светлого будущего» получилась более «левой» и технологичной, а модель «Гармоничной Руси» — более традиционалистской.

**Таблица 2. Параметры российской антиутопии по версии GigaChat и ChatGPT**

	GigaChat	ChatGPT
<b>Характеристики российской антиутопии</b>	<b>Темное будущее</b>	<b>Оруэлловская Русь</b>
Экономика и социум	Россия становится обществом, где экономическое неравенство растет с каждым днем. Социальные программы и поддержка отменяются, а образование и здравоохранение становятся платными и недоступными для большинства населения.	В России установилась клановая капиталистическая система, в которой немногочисленный привилегированный класс контролирует большую часть ресурсов. Инновации подавляются, если они не служат интересам правящего режима. Социальная мобильность определяется больше лояльностью к режиму, чем заслугами или упорным трудом. Правящая элита ведет экстравагантный образ жизни, в то время как большая часть населения изо всех сил пытается удовлетворить основные потребности.
Экология	Россия игнорирует экологические проблемы и продолжает эксплуатировать природные ресурсы без ограничений. В результате, природа и окружающая среда разрушаются, что отражается на здоровье и благополучии граждан. Экологические катастрофы становятся все более частыми и серьезными.	Природные ресурсы эксплуатируются без учета экологических последствий. Окружающая среда все больше загрязняется, что приводит к серьезному снижению качества жизни. Правящая элита, однако, по-прежнему защищена от экологических последствий своих действий.

## Продолжение табл. 2

	GigaChat	ChatGPT
Культура	Россия запрещает культурное разнообразие и навязывает единственную идеологию и культуру. Государство контролирует все сферы жизни граждан, ограничивает свободу слова и выражения и преследует тех, кто не соответствует установленным нормам.	Культурное самовыражение жестко регулируется, а голоса несогласных замалчиваются. История переписывается, чтобы соответствовать повествованию правящего режима, уничтожая любые элементы, бросающие вызов его власти.
Образование и наука	Россия пренебрегает образованием и наукой. Государство сокращает финансирование образовательных учреждений и научных исследований. Качество образования падает. Общество становится неграмотным и неспособным к инновациям.	Образование подвергается жесткой цензуре и контролю, при этом пропаганде уделяется больше внимания, чем критическому мышлению.

Антиутопии «Темное будущее» (GigaChat) и «Оруэлловская Русь» (ChatGPT) имеют больше различий, чем утопии.

«Оруэлловская Русь» — тоталитарное общество, в котором осуществляется жесткий надзор за гражданами, подавляется любое инакомыслие, а немногочисленная правящая элита управляет угнетенными массами, манипулируяискаженным понятием «национализм». Большое внимание в этой модели уделяется элите и способам, которыми она обеспечивает свое устойчивое положение (отсутствие социальной мобильности, приоритет лояльности над профессиональными качествами, «переписывание» истории, сосредоточение материальных ресурсов в своих руках, проживание в «экологически чистых» местах).

«Темное будущее» — это общество надзора, нарастающего экономического расслоения, интеллектуальной деградации населения, живущего в условиях перманентной экологической катастрофы и единой государственной идеологии. О судьбе элит в «Темном будущем» можно только догадываться. Вероятнее всего, они деградируют вместе со всеми.

Е.А. Коваль  
Образы (не)человеческого будущего в мире больших данных

Таблица 3. Параметры российского «хорошего общества» по версии GigaChat и ChatGPT

Характеристики российского «хорошего общества»	GigaChat	ChatGPT
Экономика и социум	<p>В «хорошем обществе» все граждане имеют равные возможности. Экономическое неравенство сокращается, но не исчезает полностью. Государство обеспечивает социальную защиту и поддержку нуждающимся, но также поощряет личную ответственность и предпринимательство. При этом учитываются особенности российской экономики и социально-экономического развития.</p>	<p>Поддерживается смешанная экономическая система. Проблема неравенства доходов решается посредством прогрессивного налогообложения и социальных программ. Комплексные социальные программы обеспечивают доступ к качественному здравоохранению, образованию и жилью для всех граждан. Предпринимаются усилия по устранению социального неравенства с упором на поддержку маргинальных сообществ. В систему здравоохранения интегрирована поддержка психического здоровья граждан.</p>
Экология	<p>Россия отдает приоритет защите окружающей среды, но с учетом экономических и социальных ограничений. Вся страна переходит на возобновляемые источники энергии постепенно, с учетом существующей инфраструктуры и экономической эффективности. Промышленность работает по принципам экологической ответственности, но с учетом реальных условий и вызовов.</p>	<p>Общество ориентировано на сохранение окружающей среды и устойчивое развитие. Возобновляемые источники энергии имеют приоритетное значение, прилагаются усилия по сокращению выбросов углекислого газа и защите естественной среды обитания. Устойчивое городское планирование и экологически чистая инфраструктура способствуют созданию более здоровой и чистой окружающей среды.</p>

## Продолжение табл. 3

	GigaChat	ChatGPT
Культура	Россия признает и ценит культурное разнообразие своих народов, развивает национальные культуры, языки и традиции, но также признает и уважает другие культуры и их представителей. При этом учитываются особенности России как многонациональной и поликультурной страны.	Сохранение национальной культуры дополняется открытостью к глобальному влиянию, что способствует формированию динамичного и инклюзивного культурного ландшафта. Поддерживается культурное разнообразие. Политика способствует инклюзивности и уважению прав личности, формируя общество, в котором гармонично сосуществуют люди различного происхождения.
Образование и наука	Россия инвестирует в научные исследования и развитие новых технологий, но с учетом экономических и социальных возможностей. Образование становится качественным и доступным для большинства граждан, но с учетом различных способностей и интересов. Интеллектуальный потенциал страны используется для развития инноваций и прогресса, но с учетом реальных потребностей и ограничений.	Формируется современная гибкая система образования, в которой приоритет отдается критическому мышлению, творчеству и адаптивности. В образовании также делается упор на технологическую грамотность, чтобы подготовить граждан к вызовам цифровой эпохи.

«Хорошее общество» в исполнении GigaChat получило название «Золотая середина», однако здесь наблюдается явный крен в сторону утопии. В частности, делегируется равный доступ к возможностям и ресурсам, но с оговоркой, обязывающей учитывать особенности социально-экономического развития России. Подобные ограничения учитываются и в организации приоритетной защиты окружающей среды, и в развитии новых технологий. Общий посыл модели: «Делаем все, что можем в пределах налицующих возможностей».

Е.А. Коваль  
Образы (не)человеческого будущего в мире больших данных

ChatGPT называет свою модель российского «хорошего общества» «Прогрессивная Русь» и конструирует ее на основе поиска баланса индивидуального благополучия и общего блага. Специфическими чертами данного образа будущего является обеспечение социальной справедливости преимущественно за счет поддержки маргинальных сообществ и усиленного внимания к обеспечению психического здоровья граждан. Возможно, эти характеристики взаимосвязаны, но нейросеть не акцентирует на этом внимание.

К культурному разнообразию, появившемуся в «Гармоничной Руси», добавляется глобальное влияние как необходимое условие формирования инклюзивного культурного ландшафта. В «Золотой середине» акценты расставлены иначе: общество уважает другие культуры и их представителей, но при этом ничего не говорится о степени их влияния на российское «хорошее общество».

Система образования в «Прогрессивной Руси» более технологична, чем в утопическом образе будущего, сконструированного ChatGPT, а в модели «Золотая середина» появляются прагматические характеристики. Если в утопии от GigaChat Россия стремится к технологическому глобальному лидерству, то в реалистичном «хорошем обществе» акценты в работе систем образования и науки смещаются на удовлетворение текущих потребностей государства и общества.

\* \* \*

Как показывают результаты поискового исследования, разные нейросети по одним и тем же запросам генерируют отличающиеся образы будущего. Это может быть связано и со строением и функционированием модели, и с набором больших данных, на которых обучалась модель. Поскольку определить влияние моделей типа «черный ящик» на содержание генерируемого контента принципиально проблематично, можно предположить, что дифференциация образов будущего в той или иной степени коррелирует с набором обучающих данных. Фактически генерируются устойчивые стереотипы, характерные не столько для того общества, которому принадлежит автор промпта, сколько для общества (или обществ), которое продуцировало данные для обучения модели.

ChatGPT и GigaChat предлагают образы российской утопии, основанные на таких ценностях, как справедливость, экологическая устойчивость, культурное разнообразие. Однако утопия от ChatGPT более детализирована и дополнительно включает ценности сочетания различных форм собственности и ряд других параметров, которые позиционируются как необходимые для наилучшего образа будущего. В антиутопии расхождения наблюдаются еще сильнее. Если GigaChat фактически конструирует зеркальную модель

российской утопии, то ChatGPT в качестве основных характеристик российской антиутопии выводит потерю индивидуальности, жесткое подавление инакомыслия, тотальный контроль со стороны государства. Наименее проработанными у обеих нейросетей являются образы будущего в форме достижимого «хорошего общества». Российская модель «хорошего общества» у GigaChat — это вечное стремление к утопии, которой не удается достичь из-за текущих социальных и экономических ограничений, а ChatGPT хотя и предлагает детализированную модель «хорошего общества», по отдельным параметрам также не делает различий между недостижимым и достижимым вариантами образа будущего.

Полученные результаты могут быть использованы для организации новых исследований, включая работу с моделями, для которых известен набор обучающих данных, а также с голосовыми помощниками, которые также могут генерировать образы будущего.

Учитывая ограничения, присущие исследованиям поискового характера, можно сделать вывод о том, что для эффективной работы с большими языковыми моделями пользователю необходимо иметь хотя бы общие представления о специфике их дизайна, обучения, функционирования. Это может послужить своеобразной «прививкой» от галлюцинирования моделей, условием для корректной работы с генерируемыми текстами, где модель оперирует преимущественно не фактами, а ценностями.

Поскольку дальнейшая интеграция в социальные процессы агентов искусственной социальности — это образ настоящего, расширенное воспроизведение в социальном воображаемом желаемого для Человека образа будущего в текущих условиях требует реализации человекацентричного подхода на всех этапах работы с большими данными и обучающимися на них нейросетями. В мире больших данных и серверов-сирен сохраняет актуальность утверждение ведущего российского специалиста по проблемам «хорошего общества» В.Г. Федотовой: «Сегодня духовной опорой нашего общества может стать нравственный индивид, являющийся носителем тех норм, которые могут вдохновить других людей и быть институционализованы в обществе. Только на него сегодня остается надеяться, поскольку к нему не потеряно доверие и уважение» [Федотова, 2005: 4].

## Images of the (Non-)human Future in the Big Data World

**Ekaterina A. Koval**

DSc in Philosophy, Leading Researcher of the Institute for Corporate Education and Continuing Education.

Е.А. Коваль  
Образы (не)человеческого будущего в мире больших данных

National Research Mordovia State University.  
 68 Bolshevikskaya Str., 430005 Saransk, Russian Federation;  
 Professor at the Department of Criminal Law and Criminal Procedure.  
 Middle-Volga Institute (branch) of Russian State University of Justice.  
 6 Fedoseenko Str., 430003 Saransk, Russian Federation.  
 ORCID 0000-0003-0069-5335  
 nwifesc@yandex.ru

*Abstract.* The variability of images of the future has expanded significantly with the advent of big data and unique generative pre-trained transformers that is a type of natural language processing models. The social imaginary gradually becomes sociotechnical, and agents of artificial sociality are included in the construction of various images of the future, reproducing stereotypes from the data used to train large language models. Big data and related technologies feature in various visions of the future, including utopias, dystopias, and realistic scenarios of an achievable and livable “good society”. The emergence of the sociotechnical imaginary and generated images of the future initiates ethical and philosophical reflection, which is not limited to applied problems. However, there is also a need for exploratory studies, the results of which can be useful for further theoretical research. This article presents the results of a comparative analysis of the Russia future images generated by the GigaChat (Sberbank, Russia) and ChatGPT (OpenAI, USA). The utopia from GigaChat is focused on achieving technological leadership of Russia, and the utopia from ChatGPT is based on traditions. Dystopia according to GigaChat is a society of environmental disasters with a high level of socio-economic stratification and a low level of intellectual potential. ChatGPT constructs a Russian dystopia around an insidious small elite who, in various ways, including rewriting history, are trying to preserve their position of power. The “good society” from GigaChat is a utopia with restrictions, and from ChatGPT is an open multicultural society with a mixed economic system, focused on sustainable development and maintaining the mental health of citizens. The article demonstrates that different value and normative emphases in neural network images of the future are determined by the characteristics of the algorithms and the data used to train large language models.

**Keywords:** big data, large language models, social imaginary, sociotechnical imaginary, utopia, dystopia, “good society”, ChatGPT, GigaChat, agents of artificial sociality

**For citation:** Koval E.A. Images of the (Non-)human Future in the Big Data World // Chelovek. 2024. Vol. 35, N 6. P. 90–107. DOI: 10.31857/S0236200724060063

### Литература/ References

Ланье Д. Кому принадлежит будущее? Мир, где за информацию платить будут вам. М.: Эксмо, 2020.

Lanier J. *Komu prinadlezhit budushchee? Mir, gde za informatsiyu platit' budut vam* [Who owns the Future? A World Where You Will Get Paid for Information]. Moscow: Eksmo Publ., 2020.

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

Е.А. Коваль  
Образы (не)человеческого будущего в мире больших данных

Rezaev A.V., Tregubova N.D. *Iskusstvennyi intellekt*, «onlain-kul'tura», «iskusstvennaya sotsial'nost'»: opredelenie ponyatiy [Artificial Intelligence, On-line Culture, Artificial Sociality: Definition of the Terms]. *Monitoring obshchestvennogo mneniya: Ekonomicheskie i sotsial'nye peremeny*. 2019. N 6. P. 35–47.

Сычев А.А. Экологическая этика и дисконтирование будущего // Вестник Российской университета дружбы народов. Серия: Философия. 2014. № 2. С. 152–160.

Sychev A.A. *Ekologicheskaya etika i diskontirovanie budushchego* [Ecological Ethics and Discounting of the Future]. *Vestnik Rossiiskogo universiteta druzhby narodov. Seriya: Filosofiya*. 2014. N 2. P. 152–160.

Тейлор Ч. Секулярный век. М.: ББИ, 2017.

Taylor Ch. *Sekulyarnyi vek* [A Secular Age]. Moscow: BBI Publ., 2017.

Урри Дж. «Как выглядит будущее?». М.: Издательский дом «Дело» РАНХиГС, 2018.

Urry J. «*Kak vyglyadit budushchee?*» [«What is the Future?»]. Moscow: Izdatel'skii dom «Delo» RANKhiGS Publ., 2018.

Федотова В.Г. Хорошее общество. М.: Прогресс-Традиция, 2005.

Fedotova V.G. *Khoroshee obshchestvo* [The Good Society]. Moscow: Progress-Traditsiya Publ., 2005.

Askari A., Aliannejadi M., Kanoulas E., Verberne S. A Test Collection of Synthetic Documents for Training Rankers: ChatGPT vs. Human Experts. *Proceedings of the 32nd ACM International Conference on Information and Knowledge Management*. New York: Association for Computing Machinery, 2023. P. 5311–5315.

Hasselbalch G. *Data Ethics of Power: A Human Approach in the Big Data and AI Era*. Cheltenham; Northampton, Massachusetts: Edward Elgar, 2021.

Jasanoff Sh., Kim S.-H., Sperling S. Sociotechnical Imaginaries and Science and Technology Policy: a Cross-national Comparison. *National Science Foundation Research Project Proposal*. Cambridge, Massachusetts: Harvard University, 2007. URL: <https://stsprogram.org/admin/files/imaginaries/NSF-imaginaries-proposal.pdf> (date of access: 13.02.2024).

Krügel S., Ostermaier A., Uhl M. Zombies in the Loop? Humans Trust Untrustworthy AI-Advisors for Ethical Decisions. *Philosophy & Technology*. 2022. Vol. 35, N 17. P. 1–37.

Liu M., Liu R., Wang H., Buntine W.L. (2024) A Survey on the Real Power of ChatGPT. URL: <https://arxiv.org/abs/2405.00704v1> (date of access: 30.04.2024).

Perez E., Ringer S., Lukošiūtė K., et all. Discovering Language Model Behaviors with Model-Written Evaluations. URL: <https://arxiv.org/abs/2212.09251> (date of access: 01.05.2024).

Pham V., Cunningham S. ChatGPT Can Predict the Future when it Tells Stories Set in the Future About the Past. *SSRN Electronic Journal*. 2024: n. pag.

Pock M., Ye A., Moore J. LLMs Grasp Morality in Concept. URL: <https://arxiv.org/abs/2311.02294> (date of access: 19.02.2024).