

ОБРАЗЫ ВОСПРИЯТИЯ И КОНЦЕПТУАЛИЗАЦИЯ АНТРОПОЛОГИЧЕСКИХ ВЫЗОВОВ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Т. А. Сидорова

Новосибирский государственный университет, Россия
t.sidorova@g.nsu.ru

Исследование выполнено при финансовой поддержке гранта РНФ
(проект № 21-18-00103 «Человек и новый технологический уклад.
Антропологический форсайт»).

В статье артикулируются антропологические вызовы искусственного интеллекта (ИИ) в модусе концептуализации и восприятия рисков и угроз, благ и выгод, происходящих от новой технологии.

Образы антропологических вызовов находят разные формы репрезентации в научных концептах и философской рефлексии, в визуализациях в современных видах искусства, в компьютерных играх, кинематографе, институционализированы в правилах этических руководств. Все они могут быть рассмотрены как поиск ответов на проблематизацию человека, его субъектности, целостности, открытости, которые подвергаются риску в технологиях ИИ. Образы восприятия канализированы в позиции в отношении к ИИ и одновременно определяются практиками его широкого внедрения. Концепт ИИ формируется в лексическом топосе осмысления цивилизационного вызова. Понятие «искусственный интеллект» превращается в метафору широкого порядка, порождающую множественные концептуальные модификации. Концепт ИИ, соединяя метафорическое и понятийное, выполняет функцию «оестествления», «опривычивания» технологии.

Особенностью в обобщении позиций в отношении к искусственному интеллекту является их нелинейность и целевое формирование. Рассмотрены три варианта оформления образов антропологических вызовов ИИ: алармистский, инструменталистский (профессиональный) и утилитарный (пользовательский). Коллективный ответ на антропологические вызовы ИИ вероятно будет строиться на утилитарно-прагматической основе, концептуально и институционально репрезентированный в этическом регулировании. Для нивелирования антропологических рисков действенными могут быть индивидуальные ответы на основе самосохраняющей стратегии и когнитивной гигиены, начиная со сферы образования. Разработка правил и процедур такой сохраняющей стратегии – задача, которая встает в контексте развития ИИ. Гуманитарная экспертиза нейросетей может стать частью этой стратегии.

Ключевые слова: искусственный интеллект, нейросети, концепт ИИ, антропологические вызовы, образы восприятия, этический кодекс ИИ, гуманитарная экспертиза

PERCEPTION IMAGES AND CONCEPTUALIZATION OF ANTHROPOLOGICAL CHALLENGES OF ARTIFICIAL INTELLIGENCE

Tatyana A. Sidorova

Novosibirsk State University, Novosibirsk, Russia
t.sidorova@ng.su.ru

The challenges of artificial intelligence are considered from the methodological basis of bioethical analysis of anthropological risks and threats posed by new technologies. Society exhibits a cautious attitude towards artificial intelligence technology. Anthropological challenges of artificial intelligence represent a problematic situation regarding the complexity of assessing the benefits and harms, adequate awareness of the risks and threats of new technology to humans. It is necessary to conceptually outline the anthropological challenges of AI, drawing on images of AI perception represented in art and cinema, in ethical rules, philosophical reflection, and scientific concepts. In the projection of various definitions, artificial intelligence becomes a metaphor that serves as a source of creative conceptualizations of new technology. Images of AI are identified through conceptualization, visualization, and institutionalization of risks and correspond to specific types of attitudes towards innovation in society. The peculiarity of AI perception images, both in the forms of conceptualization and in the visual or institutional objectification of these images in ethical codes, is their active and purposeful formation.

Analogous to the regulation of biotechnologies, normatively conceptualized positions regarding new technologies are divided into conservative - restrictive and prohibitive; liberal - welcoming innovations; and moderate - compromising, which often becomes the basis for ethical and legal regulation. However, sociological surveys show that those who welcome the emergence of neural networks, the widespread use of artificial intelligence, also exhibit caution and uncertainty in assessing the human future.

A three-part typology of perception images of anthropological challenges is proposed, in which non-linear opposition of positions towards AI is fixed, but vectors of possible ways of habituating and semiotization of the future are outlined. The first, alarmist type, is distinguished based on an emotionally evaluative attitude. New technologies are seen as redundant, causing alarm and fear. The second type of perception, instrumentalist, is characteristic of AI actors within a professionally formed worldview. Some concepts of the professional thesaurus become common parlance. The third type is user-oriented. For this type, it is important how the interaction between AI and humans unfolds.

The collective response to the anthropological challenges of AI is more likely to be formed on a utilitarian-pragmatic basis. Effective responses may be based on an individual self-preservation strategy, which, for example, may require adherence to cognitive hygiene in the field of education. In the context of AI devel-

opment, the task arises of developing rules and procedures for such a preservation strategy.

Keywords: artificial intelligence, neural networks, AI concept, anthropological challenges, perception images, AI ethical code, humanistic expert evaluation

DOI 10.23951/2312-7899-2024-1-102-119

Введение

Использование искусственного интеллекта (ИИ) сулит прорывы на новые уровни общественного прогресса, заманчивые экономические выгоды и технологические достижения, повышение уровня жизни и креативности, обретение новых форм управления и власти, способов продвижения интересов технологически развитых стран. Приоритетное значение в технологиях ИИ отводится разработке умных нейросетей, чтобы ускоренно использовать их в прикладных целях в госуправлении, медицине, банковской и производственной сфере, торговле, образовании. Мир столкнулся с очередным и в некоторой степени неожиданным вызовом, который провоцирует появление технологий искусственного интеллекта. На протяжении последних десятилетий среди технологических вызовов господствовал дискурс биотехнологических рисков, источником тревожных ожиданий, страхов и неопределенности были новые репродуктивные, генетические, нейро- и другие технологии. Цифровая трансформация, видоизменившая практически все формы человеческой деятельности, включая коммуникацию, казалось бы, сформировала почву для толерантного восприятия новой формы технологического улучшения человеческой природы – искусственного интеллекта. Однако пугающими являются масштабы изменений, скорость распространения систем ИИ, возможность конкурировать с человеком, вытеснять его из привычных сфер занятости, влиять на самореализацию. Искусственный интеллект недвусмысленно поставил вызов человеку как уникальной форме бытия, и это не отсроченный, а актуальный, трудноуправляемый риск, он связан с онтологическим соблазном – самоустраниением человека [Смирнов 2023, 29–30]. «Если искусственный интеллект будет предлагать человеку удобные, эффективные, выгодные решения, организовывать все более психологически комфортные жизненные условия через минимизацию внутреннего напряжения человека, связанного с вы-

бором линии его поведения и ответственности за принятое решение, то это приведет к вырождению моральных и волевых качеств человека и даже к деградации умственных способностей. Заурядному обывателю этот тренд жизни, ориентированный на потребление, когда ответственность перекладывается на кого-то (в данном случае на искусственный интеллект), будет, несомненно, удобен и в значительной степени привлекателен» [Глуздов 2022, 7].

Одновременно в восприятии ИИ отчетливо возникают сомнения: насколько его следует опасаться, какие границы и красные линии должны быть очерчены, не является ли вообще существующая настороженность первой инстинктивной реакцией, за которой последуют «натурализация», «приручение» искусственного интеллекта и сращивание его с человеческим. Практическим шагом нормализации новой технологии является разработка этических правил для акторов искусственного интеллекта. Однако к акторам относят причастных к разработке и продвижению ИИ, максимум заказчиков [Кодекс 2022], а массовый потребитель, тот самый «заурядный обыватель», то есть просто человек, – для него тоже понадобятся правила взаимодействия с ИИ, и будет ли он придерживаться правил? Будет ли он задумываться над антропологическими рисками или логика повседневности поглотит эти сомнения, отсеяв их через фильтры утилитарно обоснованной выгоды?

Перед исследователями социогуманитарных импликаций новых технологий возникает масштабная задача концептуально оформить антропологические вызовы ИИ [Perrault et al. 2019]. В условиях, когда практика развития ИИ опережает теорию даже в части работы алгоритмов и тем более в осмыслении социальных и антропологических следствий, важно систематизировать образы восприятия рисков ИИ, понять, как отражается существующая настороженность в попытках этического регулирования ИИ, в концептуализации антропологических рисков. Эта задача апеллирует к феноменологии, семиотике, герменевтике в аспекте изучения социокультурных паттернов восприятия угроз и рисков, связанных с новыми технологиями.

В изучении позиций в отношении новых технологий исследование опирается на данные социологических опросов и методологический багаж в биоэтике. Хотя биоэтика и сосредоточена на выявлении рисков, возникающих в свете развития биомедицинских технологий, тем не менее умные интеллектуальные системы универсальны, они также находят применение в медицине, и опыт биоэтического регулирования начинает распространяться

и на них. Методологическим руководством также будут служить концептуальные идеи видного российского биоэтика Б. Г. Юдина в аспекте выделения границ человеческого существования между человеком и животным, человеком и машиной, жизнью и смертью [Юдин 2011].

Репрезентации образов восприятия

Под выражением «образы восприятия антропологических вызовов» подразумевается представленность в сознании и языке тех угроз и рисков, а также благ и выгод, которые ассоциируются с искусственным интеллектом. По мнению И. М. Кондакова, «образ восприятия обусловлен как чувственно воспринимаемыми признаками, так и гипотетическими конструктами. Являясь основой для реализации практических действий по овладению объектами окружающего мира, образ восприятия также определяется характером этих действий, в процессе которых исходный образ видоизменяется, все более удовлетворяя практическим нуждам» [Кондаков 2007, 389]. Из этого следует, что образы восприятия могут быть выражены через концептуализирующие практики мышления и продуцирующую деятельность воображения в визуальных образах, создавая различные в своей семиотике способы обозначения и проживания антропологических рисков. Одновременно они репрезентированы концептуально и институционально в системах этических правил и юридических норм, которые являются ответами на антропологические вызовы ИИ. Таким образом, концептуализация, визуализация и институционализация запечатлевают образы ИИ, определяя отношение к инновациям. Ж. Бодрийяр писал: «Мы существуем внутри культуры иррадиации умов и тел знаками и образами. <...> Можно предположить, что фантастический успех искусственного разума вызван тем, что этот разум освобождает нас от разума природного, <...> освобождает от двусмысленности мысли и от неразрешимой загадки ее отношений с миром» (цит. по: [Панова 2022, 337]). Как считает Д. Е. Скворцов, образное мышление адаптирует основания мышления к подвижной антропологической реальности, объединяя образы и понятия в устойчивой динамичности познания [Скворцов 2015, 97].

Под антропологическими вызовами искусственного интеллекта будем понимать проблемную ситуацию, суть которой – в осознании угроз и рисков, в сложности оценки блага и вреда для человека, возникающих с внедрением новой технологии. Взвешивая следствия распространения ИИ для человека и общества, их сравнива-

ют с мечом: лезвия позитивных и негативных ожиданий одинаково заострены. Язык символов продуктивен в освоении новых для человека реальностей, а знаково-символьная диагностика¹ вскрывает невидимую суть нормативных трансформаций. Поскольку речь идет о вызове и мы имеем дело с модальностью, требующей ответа, то данное понятие имплицитно подразумевает поиск способов и путей преодоления того, что вызывает опасения, или использования выгод и приобретений, которые появляются с получением искусственных интеллектуальных систем.

Риски для человека, с которыми сталкивается современный мир, связаны в первую очередь с утратой человеком субъектности с точки зрения возможности принимать решения. Это вызов человеческой природе, когда у машины развивают способность, до сих пор принадлежавшую только человеку, – способность к интеллектуальной деятельности, тем самым стирая грань между человеческим и нечеловеческим, природным и искусственным. Человек, передавая свое исключительное преимущество машинному интеллекту, превращается в существо, алгоритмизированное и предсказуемое, расстается с контингентностью собственного выбора, уникальностью как основанием своей незавершенности и открытости миру. Когда человек полагается на ИИ, у него изменяются когнитивные свойства, способность к запоминанию, трансформируется память в качестве элемента самоидентификации. «С философской точки зрения с появлением искусственного интеллекта извечная стараниями классиков обоснованная взаимосвязь бытия и сознания в цифровую эпоху поставлена под вопрос иным образом. Теперь и Cogito Р. Декарта, и чистый разум И. Канта и других европейских классиков, и разум “жизненный” неклассической философии и философии всеединства, и трансцендентальное Ego, и трансцендентальная память, и творческий гений, и абсолютный дух Г. Гегеля, и мировая душа Ф. Шеллинга, и даже более современное “Сознание, прикованное к плоти” С. Зонтаг уходят в прошлое и теряют значимость. Они вынуждены уступать место новому герою – искусственному интеллекту, стремительно совершенствующемуся и уверенно обретающему всемогущество, господство над миром и власть над человечеством» [Панова 2022, 336–337]. Происходит вытеснение человека машиной из производственных процессов, исчезает целый ряд профессий,

¹ Примером служит набор процедур в концепции «семиотической диагностики», разработанной на основе оригинального подхода к исследованию информационных процессов и нелинейной динамики социокультурных систем [Мелик-Гайказян 2022].

прогнозируется, что системы ИИ благодаря своей сверхэффективности возьмут на себя большую часть человеческой работы, так что людям больше не придется зарабатывать на жизнь трудом [Knell, Rüther 2023]. Возникают новейшие формы неравенства, например в силу возраста когнитивно обусловленная недоступность паритетного взаимодействия с алгоритмами. Разработчики скорее форматируют программы под себя, нежели пытаются к ним приспособить других людей. «Идущая уже несколько десятилетий и набирающая темпы антропотехнологическая эволюция способна привести к существенным, возможно, качественным преобразованиям нашего интеллекта, творческих способностей, нашей телесности и в силу этого основного комплекса наших потребностей, а тем самым к трансформации социума» [Дубровский 2022, 103].

В свою очередь, технологии ИИ несут очевидные выгоды человеку: оптимизируют коммуникационные процессы, ускоряют обработку огромных массивов информации, обеспечивают общественную безопасность, упрощают покупки и улучшают медицинское обслуживание, избавляют человека от бюрократических процедур. Приобретения так необходимы и так велики, что принятие новых технологий происходит довольно быстро.

Институционализация рисков ИИ через создание регулирующих правил сопровождает разработку и внедрение технологии, репрезентируя через этические кодексы гуманитарную рефлексию, общественные и экспертные обсуждения. Возникновение и пользовательская адаптация технологии ИИ – процесс по преимуществу виртуализированный и распределенный, для него необходимо институциональное упорядочение не просто на уровне создания организационной архитектуры [Schultz, Seele 2022], а на уровне ментальном, с помощью институтов, являющихся «правилами игры».

Визуализация в кинематографических произведениях, компьютерных играх, кибер-арте является предваряющей и отражающей проекцией, подготавливает воображение к переменам, ищет новые предметные формы для технологий. На игры как способ репрезентации образов восприятия современной цифровой реальности, подготовившей встречу с искусственным интеллектом, следует обратить особое внимание. Хотя смысл в играх создается не через отражение мира, но через селективное моделирование его компонентов, репрезентация мира через игру формирует мышление посредством виртуальных паттернов, создает иллюзию «свободы» действий и «свободного» выбора в рамках алгоритмов, запрограммированных в видеоиграх, и становится средой для формирования образов восприятия искусственного интел-

лекта. Восприятие искусственного интеллекта подготовлено «исчезновением» физического тела в феномене кибертела. «Такое отделение от тела нередко встречается в киберкультуре, где большое количество людей проводит время в режиме неподвижного слежения, уткнувшись глазами в экран монитора» [Штайн 2010,101].

Концептуализация «мыслящих» систем

Образы восприятия, возникая как форма чувственного отношения к объекту, репрезентируются в концептах и понятиях как результат стихийной или специализированной теоретической формы рефлексии, дискурсивной «обкатки» рождающихся обозначений новых феноменов.

Предваряющим элементом анализа концептуализации антропологических рисков является экспликация понятия «искусственный интеллект». Существует несколько определений ИИ. Приведем юридически закрепленное в российском ГОСТ: «Искусственный интеллект – моделируемая (искусственно воспроизводимая) интеллектуальная деятельность мышления человека» [Андреев и др. 2022, 19]. Не претендуя на составление исчерпывающего перечня понятий, которые можно включать в поле концепта «искусственный интеллект», представим их по мере того, как они ассоциируются с антропологическими вызовами и возрастанием рисков для человека. То, что мы сегодня подразумеваем под искусственным интеллектом, может называться так: умная машина, машинный интеллект, умные алгоритмы, классический ИИ, интеллектуальная система, цифровой интеллект, органоидный интеллект, сильный и слабый ИИ, специализированный ИИ, прикладной искусственный интеллект, адаптивный ИИ, автономный ИИ, самообучающиеся интеллектуальные системы, нейронные сети, искусственные нейронные сети, генеративные нейронные сети, генетические алгоритмы, гипотетический ИИ, универсальный ИИ, общий искусственный интеллект, искусственный интеллект человеческого уровня, искусственный суперинтеллект, сверхразумный интеллект, сильный сверхразум, ИИ сверхчеловеческого уровня.

Усложнение технологии отражается в притязаниях на универсальность ИИ и соответствие уровню человеческого мышления. История становления искусственного интеллекта, бросающего вызов человеку, как понятия показывает, что термин «общий искусственный интеллект» (ОИИ) появился в 1997 году, когда Deep Blue обыграл чемпиона мира в шахматы. В рамках этой идеи

было предложено делить ИИ не на слабый и сильный, как это сделал Д. Серль, а на прикладной, или узкий, ИИ (Artificial Narrow Intelligence, ANI, или Narrow AI, или AI), универсальный, или общий (Artificial General Intelligence, AGI), и гипотетический ИИ (Artificial Superintelligence, ASI). Современную историю ИИ отсчитывают с 2010 года, когда мощность компьютеров позволила сочетать технологию больших данных (Big Data) с методами глубокого обучения (Deep Learning), которые основываются на использовании искусственных нейронных сетей [Мамина, Почебут 2022, 69–72].

Определяя, что подразумевается по искусственным интеллектом, акцент делается на дедуктивных (вычислительных, аналитических) способностях разума. Поэтому в целях концептуализации как нельзя кстати оказывается понятийная дифференциация в обозначении разумной способности человека, разработанная в философии: ум, рассудок, разум, мышление, интеллект, сознание, идеальное, когнитивная деятельность, воображение. Из этих обозначений понятие «искусственный интеллект» более точно соответствует современной исторической ступени в разработке искусственных «мыслящих» систем, в нем есть идентифицирующий признак разумной деятельности, который имеет значение для широкого прикладного использования, – способность к самообучению, то, что приближает его к человеческой способности самосознания. Алгоритмы перерастают в гибкие, реактивные синтезы, способные к генерации решений, удивляющих человека своей креативностью.

Концепт ИИ оформляется через структурирование лексического топоса: обозначение «искусственный интеллект» выступает общим, родовым понятием, превращается в «макрометафорическую концептуальную модель», порождающую множественные креативные модификации [Зыкова 2022, 316], становится дериватом для появления других понятий, обозначающих как частные вариации искусственного разума, так и степени его универсализации. Это иллюстрирует то, как концептуализация связана с образностью мышления. Метафора и понятие – два полюса концептуализации; восприятие вызовов искусственного интеллекта нуждается в метафорах, адаптирующих технические термины в обыденное употребление. Примером этому является обозначение искусственного интеллекта как нейросети, уподобляя алгоритмы нейронной сети в человеческом мозге. По мере опривычивания их стали называть «сетки», что также указывает на множественность функционала: сетки для распознавания лиц, для принятия врачебных решений,

для идентификации мигрантов и т.д. «Сетка» визуально заявляет о себе в нашем представлении, становится знаком, включая чувственный канал восприятия, присоединяет эмоциональный – пренебрежительный – оттенок и указывает на ограниченные возможности нейросетей сегодня, и еще как будто намекает на то, что будущие «сетки» будут более совершенными и интеллектуальными. Получая имена собственные, самообучаемые «сетки» опрощаются для удобства обыденного словоупотребления, как, например, нашу-мевшая нейросеть, которая впервые создала дипломную работу, сдала экзамены в американских университетах на лицензию врача, юриста, – ChatGPT. «Сетка» стала именоваться в России «чатгптопой». Так в словесной эквилибристике в Интернете происходит обживание новой технологии в языке, ее присвоение.

Типологизация образов восприятия ИИ

У А. Дж. Тойнби под вызовом понималась проблемная ситуация, решение которой становится узловым моментом цивилизационного развития. Ответ цивилизационному вызову начинается с того, как вызов будет воспринят и отражен в образно-понятийной форме. Концептуализация антропологических вызовов искусственного интеллекта берет начало в различных позициях по отношению к искусственному интеллекту, язык и семиотика которых чрезвычайно разнообразна: от академических дискуссий до киберпанка, от этических кодексов до обывательских страхов. Полагаем, что концепт приобретает завершенную форму в понятии, выполняющем нормативную функцию. Если полярно кластеризировать нормативно обозначаемые позиции в отношении к новым технологиям, опираясь на накопленный опыт регулирования биотехнологий, то они разделяются на консервативные – часто ограничительные и запретительные, либеральные – приветствующие инновации, и выделяется срединная позиция – умеренная. Как правило, умеренная позиция находит необходимый компромисс, становится основой этико-правового регулирования, чтобы обойти острые углы крайних позиций, учитывать значимые ценности, защищаемые в консервативном подходе, а также найти способы принятия и продвижения инновационных технологий. Но если применить эту трехчастную структуру к искусственному интеллекту не просто как к цифровому инструменту, а как к антропологическим следствиям, которые мы должны учитывать, определяя русло развития новой технологии и вырабатывая этику ее использования, можно заметить, что те, кто

приветствует появление нейросетей, широкое распространение искусственного интеллекта, тем не менее также проявляют настороженность и неопределенность в оценке человеческого будущего.

Социологи в целом фиксируют устойчивые тенденции принятия технологий, особенно молодежью. При этом у всех сохраняется настороженное отношение к искусственному интеллекту там, где он ассоциируется с принятием решений, где ему могут быть передоверены исключительно человеческие прерогативы. Данные опросов показывают, что образ ИИ и отношение к нему в сознании молодежи оказались противоречивыми: «максимально полезный и максимально опасный», «хорошо известный, но нераспознанный» [Ясин 2022, 199].

Опросы россиян, проведенные ВЦИОМ, демонстрируют, что растет число тех, кто доверяет ИИ. За 2022 год тех, кто доверяет, стало больше на 10 процентных пунктов, их 51%, не доверяют 32% опрошенных. Подавляющее большинство россиян видит в развитии технологий ИИ как позитивные, так и негативные последствия (91 и 93% соответственно). Одновременно треть россиян опасается замены человека технологиями искусственного интеллекта в их профессии. При оказании медицинских и образовательных услуг использование специалистами ИИ допустимо в случае, если решения принимает человек, а искусственный интеллект только помогает, – так считают 74 и 62% соответственно [ВЦИОМ 2022].

Исследователи отмечают, что культурные ценности индивидуализма, эгалитаризма, общее неприятие риска и техноскептицизм являются важными факторами отношения к ИИ. Воспринимаемая выгода влияет на отношение к использованию ИИ, но не к его управлению.

Образы восприятия антропологических вызовов могут быть также рассмотрены в трехчастной классификации. Однако ее отличие заключается в том, что фиксируется не линейное противопоставление позиций в отношении к ИИ, а векторы возможных способов опривычивания, семиотизации будущего [Мелик-Гайказян 2022], свои концептуализации и нормативные стратегии.

Особенностью образов восприятия ИИ как в формах концептуализации и метафоризации, так и в визуальном или институциональном опредмечивании этих образов в этических кодексах является их активное и целевое формирование. Агентами этого процесса выступают те, кто разрабатывает и продвигает новые технологии. В результате появляются, например, этические кодексы. Они представляют собой ответ на антропологические вызовы со

стороны морального сообщества, защищающего свои ценности, но иницируются самими разработчиками [Lauer 2021, 21]. Это, с одной стороны, служит шагом навстречу потребителям, которые испытывают опасения перед новой технологией, а с другой стороны, это институциональные регулирующие правила для различных игроков на рынке ИИ.

Представим три парадигмы, в которых могут формироваться образы восприятия антропологических вызовов ИИ. Это три основожных типа отношения к ИИ. Первый выделяется на основе эмоционально-оценочного отношения, назовем его алармизмом. Разумеется, оценки могут быть и позитивными, и негативными, и нейтральными, однако алармизм как способ восприятия ИИ широко распространен и может сопровождать в том числе восторженное приятие ИИ, что отражают амбивалентные оценки, зафиксированные в социологических опросах. Тем не менее для этого типа характерно настороженное отношение к ИИ, антропологический вызов осознается как опасность. Новые технологии рассматриваются как избыточные, вызывающие тревогу и страх. На основе алармистских образов создаются пессимистические сценарии будущего.

Второй тип восприятия свойствен носителям специализированного мировоззрения акторов ИИ, назовем его инструменталистским. Вот типичные утверждения в оптике инструменталистского видения: «системы, построенные на принципах глубинного обучения, не обладают ИИ, это не что иное, как новый и более сложный, чем программирование, способ использования компьютеров в качестве инструмента для анализа данных»; «компьютер был и будет инструментом для расширения интеллектуального потенциала человека, и главная задача заключается не в создании искусственного разума AI, а в развитии систем, которые называют *Intelligence amplification* (усиление интеллекта), *Cognitive augmentation* (когнитивное усиление) или *Machine augmented intelligence* (машинное усиление интеллекта)» [Черняк 2019]. Поскольку системы ИИ представляют собой чрезвычайно сложную по содержанию область профессиональной деятельности, то и оценка рисков и возможностей новых технологий в рамках профессиональной компетентности значительно отличается. Профессиональный дискурс стремится к популяризации разработок, и грань между специальным и неспециальным, профанным или пользовательским довольно подвижна. ИИ рассматривается в функциональном аспекте. Концептуально инструменталистское восприятие оформляется в рамках профессионального тезауруса, и часть этих понятий ста-

новится общеупотребительной. Инструменталистские образы ИИ не связывают его с субъектностью. Считается, что ИИ если и несет вред, то его источником служит оператор. ИИ – помогающий инструмент, позволяет избежать человеческих ошибок, поскольку в нем нет возможности импровизировать, нет внезапности, он более надежен в своих решениях.

Третий тип – утилитарный, или пользовательский. Антропологические вызовы в данном случае не являются предметом рефлексии. Для данного типа не принципиален вопрос о субъектности ИИ, в меньшей степени волнует вопрос о работе алгоритмов, прозрачны они или нет, важно в первую очередь то, как складывается взаимодействие с человеком. Эта версия может напоминать инструменталистскую. Однако данную категорию формируют другие носители – не те, кто разрабатывает ИИ, а те, кто его использует в качестве потребителя (акторы ИИ – частный случай пользователей). Они относятся к ИИ, стремясь извлечь пользу из созданной технологии, так, чтобы она блага приносила больше, чем вреда. Для «компетентного пользователя» важно думать о ИИ как об «оестественном интеллекте», который становится частью человеческого. Недостатки технологии связываются с институциональными системами, использующими ИИ: в здравоохранении, образовании, госуправлении есть свои пороки, именно они влияют на эффективность ИИ. Позиция транслируется через виды прикладной этики: Этики ИИ, Алгорэтики [Tsamados et al. 2022].

В выделенных типах восприятия ИИ и сопутствующих антропологических угроз не идет речь об однозначной принадлежности к консервативному, умеренному или либеральному крылу. В данном случае образ восприятия антропологических угроз ИИ – стратегия, которая определяет мыслительный, дискурсивный, семиотический горизонт индивидуального сознания, определяющего свое отношение к ИИ. Это феноменология новой реальности, возникающая в контексте распространения искусственного интеллекта. За каждой из стратегий стоят свои представления о человеке, идентичности человеческого, о его границах в отношении с машинным, не-человеческим. Это варианты, одни из которых могут задавать акцент, другие оставаться «скрытыми» и тем не менее пробивать себе путь к осуществлению [Мелик-Гайказян 2022].

Заключение

Во время работы над данной статьей на мониторе появилось предложение взять помощника на основе искусственного интел-

лекта. Из научных соображений отказаться было бы неверно: ведь нужно проверить, как это работает. Но параллельно возникли сомнения: почему так назойливо предлагают, и вообще, что это за помощник такой, и без него обходились, – которые заставили отказаться от предложения. Однако прецедент убеждает, что через некоторое время выбора не будет: так же как сегодня мы не можем отказаться от смартфона, не сможем работать без умных помощников. По сути, End Note, Mendeley и множество других поисковых библиографических ресурсов – те же алгоритмы, которые еще не называли искусственным интеллектом. С ними ученые сработались и готовы принимать следующие предложения.

Утилитаризм как этическая и теоретическая платформа широко используется в биоэтике для обоснования того, почему новые технологии переделки человеческой природы должны работать, если они созданы. Искусственный интеллект будет обеспечивать конкурентные преимущества индивидам и странам, поэтому утилитаристская пользовательская стратегия восприятия ИИ будет реализована в регуляторных практиках, позволяющих максимизировать блага. Происходящая институционализация этики в рамках ИТ-индустрии явно противоречит исследовательской свободе и независимости суждений, поскольку корпорациям невыгодно подвергать сомнению возможности нового продукта [Fried 2021]. Что касается вреда, то антропологические риски должны быть концептуализированы так, чтобы их влияние можно было оценить здесь и сейчас, как если бы они имели явно выраженный характер. Трансформации, происходящие с когнитивными способностями человека вследствие пластичности человеческой психики и высокой адаптационной способности социализированной личности, не будут восприниматься в качестве непосредственной угрозы, чтобы стать поводом для серьезных ограничений. Таким образом, коллективный ответ на антропологические вызовы ИИ, скорее, будет формироваться на утилитарно-прагматической основе. При этом действенными могут оказаться ответы, построенные на основе индивидуальной самосохраняющей стратегии, которая потребует, например, соблюдения когнитивной гигиены в сфере образования. Разработка правил и процедур такой сохраняющей стратегии – задача, которая встает в контексте развития ИИ. Частью таких процедур могут стать инициативная форма гуманитарной экспертизы конкретных «нейросеток», разрабатываемых для широкого применения, насыщение пользовательского восприятия информацией о необходимости гуманитарной оценки антропологических и социальных последствий внедрения умных помощников.

БИБЛИОГРАФИЯ

- Андреев и др. 2020 – *Андреев В. К. и др.* Искусственный интеллект в системе электронного правосудия при рассмотрении корпоративных споров // Вестник Санкт-Петербургского университета. Право. 2020. Т. 11, № 1. С. 19–34.
- ВЦИОМ 2022 – Искусственный интеллект: угроза или светлое будущее? : аналитический обзор // ВЦИОМ. 2022. 28 дек. URL: <https://wciom.ru/analytical-reviews/analiticheskii-obzor/iskusstvennyi-intellekt-ugroza-ili-svetloe-budushchee>
- Глуздов 2022 – *Глуздов Д. В.* Философско-антропологические основания взаимодействия искусственного и естественного интеллекта // Вестник Мининского университета. 2022. Т. 10, № 4. Ст. 15. URL: <https://www.minin-vestnik.ru/jour/article/view/1347/928>
- Дубровский 2022 – *Дубровский Д. И.* Развитие искусственного интеллекта и глобальный кризис земной цивилизации (к анализу социогуманитарных проблем) // Философия науки и техники. 2022. Т. 27, № 2. С. 100–107.
- Зыкова 2022 – *Зыкова И. В.* Трансдисциплинарная модель развития научных воззрений на природу лингвокреативности: философия vs психология vs семиотика vs лингвистика // Человек как открытая целостность. Новосибирск: Академиздат, 2022. С. 306–318.
- Кодекс 2022 – Кодекс этики в сфере ИИ // Альянс в сфере искусственного интеллекта. URL: <https://a-ai.ru/ethics/index.html>
- Кондаков 2007 – *Кондаков И. М.* Психология: иллюстрированный словарь. СПб.: Прайм-Еврознак, 2007.
- Мамина, Почебут 2022 – *Мамина Р. И., Почебут С. Н.* Искусственный интеллект в оптике философской методологии: образовательный трек // ДИСКУРС. 2022. Т. 8, № 1. С. 64–81.
- Мелик-Гайказян 2022 – *Мелик-Гайказян И. В.* Семиотическая диагностика расщепления траекторий мечты о прошлом и мечты о будущем // Электронный научно-образовательный журнал «История». 2022. Т. 13, № 4 (114). DOI: 10.18254/S207987840021199-7
- Панова 2022 – *Панова О. Б.* Человек эпохи smart-technologies: рефлексия над антропологической проблематикой современности (философия – наука – искусство) // Человек как открытая целостность. Новосибирск: Академиздат, 2022. С. 328–341.
- Скворцов 2015 – *Скворцов Д. Е.* Оппозиции и метафоры образного мышления в культуре // Вестник Волгоградского университета. Сер. Философия. 2015. № 2 (28). С. 95–99.

- Смирнов 2023 – Смирнов С. А. Соблазн не быть, или онтологические корни технологического аутсорсинга // Человек. 2023. Т. 34, № 1. С. 28–50.
- Черняк 2019 – Черняк Л. Сексизм и шовинизм искусственного интеллекта. Почему так сложно его побороть? // TAdviser.ru: государство, бизнес, технологии. 2019. 26 февр. URL: http://www.tadviser.ru/index.php/Статья:AI_bias
- Штайн 2010 – Штайн О. А. Трансформация телесности в современном мире // Вестник Удмуртского университета. Философия. Психология. Педагогика. 2010. Вып. 1. С. 99–102.
- Юдин 2011 – Юдин Б. Г. Границы человеческого существа как пространства техно-логических воздействий // Вопросы социальной теории. 2011. Т. V. С. 102–118.
- Ясин 2022 – Ясин М. И. Представления молодежи об искусственном интеллекте и отношение к нему // Известия Саратовского университета. Новая серия. Сер. Философия. Психология. Педагогика. 2022. Т. 22, вып. 2. С. 197–201.
- Fried 2021 – Fried I. Google fires another AI ethics leader // Axios. 2021. Feb. 20. URL: <https://www.axios.com/2021/02/19/google-fires-another-ai-ethics-leader>
- Knell, Rüther 2023 – Knell S., Rüther M. Artificial intelligence, superefficiency and the end of work: a humanistic perspective on meaning in life // AI Ethics. 2023. April. DOI: 10.1007/s43681-023-00273-w
- Lauer 2021 – Lauer D. You cannot have AI ethics without ethics // AI Ethics. 2021. Vol. 1. P. 21–25.
- Perrault R et al. 2019 – Perrault R et al. Artificial Intelligence Index Report. Stanford University, 2019.
- Schultz, Seele 2022 – Schultz M. D., Seele P. Towards AI ethics' institutionalization: knowledge bridges from business ethics to advance organizational AI ethics // AI Ethics. 2022. Vol. 3. P. 99–111.
- Tsamados et al. 2022 – Tsamados A. et al. The ethics of algorithms: key problems and solutions // AI & Soc. 2022. Vol. 37. P. 215–230.

REFERENCES

- A-AI. (2022). *AI Code of Ethics*. <https://a-ai.ru/ethics/index.html> (In Russian).
- Andreev, V. et al. (2020). Artificial intelligence in the system of electronic justice by consideration of corporate disputes. *Vestnik of Saint Petersburg University. Law*, 1, 19–34. (In Russian).

- Chernyak, L. (2019). Sexism and chauvinism of artificial intelligence. Why is it so difficult to overcome it? *TAdviser.ru*. February 26. http://www.tadviser.ru/index.php/Article:AI_bias (In Russian).
- Dubrovsky, D. I. (2022). The development of artificial intelligence and the global crisis of earthly civilization (to the analysis of socio-humanitarian problems). *Filosofiya nauki i tekhniki*, 27(2), 100–107. (In Russian).
- Fried, I. (2021). Google fires another AI ethics leader. *Axios*. Feb 20, 2021. <https://www.axios.com/2021/02/19/google-fires-another-ai-ethics-leader>
- Gluzdov, D. V. (2022). Philosophical and anthropological foundations of the interaction of artificial and natural intelligence. *Vestnik Mininskogo universiteta*, 10(4), 15.
- Knell, S., & Rütther, M. (2023). Artificial intelligence, superefficiency and the end of work: a humanistic perspective on meaning in life. *AI Ethics*, 17 April 2023. <https://doi.org/10.1007/s43681-023-00273-w>
- Kondakov, I. M. (2007). *Psikhologiya. Illyustrirovannyi slovar'* [Psychology. Illustrated Dictionary]. Prime Eurosign.
- Lauer, D. (2021). You cannot have AI ethics without ethics. *AI Ethics*, 1, 21–25.
- Mamina, R. I., & Pochebut, S. N. (2022). Artificial Intelligence in the View of Philosophical Methodology: an Educational Track. *DISCOURSE*, 8(1), 64–81. (In Russian).
- Melik-Gaykazyan, I. V. (2022). Semiotic diagnostics of the trajectory splitting between a dream of the past and dream of the future. *Istoriya*, 13:4(114). DOI: 10.18254/S207987840021199-7 (In Russian).
- Panova, O. B. (2022). Chelovek epokhi smart-technologies: refleksiya nad antropologicheskoy problematikoy sovremennosti (filosofiya – nauka – iskusstvo) [The Human of the era of smart technologies: reflections on the anthropological problems of modernity (philosophy – science – art)]. In L. P. Kiyashchenko, & T. A. Sidorova (Eds.), *Chelovek kak otkrytaya tselostnost'* [The Human as an open integrity]. Akademizdat.
- Perrault, R. et al. (2019). *Artificial Intelligence Index Report*. Stanford University.
- Schultz, M. D., & Seele, P. (2022). Towards AI ethics' institutionalization: knowledge bridges from business ethics to advance organizational AI ethics. *AI Ethics*, 3, 99–111.
- Shtayn, O. A. (2010). Body transformation in modernity. *Vestnik Udmurtskogo universiteta. Filosofiya. Psikhologiya. Pedagogika*, 1, 99–102. (In Russian).

- Skvortsov, D. E. (2015). Oppositions and metaphors of visual thinking in culture. *Vestnik Volgogradskogo universiteta. Ser. Philosophy*, 2(28), 95–99. (In Russian).
- Smirnov, S. A. (2023). The temptation of not being or ontological roots of technological outsourcing. *Chelovek*, 34(1), 28–50. (In Russian).
- Tsamados, A., et al. (2022). The ethics of algorithms: key problems and solutions. *AI & Soc*, 37, 215–230.
- WCIOM, (2022). *Artificial intelligence: a threat or a bright future? Analytical review*. <https://wciom.ru/analytical-reviews/analiticheskii-obzor/iskusstvennyi-intellekt-ugroza-ili-svetloe-budushchee> (In Russian).
- Yasin, M. I. (2022). Youth perceptions and attitudes about artificial intelligence. *Izvestiya Saratovskogo universiteta. Novaya seriya. Seriya: Filosofiya. Psikhologiya. Pedagogika*, 22(2), 197–201. (In Russian).
- Yudin, B. G. (2011). The boundaries of a human being as a space of technological and logical influences. *Voprosy sotsial'noy teorii*, 5, 102–118. (In Russian).
- Zykova, I. V. (2022). Transdistsiplinarnaya model' razvitiya nauchnykh vozzreniy na prirodu lingvokreativnosti: filosofiya vs psikhologiya vs semiotika vs lingvistika [Transdisciplinary model of developing scientific views on the nature of linguistic creativity: philosophy vs psychology vs semiotics vs linguistics]. In L. P. Kiyashchenko, & T. A. Sidorova (Eds.), *Chelovek kak otkrytaya tselostnost'* [The Human as an open integrity]. Akademizdat.

Материал поступил в редакцию 28.05.2023

Материал поступил в редакцию после рецензирования 09.09.2023