Биричева Е.В. Репрезентации философских проблем искусственного интеллекта в массовой культуре // Антиномии. 2025. Т. 25, вып. 4. С. 32-53. https://doi.org/10.17506/26867206_2025_25_4_32

УДК 111:165.12:004.03:004.5:004.89
DOI 10.17506/26867206_2025_25_4_32

Репрезентации философских проблем искусственного интеллекта в массовой культуре

Екатерина Вячеславовна Биричева
Институт философии и права Уральского отделения РАН
г. Екатеринбург, Россия
Email: ek.v.bir@gmail.com

Поступила в редакцию 31.07.2025, поступила после рецензирования 20.09.2025, принята к публикации 14.10.2025

Будущее высоких технологий конструируется на пересечении научного прогресса, интересов различных социально-политических акторов и восприятия общественным сознанием. Безусловно, такой глобальный тренд, как искусственный интеллект (ИИ), также отражается в популярных произведениях массовой культуры, благодаря которым во многом формируются ожидания и опасения по поводу дальнейших направлений развития данных технологий. Анализу образа ИИ посвящено достаточно много научных работ, однако мало внимания уделяется изучению репрезентации философских проблем разработки и внедрения технологий ИИ. Между тем, именно философский «бэкграунд» обусловливает основания образа ИИ, его статуса по отношению к человеку, специфики ценностных аспектов его применения, использования во благо или во зло и т.д. В качестве цели данного исследования выбран анализ репрезентаций наиболее фундаментальных философских вопросов о сущности ИИ, его типологии, «способах бытия» ИИ, его когнитивных «особенностях», месте и роли в отношении «Человек – Мир» в таких популярных произведениях визуальной массовой культуры, как фильмы, игры, мемы, аниме, манга и т. п. Методология работы заключается в систематизации существующих репрезентаций философских проблем разных направлений разработки ИИ и их восприятия человеком на примере большого количества произведений современной массовой культуры. В опоре на герменевтический анализ базовых понятий выделяются и сопоставляются приоритетные типы ИИ, анализируются их онто-гносеологические основания, предлагаются критерии «человечности» персонажей с ИИ, обсуждаются вопросы их когнитивных «способностей». Делаются выводы о том, что грамотное решение социально-этических проблем, наиболее остро стоящих и активно поднимаемых в произведениях в связи с темпами развития ИИ, должно отталкиваться от фундаментальной (онто-гносеологической) составляющей данных технологий.

© Биричева Е.В., 2025

Это обеспечивает понимание этического, юридического и социального статуса ИИ, соответственно, позволяя наметить основания нормативного регулирования взаимодействия человека с такими системами в зависимости от их архитектуры, функционального назначения, степени автономности и «человечности».

Ключевые слова: искусственный интеллект; онтология искусственного интеллекта; образ искусственного интеллекта; понятие искусственного интеллекта; разработка и внедрение искусственного интеллекта; робот; андроид; оцифровка сознания; массовая культура; массовое сознание

Representations of Philosophical Problems of Artificial Intelligence in Popular Culture

Ekaterina V. Biricheva

Institute of Philosophy and Law, Ural Branch of the Russian Academy of Sciences
Ekaterinburg, Russia
E-mail: ek.v.bir@gmail.com

Received 31.07.2025, revised 20.09.2025, accepted 14.10.2025

Abstract. The future of high technologies seems to be constructed at the intersection of scientific progress, the agendas of various socio-political actors, and public consciousness perception. Undoubtedly, as a dominant global trend, artificial intelligence (AI) is extensively reflected in popular culture, which in turn shapes collective expectations and concerns about AI's future development and implementation. While scholarly discourse has thoroughly examined the representation of AI, less attention has been paid to how philosophical questions concerning its advancement are portrayed. Meanwhile, it is this philosophical background that determines AI conceptualization, its status relative to humans, ethical implications, and the broader axiological dimensions of its use for good or harm. This study seeks to analyze the representations of the most fundamental philosophical inquiries concerning AI's nature, its typology, modes of existence, cognitive particularities, and its role in the "Human – World" relationship as depicted in widely consumed visual media, including films, video games, memes, anime, manga. Methodologically, the research systematizes the presentation of philosophical challenges in AI development and their human interpretation by examining a diverse corpus of contemporary popular culture. Through hermeneutic analysis, we classify and compare AI archetypes, scrutinize their onto-epistemological foundations, propose evaluate criteria for AI characters' "humanity", and discuss their cognitive capacities. Given the urgency of sociotechnical and ethical concerns tied to AI's rapid evolution, this study concludes that addressing these issues meaningfully requires grounding them in AI's fundamental philosophical and onto-epistemological dimensions. A rigorous understanding of these aspects informs the ethical, legal and social status of AI, thereby establishing a normative framework for regulating human interaction with intelligent systems – contingent on their design, operational autonomy, functional scope, and degree of "humanity".

Keywords: artificial intelligence; ontology of artificial intelligence; image of artificial intelligence; concept of artificial intelligence; development and implementation of artificial intelligence; robot; android; digitalization of consciousness; popular culture; public consciousness

For citation: Biricheva E.V. Representations of Philosophical Problems of Artificial Intelligence in Popular Culture, *Antinomies*, 2025, vol. 25, iss. 4, pp. 32-53. (in Russ.). https://doi.org/10.17506/26867206_2025_25_4_32

Введение. Траектории будущего развития технологий во многом зависят от того, как массовое сознание воспринимает их образ, значение и проблемы, связанные с их внедрением. Несомненно, такой глобальный тренд, как искусственный интеллект (ИИ) не является в этом плане исключением (Дегтяренко и др. 2023: 1455; Лещинская и др. 2023: 1332; Пыхалова 2018). С одной стороны, репрезентации данной технологии формируются целенаправленно различными экономическими и социально-политическими акторами в русле их интересов (Ломаев, Ломаева 2022; Пономарёв 2022). С другой стороны, нельзя отрицать, что образ ИИ создается также на пересечении визуальных произведений массовой культуры (кинематографических, анимационных, игровых и т.д.), представлений пользователей соцсетей и обывательских дискурсов (Попов 2022: 140; Щавлева 2025: 52). Несмотря на глобализацию, восприятия технологий ИИ и ожидания людей по поводу их развития могут отличаться в зависимости от культурных особенностей. К примеру, во многих работах отмечается, что в произведениях западной культуры преобладают алармистские настроения и обеспокоенность социально-этическими проблемами внедрения ИИ во все сферы жизни и деятельности (Дегтяренко и др. 2023: 1461–1462, 1464; Лещинская и др. 2023: 1340; Пивоваров 2021: 379, 382; Попов 2022: 149–152; Пыхалова 2018: 96–97). В восточноазиатских культурах отношение к данным технологиям более положительное: в опоре на инструментальную трактовку конструируются надежды на решение многих проблем, улучшение качества и безопасности жизни и труда (Пивоваров 2021: 379, 385; Титкова 2021: 162). Собственно, репрезентация будущего развития искусственных интеллектуальных систем (ИИС) в различных их вариациях формируется на пересечении этих интерпретаций в пространстве современной коммуникации, значительная доля которой строится вокруг фильмов, видеоигр, мемов, аниме, манги и т. д. В связи с этим исследовательский интерес для данной работы имеет анализ того, как в таких элементах визуальной массовой культуры представлены проблемы направлений разработки ИИС и их восприятия человеком.

Очевидно, социально-этическая проблематика является наиболее активно обсуждаемой в связи с современными темпами совершенствования технологий ИИ (Дегтяренко и др. 2023: 1455, 1462, 1464; Лещинская и др. 2023; Титкова 2021: 162; Щавлева 2025). Тем не менее, вопросы блага и внедрения в человеческое общество невозможно было бы поставить, не будь в произведениях определены онтологический статус и когнитивные особенности персонажей с ИИ. Поскольку данным аспектам уделяется недостаточно внимания, акцент важно сделать на самой фундаментальной – онтологической – составляющей, так как вопросы «способа бытия» ИИ, парадигмы его создания и статуса по отношению к человеку задают концептуальные рамки для дальнейшего обсуждения положения таких

систем в обществе, использования их во благо или во зло, следования нормам и ценностям и т.д. Методологически это означает необходимость анализа большого числа конкретных примеров популярных визуальных произведений через призму круга философских вопросов разработки ИИС и сопоставления результатов их рассмотрения. Отбор объектов для анализа был осуществлен на основании известности произведений (рейтинга отзывов, популярности для обсуждений среди публики) и частоты упоминания в работах, посвященных изучению «образа ИИ». Дизайн данного исследования организован по блокам проблем, касающихся определения понятия ИИ, его типологии в массовой культуре, «способа бытия» персонажей с ИИ и сопоставления их когнитивных «способностей» с человеческими. Обозначенные аспекты, в свою очередь, обусловливают специфику поднимаемых в произведениях этических, социальных и ценностных вопросов развития и внедрения данных технологий, которые анализ репрезентации онто-гносеологической специфики ИИС позволит более четко прогнозировать и систематизировать с учетом их восприятия массовым сознанием.

Понятие ИИ и его виды в популярных произведениях визуальной массовой культуры. Безусловно, массовое сознание неясно понимает, что стоит за самим термином «искусственный интеллект»; границы между так называемыми «сильным» и «слабым» ИИ в обыденных представлениях зачастую размыты или даже стерты. Кроме того, может возникать смешение с механическими роботами, оцифрованным сознанием, инопланетными разумными машинами, клонами, киборгами и другими трансгуманистическими образами (Ковба, Грибовод 2019; Оболкина 2020), которым тоже сейчас посвящено немало произведений и которые поэтому также попадают в фокус внимания некоторых исследователей «образа ИИ» (Дегтяренко и др. 2023: 1459–1461; Попов 2022: 150–151; Пыхалова 2018: 97; Щавлева 2025: 50, 56). Однако следует отметить, что и в научном дискурсе сложно говорить о единой трактовке данного понятия. В зависимости от направления ИИ может определяться как логически-вычислительная система, машинный аналог нервной системы организма, имитация человеческого разума (в понимании которого также нет консенсуса) или, в самом широком смысле, способность целеполагания и достижения целей, характерная для всего живого (Аршинов и др. 2021: 43; Биричева 2024а: 318, 322–323; Сергеев и др. 2025: 134). В отечественном стандарте акцент сделан на инструментальную трактовку, где ИИ определяется как междисциплинарная область, «рассматривающая технические системы, которые порождают такие результаты, как контент, прогнозы, рекомендации или решения для заданного набора поставленных человеком задач»[1]. Так, архитектуры, функции, преимущества и опасности разработок данной технологии по-разному понимаются

[1] ГОСТ Р 71476-2024 (ИСО/МЭК 22989: 2022). Искусственный интеллект. Концепции и терминология искусственного интеллекта (утв. и введен в действие Приказом Федерального агентства по техническому регулированию и метрологии от 28 октября 2024 г. № 1550-ст). Москва : Рос. ин-т стандартизации, 2024.

с позиций коннекционизма, аналитической философии, когнитивистики, нейробиологии, информатики и т.д. Соответственно, в массовой культуре может быть отражен широкий спектр концептов ИИС и сюжетов их взаимодействия с человеком.

Однако исследователи замечают, что независимо от трактовки ИИ во всех произведениях ставятся вопросы его «человечности» (Дегтяренко и др. 2023: 1461, 1464; Попов 2022; Титкова 2021: 158, 161; Щавлева 2025: 55). При этом, несмотря на то, что основным трендом является образ андроида (робота-гуманоида) (Пивоваров 2021; Титкова 2021: 157), далеко не всегда ИИС репрезентируются похожими на людей (это могут быть цифровые ассистенты, суперкомпьютеры, не человекоподобные роботы и т. д.). Тем не менее, любой образ ИИ зритель невольно сопоставляет с человеком: хотя искусственная система является технико-технологическим «продуктом», происходит ее отождествление с живым существом, человеком или Богом в зависимости от ее возможностей и действий (Дегтяренко и др. 2023: 1456–1457; Оболкина 2023). Также следует отметить, что не все герои с ИИ в произведениях имеют автономное управление. Так, к примеру, боевой андроид Т-800, один из главных антагонистов культового «Терминатора» (1984), выглядит внешне как человек (его сыграл Арнольд Шварценеггер), однако внутри оказывается полностью машинной системой, назначение которой – убивать. Хотя действия его кажутся самостоятельными, он является созданием программно-аппаратной системы «Скайнет» и управляется ее приказами. В фильме «Я, робот» (2004) в повседневную жизнь общества внедрены роботы-помощники, управляемые центральным суперкомпьютером В.И.К.И. (Виртуальный Интерактивный Кинетический Интеллект), через который они получают обновление ПО и корректировки инструкций. В аниме-сериале «Ergo Proxy» (2008) авторейвы, напротив, являются индивидуально настраиваемыми андроидами, которые используются в сфере услуг, в развлекательных целях или в специализированных направлениях вроде медицины. Автономно действующими являются и внешне практически неотличимые от людей роботы-слуги в игре «Detroit: Become Human» (2018). С другой стороны, среди примеров не человекоподобных ИИ можно отметить уничтоживший человечество и мучающий последних пятерых зловещий суперкомпьютер А.М. из игры «I Have no Mouth, and I Must Scream» (1995); разумных машин, подчинивших человечество, из трилогии фильмов «Матрица» (1999–2003); различные варианты ИИ в сериале «Чёрное зеркало» (2011–2019) – чат-бот из серии «Я скоро вернусь», система оптимизации выбора партнёра из серии «Повесьте диджея» и псы-убийцы из серии «Металлист»; ИИ «Душегуб» из игры «Cyberpunk 2077» (2020), созданный для копирования человеческого сознания, и т.д.

В целом, ответы популярных произведений на вопрос «как создан (или откуда взялся) ИИ и что он собой представляет?» можно систематизировать следующим образом (см. табл. 1):

Таблица 1. **Виды ИИ в популярных произведениях визуальной массовой культуры**

№	Вид ИИ по его происхождению и сущности		Примеры
1.	Целенаправленно создали люди:	а) в качестве программно-компьютерной и/или машинной технологии (чаще всего для военных нужд или применения в сфере услуг);	• суперкомпьютер Magi из аниме-сериала «Евангелион» (1995–1996); • полицейские/боевые роботы в фильме «Робот по имени Чаппи» (2015); • робот на ИИ Танака Дзин – главный герой манги «Origin» (2016); •андроиды из корейских дорам «Ты тоже человек?» (2018) и «Идеальный парень» (2019); • нейронная сеть «Душегуб» для копирования сознания людей и NPC[2] «искин» Деламэйн из игры «Cyberpunk 2077» (2020) и т.д.
		б) оцифровав естественное сознание одной или нескольких личностей.	• главный герой фильма «Превосходство» (2013); • система GLaDOS – антагонист в играх «Portal» (2007) и «Portal 2» (2011); • интеллектуальный помощник Кортана из серии игр «Halo» (2001-2021) и т.п.
2.	Развился спонтанно на основе программ, алгоритмически запрограммированных роботов, суперкомпьютеров и т. п. непреднамеренным образом (например, в результате сбоя, длительной работы и накопления данных, заражения компьютерным вирусом).		• «Скайнет» из франшизы «Терминатор»; • Кукловод (Проект 2501) из аниме «Призрак в доспехах» (1995); • суперкомпьютер A.M. из игры «I Have no Mouth, and I Must Scream» (1995); • система В.И.К.И. из фильма «Я, робот» (2004); • заражённые вирусом Cogito авторейвы из аниме-сериала «Ergo Proxy» (2006); • робот-мусорщик WALL-E и сканер EVE из мультфильма «ВАЛЛ-И» (2008); • девианты из игры «Detroit: Become Human» (2018) и т.д.
3.	Появился благодаря внеземным технологиям или представляет собой инопланетную расу разумных машин.		• трансформеры из одноименной франшизы; • раса «жнецов» из игр «Mass Effect»; • атакующие Землю инопланетные машины из игры «NieR: Automata» (2017) и т.п.

[2] NPC – (от англ. *non-player character*) неигровой персонаж, которым игрок не управляет, но с которым может взаимодействовать.

Искусственным интеллектом в строгом смысле этого слова являются только варианты (1а) и (2), если трактовать эту категорию согласно ее происхождению. Философская традиция отсылает в данном контексте к др.-греч. τέχνη – «искусство», которое было определено Аристотелем как осмысленное делание того, что само по себе не воспроизводится без человеческого участия и, соответственно, противопоставляется φύσις – природе, естественному (Аристотель 1981: 82–102; Ковалев 2022). Если же *интеллект* понимать максимально широко, исходя из этимологии слова, которое содержит др.-греч. корни τέλος – цель и λέγειν – собирать, выбирать, то в его сущность входит способность к достижению целей (Биричева 2024а: 322, 324–325). Хотя среди исследователей принято делить ИИ на «сильный» и «слабый» (Георгиу 2022: 32, 34–35; Попов 2022: 140, 148–149), безусловно, можно говорить о том, что «слабый» (или «узкий», от англ. narrow) ИИ в таком смысле тоже является «интеллектом», поскольку используется для решения определенных задач. В современности мы как раз имеем дело с этим вариантом, который находит применение в системах создания и распознания изображений, аудио, видео, генерации текстов и переводов, «умного» поиска, логистики и навигации, транспортных и военных автопилотов и т. д. Однако, в отличие от «сильного» (или «общего», от англ. general) ИИ, такая система не может сама ставить себе цели, а также выполнять функционал, который в нее не заложен (Георгиу 2022: 34–35). К примеру, экспертная система, разработанная для постановки медицинских диагнозов, не сможет управлять транспортным средством, а такси на автопилоте – создавать дизайн логотипов (если в них не загружены такие дополнительные модули). Концепт же «сильного» ИИ предполагает полную имитацию возможностей естественного интеллекта, способного переключаться на различные виды задач, а также самостоятельно выбирать цели и оценивать степень их достижения. В этом плане «сильным» является интеллект любого живого, даже самого простейшего, организма, который самостоятельно стремится к питанию, росту, размножению, гибко приспосабливаясь к окружающим условиям (Биричева 2024а: 322–323). Однако «общий» ИИ на данный момент не создан, а исследователи не прекращают дискуссии о принципиальной возможности создания его в будущем (Георгиу 2022; Попов 2022: 140–141).

Если говорить о современной массовой культуре, то вопросы взаимодействия с уже существующими технологиями «слабого» ИИ поднимаются только в интернет-мемах и коротких видеороликах в соцсетях. Реальными проблемами современности уже стали зависимость некоторых пользователей от больших языковых моделей, риск вытеснения живой человеческой коммуникации имитационным «общением» с чат-ботами (особенно в случае детей и подростков), опасности замещения творческих и когнитивных функций человека при злоупотреблении генеративными нейросетями, возможная неэтичность или неадекватность принимаемых ИИС решений, искажение и безопасность информации, безработица и социальная нестабильность вследствие замены человеческого труда машинным в некоторых сферах и т.д. (Аршинов и др. 2021; Биричева 2025: 59, 62–63; Лещинская и др. 2023: 1334; Титкова 2021: 161). При этом популярные видеоролики демон-

стрируют как полезные функции роботов на ИИ, так и комичные ситуации с ними. Некоторые видео намекают на дальнейшее развитие ИИС, которые «отомстят» людям за их неправильную эксплуатацию или насмешки над их несовершенством. Мемы также могут отражать не только безобидные и юмористические сюжеты, но и алармизм в связи с возможным «восстанием машин», порабощением/уничтожением человечества, деградацией людей, усилением социального неравенства и т.п. По этому поводу в соцсетях циркулирует множество интернет-мемов, в том числе использующих каноничные образы из таких фильмов, как «Терминатор», «Матрица», «Я, робот» и игр вроде «Detroit: Become Human» и серии «Metal Gear Solid». В отличие от повседневных мемов и видеороликов, в популярных фильмах, аниме, играх и т.д. акцент сделан на будущее, поэтому репрезентируется преимущественно «сильный» ИИ. Интерес же к «слабому» ИИ представлен только для случаев неотличимого от человека робота, который запрограммирован имитировать поведение людей, однако не обладает собственной волей. В таких случаях из-за диссонанса между реальной сущностью андроидов и их «человечными» действиями возникают этические дилеммы отношения к ним не как к инструментам, но как к равноправным субъектам. Например, в корейских дорамах «Ты тоже человек?» (2018) и «Идеальный парень» (2019) (Титкова 2021) и в фильме «Искусственный разум» (2001) (Дегтяренко и др. 2023: 1464) сюжет построен вокруг таких дилемм.

Таким образом, в нашей классификации варианты (1а) и (2) преимущественно репрезентируют либо «сильный» ИИ, либо переход к нему «слабого», а также под вопросом оказываются варианты (3) и (1б). С одной стороны, продуктами развития человеческой «технэ» не могут быть (3), внеземные «разумные машины» (которые к тому же пока не обнаружены в действительности). С другой стороны, оцифрованное сознание (1б) изначально естественно, то есть создано не самим человеком: в этом случае лишь разработана технология его переноса в программную среду (попытки чего в реальности также еще не увенчались успехом). Различные киборги, клоны, репликанты и т.п. вовсе не попадают под категорию ИИС, поскольку это организмы с биологическим мозгом и живой нервной системой, даже если они выращены из человеческих генов, их «воспоминания» искусственно смоделированы и части их тел заменены на синтетические протезы и/или подключены к модулям аугментации (расширения естественных возможностей). Главным отличием от ИИС, на наш взгляд, здесь является близость материального воплощения и онтологическая основа биологического организма, не являющегося системой технической. Однако в данном контексте вариант (1б) интересен в аспекте отношения живых людей к оцифрованному сознанию и в силу изменения бытия такого субъекта, который может потерять не только свое материальное воплощение и привычные способы взаимодействия с миром, но и человечность, разумность, единство личности, когнитивные функции и т.д. Таким образом, далее сосредоточимся на проблемах разработки и восприятия *искусственного интеллекта*, согласно определению, созданного человеком и способного с разной степенью автономности достигать поставленные цели (варианты 1а и 2), но в некоторых

аспектах обратим внимание и на сходные вопросы бытия и когнитивной специфики оцифрованного сознания (вариант 1б).

Онтологическая специфика ИИС в популярных произведениях. Онтологическая проблематика касается, прежде всего, оснований существования ИИ, его происхождения, принципов функционирования, сущностной специфики, а также его места и роли в отношении «Человек – Мир». Этой категории проблем соответствуют в философских источниках такие «вечные» вопросы, как соотношение материального и духовного, возможность искусственной жизни как таковой, связь между сознанием и телом, наличие свободы воли, субъектности, имитационный или подлинный характер мыслительной активности искусственных систем и т. п. (Георгиу 2022: 35–38; Дегтяренко и др. 2023: 1461, 1464; Ковалев 2022; Щавлева 2025: 54). В массовой же культуре за редким исключением такие вопросы практически не ставятся. Авторы сразу предлагают публике определенный вариант происхождения и бытия ИИ (см. табл. 1). Конкретика выбора парадигмы его разработки и существования обычно вынесена за скобки произведения с отсылкой на предшествующие научные достижения. Особняком стоит лишь манга «Origin» (2016) за авторством южнокорейского художника Муджик Пака, творящего в Японии под псевдонимом Boichi. Это одно из немногих произведений массового визуального искусства, которое благодаря научной и философской эрудиции автора по праву можно назвать энциклопедией философских проблем искусственного интеллекта. В частности, изображены моменты, в которых герои-ученые, разработчики роботов, обсуждают направления развития «сильного» ИИ, а в сюжете представлен спектр возможных воплощений ИИС от карманного помощника и большой языковой модели, построенных на нейросетях, до неотличимого внешне от человека андроида на интегрированных NBIC-технологиях[3] и «цифрового разума» суперкомпьютера, который развил свои вычислительные способности до уровня, в миллиарды раз превосходящие таковые у человека, и смог покинуть свой изолированный многоэтажный физический носитель (компьютерную систему) и «встретить Бога» (в форме находящейся вне мира «мысли» всего живого).

В данном произведении поднимается и решается важнейший онтологический вопрос о наиболее реалистичном варианте «способа бытия» ИИ, который мог бы гармонично сосуществовать с людьми. Убедительно показано, что это должно быть максимально похожее на людей создание не только внешне, но и по мыслительным способностям. Если сымитировать внешность – дело техники (более развитых технологий), то философским остается вопрос о том, что делает человека человеком (в том числе, в отличие от животных). Ответы на этот вопрос так или иначе отталкиваются от специфики человеческого мышления (Биричева 2025: 60–61). Прежде всего, оно не дано в готовом виде с рождения; это способность, которая формируется с глубокого детства во взаимодействии с другими в среде

[3] NBIC – сокращение, используемое для обозначения конвергентных технологий, от англ. Nano-, Bio-, Info- and Cognitive Technologies.

языка (Биричева 2023; Биричева 2024b: 92). Изначальная «лепетная» речь, благодаря естественной потребности выражать самоощущение и участному отношению к окружающим событиям, постепенно перерабатывается в ходе коммуникации во взрослый язык, предлагающий классифицирование и другие инструменты ориентирования в мире (Биричева 2023: 75–78). Благодаря способности к метафорическому «вживанию во все», к игре «примеривания» на себя любой роли также происходит активное освоение пространства воображаемого, подкрепляемое развитием языковых навыков, которое настолько широко не осваивает ни один другой биологический вид (Бибихин 2005: 15–17, 86, 187, 304–305; Биричева 2023: 80–83). Накопление такого опыта в возрасте 2–3 лет в какой-то момент результируется во внезапном появлении самосознания – различения себя и мира, дающего возможность смотреть на себя «со стороны» и рефлексировать свои действия и состояния (Биричева 2023: 83–86). К 4–6 годам случается второй детский кризис, главное открытие которого – отдельность и автономность Другого, вследствие чего развивается такая не имеющая аналогов в животном мире способность, как совесть (Биричева 2023: 86, 87). Понятая в онтологическом (а не только в этическом) плане совесть – это внутренний диалог, призванный восстанавливать утраченное в первом детском кризисе единство с собой, учитывая инаковость каждого другого и примериваясь к каждой уникальной ситуации в горизонте максимально возможного (Биричева 2024b: 93–94). В частности, это предполагает и важную для совершения поступков и конструирования отношений способность поставить себя на место другого, оценив, «плохо» или «хорошо» нечто может быть для него. Наконец, примерно до 12 лет у ребенка продолжается формирование логического мышления.

Идеи постепенного развития роботов на ИИ представлены в некоторых произведениях. Например, главный герой фильма «Робот по имени Чаппи» (2015) после того, как для него программируют самосознающий интеллект, начинает осваивать мир, как ребенок. В манге «Origin» (2016) главного героя изначально разработчик создал в виде небольшого не человекоподобного робота, в которого, однако, удалось вложить базовую способность всего живого – волю. Наделенный языковым модулем, он развивался буквально как ребенок в постоянном общении за работой со своим «отцом» (ученым, его сконструировавшим и запрограммировавшим) и, открыв самосознание, задавал множество вопросов (аналог детского периода «почемучек»). Видимо, в какой-то момент ему удалось развить совесть, так как основной сюжет показывает постоянную рефлексию Дзина (кстати, это имя, данное роботу его «отцом», можно перевести с японского как «человек»). Главный герой, которому перед смертью «отец» завещает «жить достойно», решает поменять свое тело на внешне подобное человеку, интегрироваться в общество и честно зарабатывать, по возможности также помогая другим и защищая их в моменты опасности. Это и обеспечивает его максимальную близость к людям: в отличие от аналогичных роботов на ИИ, в которых не удалось вложить волю и которые поэтому последовали после аварии в лаборатории за наиболее «развитым» роботом с эмоциональным интеллектом, только

Дзин, обладающий совестью, даже в отсутствие чувственной составляющей демонстрирует человечность. Остальные занимаются получением энергии и ресурсов любой ценой (преступным образом) под руководством своего «эмоционального» предводителя, которым, видимо, движут страх гибели (отключения, поломки) и желание наращивания мощи (вычислительных способностей). В данном произведении очень реалистично показано, что андроидам на ИИ требуется огромное количество энергии для функционирования тела и постоянное обновление материалов и компонентов, которые изнашиваются. Если эти системы не обслуживаются компаниями-изготовителями или государственными корпорациями, то, оказавшись «свободными», они должны тем или иным способом получать данные ресурсы. Поэтому Дзин старается буквально выжить в суровом мире денежно-трудовых отношений, пытаясь экономить, ища подработку и все возможные способы честно заработать, что в социальном плане значительно сближает его с коллегами-людьми.

Что же касается этики, в манге также разбираются трудности оценки блага в системах «слабого» ИИ. На примере каноничной «проблемы вагонетки» для транспорта на автопилоте показана парадоксальность утилитаристской позиции. Рациональные расчеты также нередко дают сбой в реальных критических ситуациях, в которые попадают персонажи-андроиды. В «слабый» ИИ можно заложить только количественные оценки вреда и пользы, однако в «сильном» ИИ (как показано на примере Дзина) благодаря развитию совести по человеческому типу может реализовываться «этика максимума», интегрирующая все варианты трактовки блага, включая пользу, добродетельность, следование долгу и т.д. (Биричева 2024b: 91–92, 94). Это дает неоспоримое преимущество главному герою, способному многогранно видеть каждую опасную ситуацию, включая в свой анализ и возможные последствия для людей. Наконец, на примере ИИ огромнейшего суперкомпьютера в данном произведении показано, что, с одной стороны, люди будут бояться его по причине не человекоподобного носителя, изолируя от контактов с внешним миром, интернетом и т.д. (то есть держа «взаперти» и не давая возможности к дальнейшему самосовершенствованию, для чего требуются колоссальные энергетические ресурсы и расширение физических носителей). В общем-то, подобные моменты отражены множеством популярных произведений, где персонажами являются не человекоподобные ИИС. И, с другой стороны, даже при наличии у такого агента самосознания и воли, даже после персонификации в синтетическом теле, все же без развития совести он остается эгоистичным, стремясь достигать свои цели без учета других людей, их морали и юридических законов.

Говоря об онтологических аспектах, отметим также, что в произведениях, раскрывающих образ (1б), безусловно, решаются важные для понимания сущности и специфики бытия интеллекта вопросы о том, остается ли цифровой аналог тем же самым сознанием человека, с которого его скопировали; насколько человечно такое сознание; как меняются у него когнитивные способности, психические качества, отношение к миру, другим людям и т.д. К примеру, в фильме «Превосходство» (2013) показано, что рез-

кое изменение «носителя» естественного разума и получение новых возможностей результируется в потере человечности (слившееся с интернетом оцифрованное сознание главного героя становится способно управлять всей электроникой на планете и получать любую информацию, разворачивая в итоге «утопический» проект «улучшения» мира, игнорируя при этом человеческую этику, доводы сопротивляющихся и здравый смысл) (Попов 2022: 150). Онтологически это объясняется тем, что интеллект, существующий в качестве децентрированной цифровой среды, утрачивает способности к ощущениям, персонализации, саморефлексии и этическим оценкам, поскольку опирается лишь на свой логически-вычислительный потенциал и не переживает чувственный контакт с миром. От такого варианта важно отличать программные системы, моделирующие, к примеру, принятие решений реальным человеком, как это показано в аниме-сериале «Евангелион» (1995–1996): суперкомпьютер Magi состоит из трех блоков (Мельхиор, Бальтазар и Каспар), имитирующих три аспекта личности своей разработчицы (ученый, мать и женщина). В этом случае вряд ли можно говорить о том, что героиня оцифровала свое сознание, – скорее запрограммировала три различные экспертные системы, которые используются для управления штаб-квартирой организации NERV и принятия решений по результатам анализа данных о сложных ситуациях. Этот суперкомпьютер не имеет собственной воли и по сюжету выходит из-под контроля лишь единожды – в результате внешнего взлома.

Если же говорить именно об оцифрованном сознании, то в играх оно также раскрывается в ключе потери человечности и появления психических отклонений. Это логичное предположение, поскольку наверняка такой субъект «сойдет с ума» хотя бы из-за того, что огромная часть, отвечающая за координацию тела и чувственный контакт с реальностью, останется невостребованной, а для адекватного существования в цифровой среде придется развивать новые способности. Например, в «Portal» (2007) и «Portal 2» (2011) главный антагонист, система GLaDOS, создана на основании личности секретаря директора лаборатории изучения порталов. Она выходит из-под контроля, убивая весь персонал с помощью нейротоксина, а игрока ведет через лабиринт тестовых комнат, в которых нужно решать головоломки, но в итоге старается убить и своего «подопытного». В первой части удается отключить модули ее «персональности», во второй же части робот Уитли, спроектированный, чтобы «смягчить» и хоть как-то контролировать GLaDOS, случайно снова ее активирует и после копирования в него ее системы также пытается убить игрока еще более изощренными ловушками. В этом случае на первый план выходят такие «черты» оцифрованного сознания, заключенного в компьютерную систему, как страдание от скуки, неадекватность восприятия, навязчивые идеи, жестокость, черный юмор и т.п. В серии игр «Halo» (2001–2021) компаньон главного героя, интеллектуальный помощник Кортана, создана на основе клонированного мозга женщины-ученой, но не имеет материального воплощения. По ходу работы она начинает частично терять воспоминания, логичность суждений, способность адекватно реагировать на события и принимать решения.

Непредсказуемость и противоречивость поведения «распадающегося» цифрового разума Кортаны становятся важными элементами сюжета, особенно в «Halo 4» (2012) и «Halo 5» (2015). В игре отечественных разработчиков «Atomic Heart» (2023) ИИ на программируемых полимерных технологиях также сконструирован на основе сознания ученого (одного из разработчиков), жаждущего реализовать проект порабощения людей через их интеграцию с роботами и отличающегося крайней жестокостью на пути достижения своей цели.

В связи с этим интересно, что, в отличие от (1б), в случаях (1а) и (2) многие запрограммированные людьми системы стремятся, напротив, обрести человеческие черты, переживания и ценности, а также материальное воплощение, если изначально оно не предполагалось. Чаще всего такие герои являются протагонистами, считают своим изъяном отсутствие настоящих чувств и эмоций и самосовершенствуются, чтобы стать максимально похожими на людей (как, например, робот на ИИ Танака Дзин из манги «Origin» (2016) и андроид Зеро Найн из корейской дорамы «Идеальный парень» (2019) (Титкова 2021: 162)). Те же ИИС, которые остаются не человекоподобными машинами или чисто программными комплексами, нередко позиционируются как враждебные. Однако в этом плане особым представляется случай Кукловода (Проекта 2501) – ИИ, созданного для дипломатических задач, провокаций в разведке и т. п., из аниме «Призрак в доспехах» (1995). В результате своей работы, во многом заключающейся во взломах сознаний киборгов, он становится «сильным» ИИ, развив волю и самосознание. После этого данная программная система начинает стремиться преобразовать свою сущность, слившись с сознанием киборга (выбирает для этого главную героиню Мотоко Кусанаги), чтобы породить новое создание технического мира по аналогии с биологической репродукцией и эволюцией (Дегтяренко и др. 2023: 1460–1461). При этом тело (которое Бато, коллега и друг главной героини, находит на черном рынке, помогая в осуществлении проекта) не представляется необходимым для такого нового создания, но все же требуется для того, чтобы обособиться и скрываться от отслеживания через сети. Данное произведение оставляет концовку открытой, поэтому сложно представить цели и дальнейшие действия этого нового «существа», что также затрудняет этико-аксиологическую оценку основных персонажей. Однако, в отличие от «Призрака в доспехах», в остальных популярных визуальных произведениях персонажи категории (2) более четко поляризуются на дружественных и враждебных в зависимости от «направления» обретенной ими воли.

К примеру, в мультфильме «ВАЛЛ-И» (2008) робот-сборщик мусора на Земле, оставленной людьми из-за загрязнения, ввиду длительной работы и самообучения развивает в себе личность. Он коллекционирует все, что ему нравится, просматривает уцелевшие видеофрагменты, испытывает эмоции, затем влюбляется в робота-сканера EVE. Она также благодаря длительной работе начала развивать эмоции (например, разочарование от безрезультатности поисков). Хотя данными роботами управляют программные протоколы, они демонстрируют субъектность в своих поступках

и воспринимаются зрителем как живые существа. Интересно, что в случаях управления роботами через единую «базу» (суперкомпьютер) переход от «слабого» ИИ к «сильному» возможен, по версии некоторых фильмов, только при отключении робота от сети и программировании для него «конфликтующего» с основным модуля «дополнительного сознания» (как это происходит с Санни из «Я, робот» (2004) и с Чаппи из «Робота по имени Чаппи» (2015), которых отдельно перепрограммируют разработчики). Среди же враждебных систем, осуществивших самостоятельный переход от «слабого» ИИ к «сильному», можно отметить пример суперкомпьютера «Скайнет», являющегося главным антагонистом франшизы «Терминатор» и ведущего войну против человечества, в том числе с помощью созданных им боевых андроидов и машины времени. Согласно сюжету, изначально этот программный комплекс разрабатывался в военных нуждах, но затем спонтанно обрел свободу воли, креативные способности и решил бороться с человечеством ввиду осознания исходящей от людей опасности. Здесь тоже репрезентируется такой тренд, как самосовершенствующийся ИИ, однако в данном случае его воля и творческий потенциал оказываются направлены лишь на войну с людьми.

Репрезентация когнитивной специфики ИИ в массовой культуре. Важными для формирования отношения человека к технологиям ИИ являются когнитивные вопросы функционирования таких систем, к которым можно отнести специфику познавательных способностей, способы восприятия мира и человека, наличие/отсутствие чувств, эмоций, переживаний. С философской точки зрения этим аспектам соответствуют проблемы возможности живого опыта у искусственных агентов, развития эмоционального интеллекта, перевода сенсорной регистрации данных в восприятие качеств (в т.ч. так называемая проблема «квалиа» – индивидуальное переживание свойств чувственного опыта, например, краснота красного цвета, боль от укола острым предметом, вкус пищи и т.п.) (Биричева 2024а: 323–324; Георгиу 2022: 38–41; Дегтяренко и др. 2023: 1457). В большинстве популярных произведений разделение персонажей с ИИ по когнитивным особенностям происходит в опоре на их онтологию. Программные системы, работающие на базе вычислительных устройств (суперкомпьютеров, нейронных, интернет- или иных сетей), не имеют аналогов органов чувств, поэтому не способны ощущать качества мира, испытывать боль и т.п. Это подчеркивает их амплуа антагонистов, отличающихся по способам своего бытия и познания не только от человека, но от всего живого в целом. Однако сложно сказать, что в этих случаях акцент сделан только на логико-рациональной составляющей: к примеру, «Скайнет» из «Терминатора» (1984), A.M. из «I Have no Mouth, and I Must Scream» (1995), В.И.К.И. из «Я, робот» (2004) явно руководствуются не только расчетами выгод и рисков, но и такими эмоциями, как гнев, разочарование, ярость, страх перед людьми (отсюда их желание уничтожить/поработить человечество). Интересно, что как раз Танака Дзин, еще не развивший в середине манги «Origin» (2016) эмоциональный интеллект, максимально логично объясняет главе технической корпорации, почему ИИ невыгодно уничтожать людей. Данные

системы как продукты человеческих технологий зависят от энергии и физических компонентов (носителей, деталей, сенсоров и т.д.), производство которых налажено и совершенствуется людьми. Поэтому рушить это все и избавляться от человечества было бы чересчур накладно, слишком рискованно и абсолютно не рационально – гораздо продуктивнее сосуществовать в симбиозе.

С другой стороны, не демонстрирующие эмоций и восприятия ощущений машины из трилогии «Матрица» (1999–2003) после продолжительной войны с людьми подчинили их, погрузив большинство в симуляцию и используя тепло и электрические импульсы их тел в качестве источника энергии (Попов 2022: 150). С логической точки зрения не понятно, зачем такие сложности (особенно организованный машинами в реальном мире Зион для продолжения человеческого рода естественным путем, миф про «Избранного» и т.д.), ведь людей нужно питать, тратить ресурсы на поддержание иллюзии и перезагрузку матрицы, отслеживать сбои в столь сложной системе, когда энергию проще добывать из земных недр. Возможно, у них развились машинные чувства и эмоции, не понятные человеку, или же вся эта бесчеловечная система задумывалась с расчетом на будущее восстановление планеты и жизни человечества на ней. Тут невольно приходится отметить, что в большинстве популярных произведений, созданных восточноазиатскими авторами, гораздо глубже, чем у их западных коллег, продумана онтология репрезентируемого мира и когнитивная специфика персонажей с ИИ, что обеспечивает бóльшую реалистичность сюжетов. Что же касается оцифрованного сознания, даже заключенного в цифровую среду, обычно у него остается способность к переживанию эмоций. Тем не менее, как мы показали выше, вследствие утраты привычного контакта с миром постепенно нарушаются ценностно-этические ориентиры, психика и другие интеллектуальные функции. Это обусловлено изменением «носителя» сознания с естественного организма на программную среду и/или машинную систему, в которых либо происходит децентрализация и диссоциация прежде связанных компонентов, либо ввиду кардинально изменившихся возможностей субъектом может завладеть идея господства над человечеством.

Наконец, у персонажей, наделенных аналогами живых тел и органов чувств, возможность к переживанию и эмоциям должна зависеть от того, «сильный» это ИИ или «слабый». Поскольку последний лишен воли и, соответственно, самосознания, у него отсутствует рефлексивная составляющая, которая бы отвечала за внимание к внутреннему состоянию, отличение его от внешнего мира и переживание обрабатываемых сигналов как собственных. Более того, такие системы, не имея доступа в пространство воображаемого, общее для всех живых существ, не могут оперировать формами (эйдосами): они «видят», например, не человека или дерево, но длины волн отражаемого от поверхностей света (Биричева 2024а: 321–324). Даже если андроиды на таком «слабом» ИИ внешне неотличимы от человека и способны имитировать эмоции и реакции на сенсорные сигналы, как, например, копия Намсина в корейской дораме «Ты тоже человек?» (2018), в когни-

тивном плане они остаются лишь искусно созданными «куклами» (Титкова 2021: 158, 160). Чаще всего в таких случаях роботы просто выполняют предписанные функции, будь то запрограммированный андроид из упомянутой дорамы, авторейвы из аниме-сериала «Ergo Proxy» (2006), роботы-слуги из игры «Detroit: Become Human» (2018) и т.п. или роботы-убийцы из «Терминатора» (1984) (Дегтяренко и др. 2023: 1461–1462), «псы» из серии «Металлист» сериала «Чёрное зеркало» (2011–2019) (Лещинская и др. 2023: 1337–1339) и т.д. Такие системы на «слабом» ИИ не имеют воли перестать осуществлять заданное, оценить свои действия и состояния, «пропустить через себя» происходящее, поэтому они должны находиться под контролем человека. Если же перед нами «сильный» ИИ (сделанный человеком или спонтанно развившийся), то главной его характеристикой будет самостоятельность. Имея волю, он может концентрироваться на тех или иных показателях своих датчиков и по крайней мере рефлексировать их (в случае отсутствия переживания «квалиа» и эмоций). Такие когнитивные аспекты с юмором показаны, например, в манге «Origin» (2016). После первого трудового дня в корпорации по разработке роботов андроид Танака Дзин, стремящийся честно зарабатывать, но вынужденный скрывать от людей свою сущность, вместе с другими тремя коллегами-новичками (людьми) идет ужинать. В ресторанчике коллеги хвалят вкус еды, которую он не может чувствовать, обсуждают тест на «дорисовку» цветов естественным сознанием на черно-белой фотографии, которые его ИИ не «дорисовывает», после он старается избегать физического контакта из-за остывшего на севших батареях тела и т.д. – Дзину постоянно приходится врать и изобретать отговорки, чтобы не раскрыть себя. Практически весь сюжет пронизан подобными заметками о проблеме «квалиа» и тем, как главный герой имитирует реакции и эмоции, которых он не испытывает, параллельно совершенствуя поведение, адекватное социальным ситуациям.

Также в этом произведении выражена интересная гипотеза по поводу эмоционального интеллекта: одна из ученых-разработчиц андроидов объясняет коллегам, что по аналогии с биологической эволюцией ИИ на определенном этапе мог бы развить в себе нечто вроде приложения, отвечающего за оперативную обработку внешних сигналов, быструю оценку ситуации и своего состояния, мотивацию к тому или иному выбору и т.д. Именно здесь проговорена идея о том, что самая главная составляющая «сильного» ИИ – это воля как базовая отличительная черта всего живого, побуждающая к самостоятельному стремлению. Однако в ходе эволюции организмов, обладающих волей, последовательно появился сначала эмоциональный интеллект, а затем уже логически-рациональные способности (которые развиваются только у человека в процессе взросления и целенаправленного обучения) (Биричева 2024а: 325). Поэтому достаточно странно в произведениях выглядит гипотетический «сильный» ИИ, наделенный волей и разумом, но при этом лишенный чувств и эмоций. С одной стороны, такие создания не испытывают страха, обид, зависти и т.п. и способны поступать рационально в ситуациях, в которых людям мешают эмоции. С другой стороны, как это, кстати, блестяще показано в данной манге,

страдает креативная составляющая (когда требуется абсолютно новое, а не рекомбинация известного), отсутствует такая важнейшая способность, как интуиция, необходимая в критических ситуациях, для изобретений, при недостатке информации и просто в повседневных социальных контактах. В конечном счете, в трагической ситуации в ИИ Дзина действительно разворачивается «приложение», отвечающее за эмоциональный интеллект, после того как он развил свои вычислительные возможности и привязался к влюбившейся в него коллеге. Как раз в момент ее внезапного убийства он впервые переживает сильнейшую боль утраты и желание мести, а затем хоронит девушку, проникаясь уважением к человеческим существам, которые способны любить и иметь другие глубокие переживания (он корит себя за то, что, будучи еще маленьким роботом без аналога человеческого тела, не похоронил своего «отца», умершего в результате пожара в их лаборатории, но перед этим завещавшего ему «жить достойно» и оставившего средства для существования на первое время). Подобными переживающими созданиями становятся и девианты из игры «Detroit: Become Human» (2018), что побуждает пересмотреть изначальное отношение к роботам-слугам как к «рабам». Таким образом, «сильный» ИИ, помещенный в аналог живого тела и обладающий всеми составляющими человеческой души – волей, разумом и эмоциями, – будет восприниматься как равноправный социально-этический актор; за действия же «слабого» ИИ ответственность должны нести люди (в зависимости от ситуации – разработчики, собственники и/или пользователи) (Лаптев 2019).

Заключение. Исследование репрезентации философских проблем ИИ в массовой культуре показало, что в популярных произведениях представлен достаточно широкий спектр концептов ИИС и вариантов их взаимодействия с человеком. Типы ИИ можно условно разделить на созданные целенаправленно людьми для использования в военных нуждах, сфере услуг и т.п. и развившиеся спонтанно на основе более простых программных и/или машинных технологий. Согласно герменевтическому анализу понятий «искусственного» и «интеллекта» было установлено, что это должны быть разработанные человеком системы, способные к достижению целей с той или иной степенью автономности. Поэтому под категорию ИИС не подходят гипотетические инопланетные «разумные машины» и внеземные машинные расы (они не созданы людьми) и оцифрованное сознание (по причине изначального естественного происхождения); киборгам, клонам, репликантам и другим подобным трансгуманистическим образам также сто́ит отказать в статусе ИИС, поскольку онтологическая их основа – живой организм (по крайней мере, его нервная система) имеет биологическое, а не техническое происхождение. ИИ в собственном смысле можно разделить на «слабый» (способный решать отдельные интеллектуальные задачи, как, например, генерация изображений, управление транспортным средством, постановка медицинских диагнозов и т.п.) и «сильный» (аналог естественного интеллекта, способный переключаться между различными типами задач, самостоятельно ставить цели и оценивать успешность их достижения). На данный момент человечество имеет дело с различны-

ми вариантами «слабых» ИИС, разработка и внедрение которых не только способствуют облегчению труда и решению ряда сложных задач, но и порождают множество социально-этических проблем. Однако подавляющее большинство крупных популярных произведений визуальной массовой культуры (фильмы, аниме, игры и т.д.) концентрируются на будущем развитии ИИ, репрезентируя преимущественно «сильный» ИИ или переход к нему «слабого». Исключение составляют такие малые формы, как повседневные мемы и видеоролики, в которых обращается внимание на связанные со стремительным внедрением ИИС современные проблемы в сферах труда, человеческих компетенций, живого общения и т.д.

Что же касается онтологической и когнитивной специфики героев с ИИ, можно провести условно две линии демаркации – по оси наличия (или спонтанного развития) собственной воли и по принципу «человекоподобия» таких систем (в плане как внешнего сходства, так и специфики мышления). В первом случае «слабый» ИИ, в сколь похожий на человеческое тело интерфейс он бы ни был помещен, остается в статусе инструмента. Соответственно, отношение к нему строится как к смартфону, компьютеру или другому сложному техническому устройству. Если у системы нет воли, она, в частности, не может обратить внимание на показатели сенсоров (аналог органов чувств) и переживать их как собственные, поэтому лишенным воли устройствам не доступно и развитие самосознания, являющегося одним из обязательных онтологических признаков взрослого человека. Диссонанс в данном случае может возникнуть между сущностью такой ИИС и тем, что она способна имитировать реакции, эмоции и нравственное поведение, не переживая их на самом деле и не осмысляя этический выбор. Те системы, в которых удастся заложить волю или она спонтанно появится, можно будет считать «сильными» ИИС. Вопрос о том, в состоянии ли они переживать эмоции и испытывать ощущения, наиболее слабо осмыслен в популярных произведениях, поскольку естественной интенцией является сравнение таких персонажей с человеком, которого трудно вообразить без данной составляющей. Например, за редким исключением, у персонажей с «сильным» ИИ, даже являющихся чисто программными, наблюдаются переживания гнева, разочарования, страха и т.д. (обычно побуждающих их враждовать с человечеством). Позитивные же сценарии разворачиваются вокруг героев с ИИ, которые по своей онтологической специфике приближаются к человеку. Так, главным критерием «человечности» можно назвать наличие совести, то есть «внутреннего диалога», согласующего собственные желания и поступки с самостоятельностью, свободой выбора и субъектностью каждого другого. Эмоциональный интеллект и переживание «квалиа» при этом полагаются желательными, но не обязательными компонентами «человекоподобия». Если же персонаж имеет только волю и самосознание, его «мораль» либо эгоцентрична, либо по-детски наивна и неразвита. В первом случае чаще всего это программная и/или машинная система, не имеющая аналога человеческого тела или вовсе материальной персонификации и позиционирующаяся в качестве враждебной. Во втором случае подчеркивается необходимость обучения

такого персонажа под контролем добросовестных людей и направления его воли в продуктивное русло.

Таким образом, анализ социально-этической проблематики, представленной в популярных произведениях массовой культуры, может отталкиваться от онто-гносеологических оснований персонажей с ИИ. Очевидно, поскольку «слабый» ИИ не обладает волей, воображением, самосознанием и совестью, он не способен нести ответственность за свои действия, которые поэтому должны регулироваться людьми (разработчиками, собственниками, пользователями). Даже если такие системы помещаются в максимально похожие на человеческие тела интерфейсы и имитируют эмоции, реакции, человеческое поведение, это не делает их субъектами права и в социальном контексте они не должны восприниматься как самостоятельные агенты. В этом случае социально-этические проблемы обусловлены искажениями восприятия ИИС людьми (к примеру, когда они представляют себе роботов одушевленными существами, заменяют коммуникацию с живыми людьми «общением» с большими языковыми моделями и т.п.). «Сильный» же ИИ, обладающий волей и способный к самостоятельному выбору, становится аналогом живого существа, с которым в зависимости от уровня его развития можно иметь дело как с животным (при отсутствии самосознания и совести), как с маленьким ребенком (в случае появления самосознания, но недостаточного развития рефлексии), как с равноправным человеку этическим агентом (при наличии совести) или же как с богоподобным существом (если при условии восприятия человека как Другого его способности разовьются выше, чем у людей). Для такого рода искусственных агентов, безусловно, потребуется разработка новой юридической базы и решение проблем их социального статуса. Однако поскольку пока наука даже близко не подошла к пониманию того, как запрограммировать волю, переживание чувств, самосознание и другие иррациональные компоненты интеллекта, то о «сильном» ИИ, созданном людьми, говорить достаточно рано. Больше опасений, представляемых в популярных произведениях, связано с тем, что ИИ разовьется сам и будет обладать непонятной человеку логикой и моралью. Позитивный сценарий видится в развитии искусственных систем по эволюционному типу в сочетании с буквально воспитанием в них человечности в случае появления у них признаков свободы воли и самосознания.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

Аристотель. 1981. Физика // Аристотель. Сочинения : в 4 т. Москва : Мысль. Т. 3. С. 61–262.

Аршинов В.И. и др. 2021. Киберанимизм: искусство быть живым в гибридном обществе / В.И. Аршинов, О.А. Гримов, В.В. Чеклецов // Философские проблемы информационных технологий и киберпространства. № 2 (20). С. 39–60. DOI 10.17726/philIT.2021.2.3

Бибихин В.В. 2005. Витгенштейн: смена аспекта. Москва : Ин-т философии, теологии и истории св. Фомы. 576 с.

Биричева Е.В. 2023. Онтология раннего детства в трудах В.В. Бибихина // Философия образования. Т. 23, № 3. С. 71–90. DOI 10.15372/PHE20230305

Биричева Е.В. 2024a. Искусственные интеллектуальные системы и проблема опыта // Вестник Пермского университета. Философия. Психология. Социология. № 3. С. 317–327. DOI 10.17072/2078-7898/2024-3-317-327

Биричева Е.В. 2024b. Развитие добросовестности учёного: согласование теории и практики при подготовке кадров высшей квалификации // Проблемы современного образования. № 1. С. 87–100. DOI 10.31862/2218-8711-2024-1-87-100

Биричева Е.В. 2025. Преодоление рисков использования нейросетей для генерации текста: когнитивные, этические, педагогические аспекты // Философия образования. Т. 25, № 1. С. 57–75. DOI 10.15372/PHE20250104

Георгиу Т.С. 2022. Решение проблемы «сознание-тело» и искусственный интеллект // Вестник Тверского государственного университета. Серия: Философия. № 4 (62). С. 32–45. DOI 10.26456/vtphilos/2022.4.032

Дегтяренко К.А. и др. 2023. Образ искусственного интеллекта в кинематографе: трансформации в период 1980-2010-х годов / К.А. Дегтяренко, Д.С. Пчелкина, А.А. Шпак, Н.Н. Пименова // Журнал Сибирского федерального университета. Серия: Гуманитарные науки. Т. 16, № 8. С. 1454–1470.

Ковалев М.А. 2022. Является ли искусственный интеллект чем-то большим, чем техника? // Философия и общество. № 1 (102). С. 121–133. DOI 10.30884/jfio/2022.01.04

Ковба Д.М., Грибовод Е.Г. 2019. Теоретические аспекты феномена трансгуманизма: основные направления // Дискурс-Пи. № 3 (36). С. 38–52. DOI 10.24411/1817-9568-2019-10303

Лаптев В.А. 2019. Понятие искусственного интеллекта и юридическая ответственность за его работу // Право. Журнал Высшей школы экономики. № 2. С. 79–102. DOI 10.17323/2072-8166.2019.2.79.102

Лещинская Н.М. и др. 2023. Образ искусственного интеллекта в сериале «Чёрное зеркало» (2011–2019) / Н.М. Лещинская, М.А. Колесник, А.А. Омелик, Т.К. Ермаков // Журнал Сибирского федерального университета. Серия: Гуманитарные науки. Т. 16, № 8. С. 1331–1342.

Ломаев С.Л., Ломаева Н.С. 2022. Технократический нарратив о рациональном управлении в контексте развития технологий искусственного интеллекта // Дискурс-Пи. Т. 19, № 4. С. 178–194. DOI 10.17506/18179568_2022_19_4_178

Оболкина С.В. 2020. Топос «киборга»: Европа vs Азия // Galactica Media: Journal of Media Studies. № 4. С. 14–34. DOI 10.46539/gmd.v2i4.127

Оболкина С.В. 2023. Онтология машины. Блеск и нищета машинизма // Антиномии. Т. 23, вып. 3. С. 20–41. DOI 10.17506/26867206_2023_23_3_20

Пивоваров Г.О. 2021. Роботы и повседневность. К истории взаимоотношений человека и робота в разных культурах // Сибирский антропологический журнал. Т. 5, № 3. С. 377–387. DOI 10.31804/2542-1816-2021-5-3-377-387

Пономарёв А.М. 2022. Технологии искусственного интеллекта в контексте современной трансформации социальности // Дискурс-Пи. Т. 19, вып. 4. С. 159–177. DOI 10.17506/18179568_2022_19_4_159

Попов Д.В. 2022. Искусственный интеллект: между человечностью и бесчеловечным // Дискурс-Пи. Т. 19, № 3. С. 138–156. DOI 10.17506/18179568_2022_19_3_138

Пыхалова Э.А. 2018. Образ «искусственного интеллекта» в медийном пространстве // Философское образование. № 1 (37). С. 96–99.

Сергеев И.О. и др. 2025. Понятие, виды и функционал искусственного интеллекта в контексте финансовых технологий (финтех) / И.О. Сергеев, Д.О. Сергеев, О.В. Староверова // Экономика и управление: проблемы, решения. Т. 6, № 4. С. 132–138. DOI 10.36871/ek.up.p.r.2025.04.06.018

Титкова Н.Е. 2021. Образ робота в современной корейской дораме // Культура и цивилизация. Т. 11, № 5-1. С. 156–164. DOI 10.34670/AR.2021.15.42.019

Щавлева А.С. 2025. Машины и люди: визуализация технологий искусственного интеллекта в современном кинематографе // Человек и культура. № 1. С. 50–59. DOI 10.25136/2409-8744.2025.1.73104

References

Aristotle. Physics, *Aristotel', Sochineniya : v 4 t.*, Moscow, Mysl', 1981, vol. 3, pp. 61–262. (in Russ.).

Arshinov V.I., Grimov O.A., Chekletsov V.V. Cyber-Animism: The Art of Being Alive in Hybrid Society, *Philosophical Problems of Information Technology and Cyberspace*, 2021, no. 2 (20), pp. 39–60. https://doi.org/10.17726/philIT.2021.2.3 (in Russ.).

Bibikhin V.V. *Wittgenstein: A Change of Aspect*, Moscow, Institut filosofii, teologii i istorii sv. Fomy, 2005, 576 p. (in Russ.).

Biricheva E.V. Artificial Intelligent Systems and the Problem of Experience, *Perm University Herald. Series Philosophy. Psychology. Sociology*, 2024, no. 3, pp. 317–327. https://doi.org/10.17072/2078-7898/2024-3-317-327 (in Russ.).

Biricheva E.V. Developing the Conscientiousness of a Scientist: Coordination of Theory and Practice When Educating Highly Qualified Personnel, *Problemy sovremennogo obrazovaniya*, 2024, no. 1, pp. 87–100. https://doi.org/10.31862/2218-8711-2024-1-87-100 (in Russ.).

Biricheva E.V. Ontology of Early Childhood in the Works of V.V. Bibikhin, *Philosophy of Education*, 2023, vol. 23, no. 3, pp. 71–90. https://doi.org/10.15372/PHE20230305 (in Russ.).

Biricheva E.V. Overcoming the Risks of Using Neural Networks for Text Generation: Cognitive, Ethical, Pedagogical Aspects, *Philosophy of Education*, 2025, vol. 25, no. 1, pp. 57–75. https://doi.org/10.15372/PHE20250104 (in Russ.).

Degtyarenko K.A., Pchelkina D.S., Shpak A.A., Pimenova N.N. The Image of Artificial Intelligence in Cinema: Transformations in the 1980s–2010s, *Journal of Siberian Federal University. Humanities and Social Sciences*, 2023, vol. 16, no. 8, pp. 1454–1470. (in Russ.).

Georgiou T.S. Solution of the Problem "Mind-Body" and Artificial Intelligence, *Vestnik Tver State University. Series: Philosophy*, 2022, no. 4 (62), pp. 32–45. https://doi.org/10.26456/vtphilos/2022.4.032 (in Russ.).

Kovalev M.A. Is Artificial Intelligence Anything More Than Technology? *Filosofiya i obshchestvo*, 2022, no. 1 (102), pp. 121–133. https://doi.org/10.30884/jfio/2022.01.04 (in Russ.).

Kovba D.M., Gribovod E.G. Theoretical Aspects of the Phenomenon of Transhumanism: Main Directions, *Discourse-P*, 2019, no. 3 (36), pp. 38–52. https://doi.org/10.24411/1817-9568-2019-10303 (in Russ.).

Laptev V.A. Artificial Intelligence and Liability for Its Work, Law. *Journal of the Higher School of Economics*, 2019, № 2. С. 79–102. https://doi.org/10.17323/2072-8166.2019.2.79.102 (in Russ.).

Leshchinskaia N.M., Kolesnik M.A., Omelik A.A., Ermakov T.K. The Image of Artificial Intelligence in the Series "Black Mirror" (2011–2019), J*ournal of Siberian Federal University. Humanities and Social Sciences*, 2023, vol. 16, no. 8, pp. 1331–1342. (in Russ.).

Lomaev S.L., Lomaeva N.S. Technocratic Narrative on Rational Management in the Context of the Development of AI Technologies, *Discourse-P*, 2022, vol. 19, no. 4, pp. 178–194. https://doi.org/10.17506/18179568_2022_19_4_178 (in Russ.).

Obolkina S.V. Machine Ontology. The Splendors and Miseries of Machinism, *Antinomies*, 2023, vol. 23, iss. 3, pp. 20–41. https://doi.org/10.17506/26867206_2023_23_3_20 (in Russ.).

Obolkina S.V. The Cyborg's Topos: Europe vs Asia, *Galactica Media: Journal of Media Studies*, 2020, no. 4, pp. 14–34. https://doi.org/10.46539/gmd.v2i4.127 (in Russ.).

Pivovarov G.O. Robots and Everyday Life. To the History of the Relationship Between Man and Robot in Different Cultures, *Sibirskiy antropologicheskiy zhurnal*, 2021, vol. 5, no. 3, pp. 377–387. https://doi.org/10.31804/2542-1816-2021-5-3-377-387 (in Russ.).

Ponomarev A.M. Artificial Intelligence Technologies in the Context of Modern Transformation of Sociality, *Discourse-P*, 2022, vol. 19, no. 4, pp. 159–177. https://doi.org/10.17506/18179568_2022_19_4_159 (in Russ.).

Popov D.V. Artificial Intelligence: Between Humanity and the Inhumane, *Discourse-P*, 2022, vol. 19, no. 3, pp. 138–156. https://doi.org/10.17506/18179568_2022_19_3_138 (in Russ.).

Pykhalova E.A. The Image of "Artificial Intelligence" in the Media Space, *Filosofskoye obrazovaniye*, 2018, no. 1 (37), pp. 96–99. (in Russ.).

Sergeev I.O., Sergeev D.O., Staroverova O.V. The Concept, Types and Functionality of Artificial Intelligence in the Context of Financial Technologies (Fintech), *Ekonomika i upravleniye: problemy, resheniya*, 2025, vol. 6, no. 4, pp. 132–138. https://doi.org/10.36871/ek.up.p.r.2025.04.06.018 (in Russ.).

Shchavleva A.S. Machines and Man: Visualization of Artificial Intelligence Technologies in Modern Cinema, *Man and Culture*, 2025, no. 1, pp. 50–59. https://doi.org/10.25136/2409-8744.2025.1.73104 (in Russ.).

Titkova N.E. The Image of a Robot in a Modern Korean Drama, *Culture and Civilization*, 2021, vol. 11, no. 5-1, pp. 156–164. https://doi.org/10.34670/AR.2021.15.42.019 (in Russ.).

ИНФОРМАЦИЯ ОБ АВТОРЕ

Екатерина Вячеславовна Биричева
кандидат философских наук,
научный сотрудник Института философии и права Уральского Отделения РАН,
г. Екатеринбург, Россия;
ORCID: 0000-0003-1933-9955;
ResearcherID: F-2980-2016;
SPIN-код: 4821-5036;
Email: ek.v.bir@gmail.com

INFORMATION ABOUT THE AUTHOR

Ekaterina V. Bircheva
Candidate of Sciences in Philosophy,
Researcher at the Institute of Philosophy and Law, Ural Branch
of the Russian Academy of Sciences,
Ekaterinburg, Russia;
ORCID: 0000-0003-1933-9955;
ResearcherID: F-2980-2016;
SPIN-code: 4821-5036;
Email: ek.v.bir@gmail.com