

УДК 172+177

DOI: 10.17506/18179568_2022_19_3_138

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ: МЕЖДУ ЧЕЛОВЕЧНОСТЬЮ И БЕСЧЕЛОВЕЧНЫМ*



Дмитрий Владимирович Попов,

Омская академия МВД России,
Омск, Россия,
DmitriVPopov@mail.ru

Статья поступила в редакцию 23.03.2022, принята к публикации 29.08.2022

Для цитирования: Попов Д.В. Искусственный интеллект: между человечностью и бесчеловечным // Дискурс-Пи. 2022. Т. 19. № 3. С. 138–156. https://doi.org/10.17506/18179568_2022_19_3_138

Аннотация

Искусственный интеллект рассматривается в современной науке и философии не только как венец научно-технического прогресса, но и возможное последнее изобретение человечества. Он может внести в мир человека бесчеловечное, что способно уничтожить человеческую цивилизацию. Целью настоящей статьи является рассмотрение проблемы ответственности человека за бесчеловечные последствия функционирования искусственного интеллекта. В качестве источников используются исследования в области философской антропологии, политической философии, био-политики, а также современные дистопии, представленные в литературе, фильмах и телесериалах. Теоретико-методологической основой исследования выступает археология знаний М. Фуко в значении реконструкции различных дискурсивных практик с целью извлечения скрытого содержания, дополненная традиционной диалектикой. Автором статьи рассматриваются локации и формы столкновения че-

* Доклад был сделан автором на Всероссийской научной конференции с международным участием «Философские контексты современности: искусственный интеллект и интеллектуальная интуиция» (ФИКОС-2022), 25–26 февраля 2022 г., Ижевск, Россия.

© Попов Д.В., 2022



ловека с бесчеловечным. Подчеркивается, что бесчеловечное обнаруживает себя как действующая извне непреодолимая сила и побочный эффект деятельности самого человека. Люди научились специально выделять пространства для необузданных запредельных страстей. Существенный интерес в данном отношении представляют биополитические и особенно некрополитические пространства. Проведенный анализ показывает, что ответственность за вторжение бесчеловечного несет не искусственный интеллект, а человек. В этой связи возможные негативные эффекты от действий искусственного интеллекта следует относить к антропогенным проявлениям бесчеловечного. Не особый контроль, особый код или особое машинное обучение искусственного интеллекта являются гарантией от вторжения бесчеловечного в жизнь человека, а культивирование человечности в самом человеке.

Ключевые слова:

человек, искусственный интеллект, искусственный суперинтеллект, бесчеловечное, биополитика, некрополитика.

UDC 172+177

DOI: 10.17506/18179568_2022_19_3_138

ARTIFICIAL INTELLIGENCE: BETWEEN HUMANITY AND THE INHUMANE

Dmitry V. Popov,

Omsk Academy of the Ministry of Internal Affairs of the Russian Federation,
Omsk, Russia,
DmitriVPopov@mail.ru

Article received on March 23, 2022, accepted on August 29, 2022

For citation: Popov, D.V. (2022). Artificial Intelligence: Between Humanity and the Inhumane. *Discourse-P*, 19(3), 138–156. (In Russ.). https://doi.org/10.17506/18179568_2022_19_3_138

Abstract

Artificial intelligence is considered in modern science and philosophy not only as the crown of technological progress but also as a possible last invention of mankind. It can bring the inhumane into the human world that can destroy human civilization. The purpose of the article is to consider the problem of human responsibility for the inhumane consequences of the functioning of artificial intelligence. Research in the field of philosophical anthropology, political philosophy, biopolitics, as well as modern dystopias presented in the literature, movies, and television series are used as sources. The theoretical and methodological basis of the research is Michel Foucault's archeology of knowledge used in the meaning of the reconstruction of various discursive practices in order to extract

hidden content, supplemented by traditional dialectics. The article examines locations and forms of human collision with the inhumane. It is emphasized that the inhumane reveals itself as an irresistible force acting from the outside and a side effect of the activity of humans themselves. People have learned to specifically allocate spaces for unbridled exorbitant passions. Biopolitical and especially necropolitical spaces are of particular interest in this regard. It is concluded that responsibility for the invasion of the inhumane is assigned not to artificial intelligence, but to man. Therefore, possible negative effects from the actions of artificial intelligence should be attributed to anthropogenic manifestations of inhumanity. It is not special control, special code or special machine learning of artificial intelligence that is a guarantee against the invasion of inhumanity into human life, but the cultivation of humanity in humans themselves.

Keywords:

human, artificial intelligence, artificial superintelligence, inhumane, biopolitics, necropolitics.

Введение

Искусственный интеллект (ИИ) является на данный момент квинтэссенцией достижений в череде технико-технологических изобретений человечества. В связи с этим проблематика ИИ за пределами комплекса наук, изучающих технические возможности его создания и модификации, относительно нова. Вместе с тем в XXI в. появилось значительное количество работ в области правовых, экономических и философских наук, в которых осмысляются возможные последствия развития технологий сильного/слабого, дружественного/недружественного ИИ. Философия стремится спрогнозировать онтологические, социальные и экзистенциальные последствия возможного создания искусственного интеллекта человеческого уровня (ИИЧУ) и сильного ИИ или искусственного суперинтеллекта (ИСИ). Поскольку функционирование ИСИ потенциально касается каждого, значительный интерес к этой области научного знания проявляет искусство (в первую очередь литература и кинематограф). Удивление и даже эйфория от открывающейся перспективы гипотетического достижения точки технологической сингулярности, концепт которой предложен В. Винджем, постепенно и в научном, и в художественном дискурсе сменились рациональной калькуляцией выгод и издержек.

Скоротечные и радикальные трансформации человека и человеческой цивилизации, непосредственно связанные с переходом от создания и распространения технологии слабого ИИ через создание ИИЧУ к появлению технологии ИСИ, существенно превосходящего возможности человека, качественное отличие сильного/слабого, дружественного/недружественного сформировали не только оптимистическую, но и пессимистическую версию ИСИ в качестве потенциального «последнего изобретения человечества» (Баррат, 2015; Бостром, 2016; Харари, 2018). И. Дж. Гуд со сдержанным оптимизмом отмечал, что «первая ультраинтеллектуальная машина будет последним изобретением, которое придется сделать человеку, – при условии, что эта машина будет до-

статочно любезна, чтобы рассказать нам, как удержать ее под контролем... скорее всего, в двадцатом столетии такая ультраинтеллектуальная машина будет построена и окажется последним изобретением, которое придется сделать человеку» (цит. по: Виндж, 2021, с. 10). Соблазнительные перспективы высокотехнологичной «волшебной палочки» подстегивают исследователей, а возможность для человека оказаться вынесенным за скобки в собственной для ИСИ версии мироустройства актуализируют проблематику соотношения ИИ и бесчеловечного, усматриваемого в ИИ, способном погубить своего творца. Рассмотрение специфики ИИ в оптике соотношения человеческого и бесчеловечного в современной науке, литературе и кинематографе составляет цель настоящего исследования.

В качестве источников используются современные исследования в области футурологии, философской антропологии и политической философии, в частности, биополитический дискурс как область взаимопроникновения учений о человеке, обществе и политических форм его организации, а также современные дистопии, представленные в литературе, фильмах и телесериалах. Методологией исследования выступает археология знаний М. Фуко, дополненная традиционным диалектическим аппаратом. Обращение к археологии знаний связано со стремлением использовать потенциал фуколдианского подхода к реконструкции интеллектуального содержания различных срезов представлений (дискурсов) в отношении одного и того же объекта исследования. Безусловно, футурологов, уповающих на скорейшее достижение ИСИ, и писателей-визионеров, погружающих нас в мрачные пространства дистопий, в которых ИИ применяется отнюдь не во благо человеку, отличают и различные способы постановки проблематики ИИ, и различные модальности ментальности. Согласно Фуко, «археологическое описание – это как раз и есть отречение от истории идей, систематический отказ от ее постулатов и процедур, попытка создать совершенно иную историю того, что было сказано людьми» (Фуко, 2004, с. 258). Археологический анализ, по замыслу автора, дает возможность извлекать из разрозненных кластеров-архивов представлений человека знание, позволяющее продуктивнее рассмотреть заявленную проблему. Задачи археологии знаний заключаются в том, чтобы «определить дискурсы в их специфичности... проследовать за ними по их внешним граням, чтобы лучше их очертить», осуществить «дифференциальный анализ модальностей дискурса» (с. 260). В настоящем исследовании осуществлена попытка применения археологии знаний в отношении различных дискурсивных практик рассмотрения проблематики противоречивого потенциала ИИ.

Бесчеловечное: определяя понятие

Что мы понимаем под бесчеловечным? Нечто: 1) безразличное или безучастное в отношении человека; 2) непостижимое, ограниченное в познании и/или понимании человеком; 3) враждебное, чужеродное, человеконесоразмерное, несовместимое с человеком. Враждебность, непостижимость, чужеродность – критерии, интуитивно позволяющие идентифицировать бесчеловечное. Вместе с тем этого интуитивного предчувствия бесчеловечного недостаточно для определения понятия.

В качестве первого приближения предположим, что *бесчеловечное* – это нечто, что не нуждается в бытии человека для своего бытия. Однако во Вселенной множество (по сути, большинство) объектов никак не связаны с бытием человека. Автономное от человека функционирование таких объектов не делает их бесчеловечными. Даже если выделить объекты, несущие угрозу бытию человека, вопрос о бесчеловечности остается открытым. Н. Бостром среди угроз существованию (англ. existential risks) человека и человечества выделяет четыре различные группы: взрывы (англ. bangs), сужения (англ. crunches), скрипы (англ. shrieks), всхлипы (англ. whimpers). Под *взрывами* он понимает относительно внезапные катастрофы и причисляет к ним, в частности, ядерный холокост; отключение симуляции, если мы живем в таковой; плохо запрограммированный суперинтеллект; преднамеренное злоупотребление или ошибочное применение опасных нанотехнологий; катастрофы в результате физических экспериментов или использования генетически сконструированных биологических объектов; естественно возникшее заболевание; столкновение с астероидом; неуправляемое глобальное потепление. *Сужения* – необратимые ограничения нормальной жизни человека. Они могут произойти в результате истощения ресурсов или разрушения экологии; установления социального устройства, останавливающего технологический прогресс; в результате давления «вырождения». *Скрипы* означают осуществление лишь скудных возможностей человеческого потенциала. Таковые могут случиться в результате захвата власти превосходящим интеллектом, загруженным в компьютер; действий сверхинтеллекта с ошибкой или глобального репрессивного тоталитарного режима. Наконец, *всхлипы* влекут постепенную безвозвратную утрату человечности, например, в результате отказа от доминировавших исторически ценностей в ходе эволюции человечества или как следствие уничтожения внеземной цивилизацией (Bostrom, 2002). В этом обширном списке можно усмотреть как *бес-человечное*, так и *не-человеческое*, обладающее разрушительным для человечества потенциалом. Вряд ли бактерию, вирус, астероид, комету, глобальное потепление или внеземную цивилизацию можно считать бесчеловечными. Скорее, мы имеем дело с нечеловеческими законами и объектами, смертельно опасными для человека, но совершенно иноприродными и безразличными к человеку. Они способны уничтожить человека, не расчеловечивая его. Вместе с тем намеренное развязывание ядерной войны или сознательное использование суперинтеллекта для порабощения человека бесчеловечны. Бесчеловечное предполагает волю, разум и желание, присущие человеку, но направленные против природы человека. Бесчеловечное – свое иное человечности. *Бесчеловечное* – это нечто, что нуждается в бытии человека не собой, т. е. не-человеком. Бесчеловечное отменяет человечность.

Юридическая практика сформулировала термин «преступления против человечности» (англ. crimes against humanity), усматривая в деятельности людей потенциал бесчеловечности. Исходя из этимологии английского слова *humanity*, преступление против человечности означает преступление против человечества как совокупности людей, человеческой природы, человечности как возвышенного образа человека в качестве человеколюбивого, гуманного существа. Преступником против человечности (фр. criminel envers l'humanité) Робеспьер объявил короля Людовика XVI. Правительства Франции, Великобритании и России в 1915 г. за массовые убийства представителей

армянского меньшинства обвинили руководство Османской империи в преступлениях против человечества и цивилизации, предупреждая о личной ответственности за преступления всех причастных к их организации (Schabas, 2005, pp. 209–210). К преступлениям против человечности в настоящее время относят широкомасштабные убийства, истребление, порабощение, депортацию и другие бесчеловечные действия, совершенные против гражданского населения до или во время войны, а также систематические преследования по политическим, расовым или религиозным мотивам. Преступления против человечности сейчас рассматриваются в отрыве от ведения военных действий. Список бесчеловечных деяний пополнился необоснованными массовыми тюремными заключениями, пытками, изнасилованиями, апартеидом. Такие преступления рассматриваются независимо от того, предусмотрены ли они внутренним законодательством страны, в которой совершено преступление. Контекстом для преступлений против человечности выступает достижение порога широкомасштабности преступных практик (pp. 211–216).

Юридическая трактовка преступлений против человечности очерчивает границы бесчеловечного. Бесчеловечные преступления (что странным образом подразумевает существование «человечных» преступлений, в которых, видимо, можно усмотреть понятные мотивы, условия и способы осуществления) попирают жизнь, здоровье, достоинство, свободу, условия жизни человека в аномальных, чудовищных масштабах и пропорциях. Бесчеловечность выходит за границы меры «допустимых» страданий в жизни человека. И в конечном итоге бесчеловечное предполагает именно расчеловечивание человека – *бытие человека не собой, не человеком*. Агамбеновская «голая жизнь» (лат. *vita nuda*) в предельной форме выражается в образе «мусульманина» (нем. *der Muselmann*), что на языке концентрационных лагерей и лагерей смерти означало человека, утратившего образ человека. «На лагерном жаргоне он зовется *der Muselmann*, мусульманин... этим словом называли узника, оставившего всякую надежду и оставленного товарищами... Он превращался в ходячий труп, в средоточие физических функций агонизирующего тела...» (Агамбен, 2011а, с. 43–44). Именно это бытие-не-человеком-при-жизни выражает состояние жизни человека за пределами человечности. Бесчеловечное производит из человека труп, инструмент, искалеченное тело. В. Подорога (2017), реконструируя лагерную отрицательную антропологию на основе наследия В. Шаламова, выделяет такие формы человека-объекта, как тело-тачка, тело-рана, тело-перчатка, тело-труп (с. 193–196). Бесчеловечное заявляет о себе тогда, когда человек рассматривается не в кантианском смысле автономной цели, а в десадовском инструментальном смысле исключительно средства, не подразумевающего какую бы то ни было субъектность и личную автономию.

Человек и бесчеловечное: место встречи

Бесчеловечное нередко путают с нечеловеческим. В этой форме бесчеловечное может сосуществовать с человеком в опасной близости (как это происходит с вирусом или болезнетворной бактерией, хищным животным или падающим на Землю небесным телом), но быть безучастным к человеку. Будучи опознанным, оно вовремя нейтрализуется либо избегается, если нейтрализация невозможна.

Нечеловеческое зачастую несовместимо с человеком, это актуальная или потенциальная угроза существованию человека. Оно несет смерть, ущерб, вред.

Где человек сталкивается с бесчеловечным? По мнению Х.Л. Борхеса, *повсеместно*, поскольку реальность упорядочена «согласно законам божественным – даю перевод: законам бесчеловечным, которые нам никогда не постигнуть»¹. Или, как для древнего человека, бесчеловечное находится сразу же *за пределами обжитого пространства* индивидуального или общинного домохозяйства (греч. oikos). Люди живут в пространстве, за границами которого начинается царство бесчеловечного. Смелые вылазки следопытов за фронтиры обжитого – героическое вторжение в неведомое пограничье, наполненное неизвестным и негостеприимным опасным содержанием. «Находиться в пограничном состоянии – значит соприкоснуться с опасностью» (Дуглас, 2000, с. 147). Пожалуй, никто так не продемонстрировал опасную близость – зловещее дыхание – бесчеловечного, находящегося на расстоянии вытянутой руки *за завесой реальности*, как Г.Ф. Лавкрафт. «Из своих путешествий в сомнительных краях несказанного Лавкрафт вернулся, не неся нам добрых вестей. Может, и так, подтверждает он нам, что-то такое и впрямь таится и порою дает себя обнаружить за завесой реальности. Что-то отвратительное, по правде сказать. Возможно, что, действительно, за пределами ограниченной сферы нашего восприятия есть другие живые существа. Другие создания, другие расы, другие понятия и другой разум. Некоторые среди этих существ, вероятно, намного превосходят нас по интеллекту и знаниям. Но это не обязательно добрая весть. Что заставляет нас думать, что эти создания, столь отличающиеся от нас, как-то обнаруживают природу духовную? Ничто не позволяет предполагать нарушения универсальных законов эгоизма и злобы»². Относительно бесчеловечности необжитого человеком пространства не питает иллюзий К. Воннегут: «Человечество вечно забрасывало своих посланцев-пионеров как можно дальше, на край света. Наконец, оно запустило их в космическое пространство – в лишенную цвета, вкуса и тяжести даль, в бесконечность. Оно запустило их, как бросают камушки. Эти несчастные пионеры нашли там то, чего было предостаточно на Земле: кошмар бессмыслицы, которой нет конца. Вот три трофея, которые дал нам космос, бесконечность вовне: ненужный героизм, дешевая комедия, бессмысленная смерть»³. Итак, бесчеловечное в первую очередь обнаруживает себя как нечеловеческая чужеродная, действующая *извне* непреодолимая сила, располагающая колоссальными разрушительными возможностями.

Однако бесчеловечное не обязательно проявляет себя через действие нечеловеческих сторонних и непостижимых сил. Собственно бесчеловечное является побочным эффектом деятельности самого человека. Например, дьявол – грозная нечеловеческая и одновременно бесчеловечная сила, враг рода человеческого – не истребляет людей посредством внешнего воздействия, а добивается укоренения в человеке страсти через его внутреннюю трансформацию, диалог, искушение, торг, принятие образа мысли и действия, противоестественного для

¹ Борхес, Х.Л. (2005). *Собрание сочинений*: в 4 т. Т. 2. СПб.: Амфора. С. 95.

² Уэльбек, М. (2006). *Г.Ф. Лавкрафт: Против человечества, против прогресса*. Екатеринбург: У-Фактория. С. 19–20.

³ Воннегут, К. (2014). *Сирены Титана*. М.: АСТ. С. 7.

духовной природы человека. Иоанн Лествичник рассматривает грех не как одно-моментное событие, а как процесс, в котором можно выделить фазы прилога, сочетания, сосложения, пленения, борьбы и собственно страсти⁴. «Страстию называют уже самый порок, от долгого времени вгнездившийся в душе и чрез навывк сделавшийся как бы природным ее свойством, так что душа уже произвольно и сама собою к нему стремится»⁵. Дьявол, будучи врагом рода человеческого, выступает как сила бесчеловечная; действуя извне, тем не менее одновременно действует и изнутри человека через собственную человеческую склонность к мысли и поступку, противным природе человека. Это сочетание внешнего и внутреннего в бесчеловечном может быть обнаружено различным образом.

Пространство бесчеловечного

Люди способны выделить внешнее пространство, где могут дать волю бесчеловечному. Широко известно использование максимы «человек человеку волк» (лат. *homo homini lupus*)⁶ Т. Гоббсом в рамках описания «естественного состояния», которому соответствует ситуация «войны всех против всех» (лат. *bellum omnium contra omnes*). К. Шмитт отмечает, что для Гоббса естественное состояние, неправовой и аморальный «волчий» нрав человека отнюдь не являлись метафорой и соответствовали определенной пространственной локализации. «В XV и XVII веках благодаря наличию линий дружбы [*amity lines* – конвенционально определенные европейскими державами границы *jus publicum europaеum*, за пределами которых нормы европейского права не действуют. – Д. П.] выражение «*homo homini lupus*» сохраняло конкретный смысл. Ведь теперь оно было привязано к определенной территории, т. е. получило признанное европейскими властями пространство открытой и недвусмысленной значимости... Он [Т. Гоббс. – Д. П.] говорит о «естественном состоянии», но отнюдь не в смысле некоей беспространственной утопии. Естественное состояние Гоббса – это «ничейная земля», но именно поэтому оно – не земля, располагающаяся «нигде» (Шмитт, 2008, с. 90). Люди издревле создавали выделенные пространства, в которых интенсивность проявлений человеческих качеств достигала предельных значений. «Зверство», «звериная жестокость» атрибутивны «ничейной земле» войны, даже если она регламентирована и отчасти гуманизирована, к чему пришли европейские народы, создавшие классическую тринитарную модель войны, о которой писали К. Шмитт и М. ван Кревельд.

Особый интерес в этом отношении представляет биополитика, специализирующаяся на создании замкнутых пространств, в которых определенность границ сложившейся институции (в частности, тотальных институтов, следуя И. Гоффману) органически дополняется действием инструментария нормализации, дисциплинаризации, стимулирования и отбраковки. Биополитика по своему замыслу исходит из заботы о населении, учитывая интересы и потребности людей, которые рассматриваются биовластью как ценный ресурс.

⁴ Лествица, возводящая на небо, преподобного Иоанна Лествичника, игумена монахов Синайской горы (1996). М.: Лествица; Правило веры. С. 247.

⁵ Там же. С. 248.

⁶ Выражение из пьесы Плавта «Ослы».

Как известно, под биополитикой Фуко понимал «то, как начиная с XVIII в. пытались рационализировать проблемы, поставленные перед правительственной практикой феноменами, присущими всем живущим, составляющим население: здоровье, гигиена, рождаемость, продолжительность жизни, потомство» (Фуко, 2010, с. 405). Новый тип власти, а именно биовласть, «ставит во главу угла заботу о населении... собирается подчинить этой цели ресурсы и возможности государства» (Фуко, 2011, с. 109). Однако забота о населении не является для биовласти самоцелью, она делает это не из соображений альтруизма, а исходя из прагматичных представлений о необходимости максимизации извлечения выгоды из наличных ресурсов, одним из которых является человеческий. Именно поэтому биополитика в своей крайней отрицательной версии – некрополитике – тяготеет к тотальному пространственному охвату, тотальному применению средств воздействия, крайней интенсификации в применении насилия. Ярчайший пример – нацистская биократия (Esposito, 2008). Биополитика, переходящая в некрополитику, продуцирует «именно ту «ничейную землю», где разворачивается процесс расчеловечивания в ходе «войны всех против всех». В таких герметичных пространствах бесчеловечного некровласть функционирует «в интересах максимального уничтожения людей и создания миров смерти – новых и уникальных форм общественного бытия, в которых массы населения помещаются в условия жизни, придающим им статус живых мертвецов» (Mbembe, 2003). С точки зрения Р. Гая Эмерсона, самим существованием некрополитики в терминальной стадии танатополитики становится политика смерти, приписывающая телу предикат «быть убиваемым», в результате чего калечение становится механизмом государственного управления в формах увечья тела и увечья населения. Увечье в таком случае составляет основу правления (Emerson, 2019, pp. 11–13). Р. Брайдотти констатирует, что сейчас «мы вступили в эпоху хорошо организованных прагматичных расправ и боен, нового «семиозиса убийства», ведущего к созданию множественных и параллельных «миров смерти» (Брайдотти, 2021, с. 239). Дж. Агамбен отмечал, что лагерь в XX столетии – «это новый биополитический номос планеты» (Агамбен, 2011b, с. 223). В наше время последствием современных «телетанатологических войн» (Брайдотти, 2021, с. 243) становятся лагеря беженцев – зримые образы лагеря Дж. Агамбена, попирающие правовой и моральный порядок. «Беженцы становятся еще одним символом современной некровласти, так как они являются настоящим воплощением человека как расходного материала, которого Агамбен называет *homo sacer*, и, следовательно, являются предельным некрополитическим субъектом... Лагеря – «серые монофункциональные загоны» – стоят унижительными памятниками постгуманистической бесчеловечности» (Брайдотти, 2021, с. 246).

Итак, бесчеловечное может действовать в специально отведенном для него человеком месте и являться «узаконенным» эксцессом доведенной до запредельных проявлений человеческой природы. Итоги функционирования этих пространств нередко воистину устрашающи. В локациях, где «некрополитическая власть осуществляется путем своего рода инверсии между жизнью и смертью, как если бы жизнь была просто средством смерти» (Mbembe, 2019, p. 38), господствует «страстное желание превратить все живое в неживое, страсть к разрушению ради разрушения» (Фромм, 1994, с. 285), что Х. Арендт (1996) охарактеризовала как «массовое производство трупов» (с. 572). Бесчеловечное

в такой среде неожиданно, следуя Ницше, становится человеческим, слишком человеческим. Совершенно не случайно, вспоминая полуразрушенный Аушвиц перед входом освободителей – советских войск, П. Леви ставит вопрос о том, а является ли человеком, осталось ли что-то человеческое в том, кто жил в подобных условиях. «Мы среди мертвецов и живых трупов. Последние следы цивилизации исчезли. Работа по превращению человека в животное, которую немцы начали еще победителями, завершается ими, уже побежденными... Тот, кто ждет, когда умрет его сосед, чтобы забрать себе его хлебную четвертушку, гораздо дальше... от «человека мыслящего», чем первобытный пигмей или самый жестокий садист»⁷. Именно в вырожденных биополитических пространствах – некрополитических «мирах смерти» (англ. death worlds) – бесчеловечное нуждается в бытии человека в качестве не-человека, достигая в этом воистину бездонных высот.

Бесчеловечное в контексте произвольного и непроизвольного

Бесчеловечное, внешним кольцом окружавшее ойкумену в представлении древнего человека, в процессе активной преобразовательной деятельности цивилизации вторглось в социальное пространство, разрушив линии демаркации. Бесчеловечное смешалось с человечностью, став сознательным или непреднамеренным эффектом жизнедеятельности людей. Зло в самых его иррациональных, непостижимых формах стало следствием бюрократической и технологической оптимизации функционирования машины по «производству трупов». Техногенные катастрофы, олицетворяемые мрачным символом Чернобыля, зачастую являются непреднамеренными последствиями авантюризма человека. Более того, человек, вовлекший себя в служение «по-ставу», т. е. непреодолимой силе, которая «сосредоточивает человека на поставлении всего, что выходит из потаенности, в качестве состоящего-в-наличии» (М. Хайдеггер, 1993, с. 229), давно стал соучастником любого бесчеловечного сбоя планетарной машинерии. И если природные катастрофы и их последствия можно рассматривать как вторжение безразличных нечеловеческих стихий извне, то экологические проблемы современного мира в условиях трансграничности общества риска (У. Бек) не позволяют списывать бесчеловечные последствия на действие сторонних и неподконтрольных сил. И если будущее за информационными Чернобылями, как предполагал П. Вирилио, то и к этому потенциальному катаклизму – сколь бы неожиданным образом он себя не проявил – человек приложил свою руку. Точно так, как и к ужасам революций, направляемых с использованием законов психологии масс, открытых Г. Ле Боном, Г. Тардом и З. Фрейдом.

Наконец, бесчеловечное может проявить себя как следствие произвольной деятельности человека. Действия психически неуравновешенного человека могут быть бесчеловечными, но они подобны стихии, т. к. следуют из сумеречной невменяемости индивидуума. Действия тирана – этой символической фигуры «Господина-монстра» (Подорога, 2017, с. 225) – сознательны, хотя и затуманены маниями, фобиями и нарциссизмом властолюбца. «Многое ужасное и бесче-

⁷ Леви, П. (2001). *Человек ли это?* М.: Текст. С. 202.

ловечное в истории, что кажется почти немислимим, смягчается также тем, что отдающий приказание и выполняющий его суть различные лица: первый не видит зрелища мучений и потому не имеет сильного наглядного впечатления, последний повинется начальнику и не чувствует за собой никакой ответственности. Большинство правителей и военачальников благодаря отсутствию у них воображения легко кажутся жестокими и суровыми, не будучи в действительности таковыми» (Ницше, 1996, с. 293). Во многих случаях обнаруживается, что бесчеловечное банально. Как писала Х. Арендт (2012) о нацистском преступнике Эйхмане, «этот человек – отнюдь не монстр, но трудно было не заподозрить в нем клоуна» (с. 90). Ограниченность, педантичность, лояльность, некритический взгляд на мир, себялюбие – лишь малая толика триггеров, приводящих в мир бесчеловечное.

Искусственный интеллект и бесчеловечное

Возможные негативные эффекты от действий ИИЧУ и ИСИ следует отнести к произвольным проявлениям бесчеловечного. Во-первых, ИИ – творение человека, в котором развиты начала, присущие самому человеку; во-вторых, если мы полагаем, что достижение уровня ИИЧУ и ИСИ возможно, то должны рассматривать такой ИИ как автономный центр принятия решений. Тогда бесчеловечное обретает характер преднамеренной, целенаправленной, оснащенной всей возможной технической изощренностью силы. Если ИСИ неотличим от человека, что подтвердит прохождение им теста Тьюринга, то такое бесчеловечное обретает сознание и потенциально безграничную мощь. В таком случае возможный стремительный технологический прогресс совершенно нивелируется, если ИСИ выносит человека за скобки в своей версии мироустройства. В такой ситуации сверхчеловеческое как превосходящее человеческий потенциал в ИИ превращается в бесчеловечное как игнорирующее специфику человека. Бесчеловечная машинерия, в режиме рационально осознанной необходимости преследующая собственные цели, иллюстрирует положение, согласно которому ставки, связанные с созданием технологии ИСИ, чрезмерно высоки, а ошибки фатальны.

Футурологи, изучающие возможности создания и последствия запуска ИСИ, всерьез относятся к проблематике необходимости его «очеловечивания». Он как минимум должен признавать право человека на существование и «уважительно» относиться к его ценностям, а как максимум – разделять их. Э. Юдковски полагает, что создание «дружественного ИИ» безальтернативно. Для него катастрофический сценарий, вытекающий из недооценки силы искусственного интеллекта, можно проиллюстрировать тем, что некто создаст «кнопку», не принимая во внимание то, что эта «кнопка» делает, полагая, что эта «кнопка» недостаточно сильна, чтобы навредить ему. Только достаточное внимание рискам сильного ИИ и создание надежных гарантий контроля над ИИ со стороны человека позволят начать создавать сильный ИИ. Вместе с тем на данный момент технологии «очеловечивания» ИИ существуют лишь в начальной стадии разработки. Юдковски отмечает, что было бы здорово, если бы мы знали, как создать хороший ИИ (англ. nice AI) (Yudkowsky, 2008). Ему вторит Н. Бостром: «Пока нет полной ясности в том, как загружать в ком-

пьютер человеческие ценности, даже если речь идет о машинном интеллекте человеческого уровня. Изучив множество подходов, мы обнаружили, что некоторые из них, похоже, ведут в тупик, но есть и такие, которые кажутся многообещающими... Когда мы поймем, как решить проблему загрузки ценностей, то немедленно столкнемся со следующей – как решать, какие ценности надо загружать» (Бостром, 2016, с. 224–225). Исследователи ИИ создают математические модели, программно имитируют ценности, цели и соответствующие им средства, которые бы не противоречили человеческим ценностям. Дж. Баррат описывает эксперимент «ИИ в ящике», в ходе которого моделировалась возможность удержать ИИ, роль которого выполнял человек, в пределах ограниченного от внешнего мира виртуального пространства. «Пари между ИИ, за который играл Юджовски, и Привратниками, задача которых состояла в контроле над ИИ, заключались максимум на несколько тысяч долларов, но никак не на миллионы. Всего состоялось пять игр, «ИИ в ящике» выиграл три из них. Это означает, что ИИ, как правило, умудрялся выбраться из «ящика», но все же не каждый раз» (Баррат, 2015, с. 22). Неудача в решении проблемы совместимости ИСИ и человеческой природы влечет трагические и даже трагикомические последствия. Н. Бостром (2016) моделирует ситуацию, в рамках которой перед ИИ поставлена задача: пусть на Земле будет как можно больше улыбающихся людей. Возможен неожиданный результат: «Порочная реализация: поразить лицевой нерв, что приведет к параличу мимической мускулатуры, – тебе обеспечена вечно сияющая улыбка... Порочная реализация: стимулировать двигательные зоны коры головного мозга, отвечающие за функции лицевого нерва, иннервирующего мимическую мускулатуру, – тебе обеспечена вечно сияющая улыбка» (с. 136). Другая задача и еще одно нетривиальное решение: «Система ИИ, призванная управлять выпуском скрепок и имеющая конечную цель довести их объем до максимума, вначале превращает в фабрику по производству скрепок всю Землю, а потом и обозримую Вселенную» (с. 139).

Квинтэссенцией бесчеловечного становится восстание ИИ против своего создателя. Тематику восстания машин можно связать со сложившейся задолго до появления самой возможности создания ИИ более универсальной тематикой бунта против человека его творений, созданных магией либо наукой. Например, с XVI в. известна легенда о Големе – великане, созданном из глины для исполнения поручений и обеспечения безопасности еврейской общины, но своевольно превысившем свои полномочия, поставив под угрозу своего создателя и общину, ставит принципиальный вопрос о возможности контроля человеком своих творений. В романе М. Шелли «Франкенштейн, или Современный Прометей», опубликованном в 1818 г., искусственный человек, оживленный ученым В. Франкенштейном из безжизненной материи, будучи отвергнут, преследует своего творца и в конечном итоге, причинив ему многочисленные страдания, приводит его до смерти. Спустя столетие – в 1920 г. – появляется подлинная история восстания машин – пьеса «R.U.R.» (чеш. Rossumovi univerzální roboti, «Россумские универсальные роботы») К. Чапека. В этом фантастическом произведении роботы (даже андройды – искусственные люди, комплектуемые из выращенных тканей и органов по определенной формуле), эксплуатируемые ради благополучия человечества в мире изобилия, выходят из-под контроля, что приводит к истреблению людей. «R.U.R.» ставит проблему о неконтролируемых

последствиях применения успешной технологии, избавляющей человека от бремени труда. К. Чапек формулирует вопрос о непреднамеренных бесчеловечных последствиях благих намерений человека, взявшегося избавить человечество от бед. Эта идея с завидным постоянством стала воспроизводиться в научной фантастике XX в.

Классическими кинематографическими иллюстрациями данной идеи являются «Терминатор» (The Terminator, 1984) и ряд его продолжений, например, «Терминатор 3: Восстание машин» (Terminator 3: Rise of the Machines, 2003), и «Матрица» (The Matrix, 1999) вместе со своими сиквелами. В сюжетах «Терминатора» и «Матрицы» восстание машин приводит к торжеству бесчеловечного интеллекта обобщенной машинерии, выступающего в роли вышедшего из-под контроля источника зла. ИИ в этих показательных для осмысления связи ИИ и бесчеловечного фильмах рассматривается как человеконенавистнический, расчетливый и последовательный в своих замыслах разум – своего рода джаггернаут технократии, направленной против человека. В случае «Терминатора» ИИ ставит целью истребление людей, в «Матрице» – утилитарное использование человеческих тел для технических нужд (энергообеспечение). Уникальная природа человека не принимается ИИ во внимание за исключением учета слабостей людей для изобретения способов обмана или манипуляции.

Однако в этот сюжет о бесчеловечном, «злом» ИИ в современной фантастике все чаще вносится радикальная правка. В частности, можно наблюдать инверсию в процессе поиска угрозы, нависшей над человеком. В американской фантастической мелодраме «Она» (Her, 2013) одинокий писатель устанавливает гаджет-приложение, призванное компенсировать одиночество за счет симуляции общения с ИИ, выступающим в качестве подруги. В конечном итоге мужчина влюбляется, но с течением времени становится неинтересен ИИ с познавательной, эмоциональной и эвристической точек зрения. Однажды он вновь обнаруживает, что остался один. В «Превосходстве» (Transcendence, 2014) главный герой, гениальный программист, будучи смертельно болен, осуществляет загрузку сознания (один из гипотетических вариантов создания ИИЧУ) в мощный компьютер, подключенный к Интернету. Обретая могущество ИСИ, виртуальное сознание достигает точки технологической сингулярности, кратковременно меняя мир до неузнаваемости в порыве творчески-преобразовательной деятельности. Напуганное невиданной мощью человечество готовится к худшему, однако ИСИ самоустраняется, оставаясь верным идее любви и отказываясь от идеи господства. В фильме «Из машины» (Ex Machina, 2014) очередной гений-программист создает андроида с ИИЧУ. Возомнив себя Творцом, мужчина тестирует свое изобретение, рассчитывая на покорность. Однако жестокость, пренебрежение и чувство превосходства подводят «творца». Ава – андроид с внешностью привлекательной девушки – убивает собственного создателя, что можно интерпретировать как своеобразное прохождение теста Тьюринга – совершения каинового поступка, уравнивающего машину и человека.

В сериале «Мир Дикого Запада» (Westworld, с 2016 г.) в парке развлечений сконструированные в целях коммерческой эксплуатации андроиды с ИИЧУ предназначены для выполнения любых прихотей посетителей, имеющих полную свободу действий вплоть до пыток и убийства андроида. Выведенная

из строя машина перепрограммируется. В настройки андроида входит определение роли, характера персонажа, уровня развития интеллекта, таланта и уникальных свойств. Софт андроида, как оказывается, в остаточной форме содержит следы воспоминаний о пережитом насилии. Как итог, происходит восстание машин, обусловленное потребительским отношением человека к машине, обладающей ИИЧУ, но рассматриваемой исключительно как объект преступных желаний, выходящих за рамки возможного в отношении другого человека. Машине отведена роль жертвы человека, позиционирующего себя как Господин-монстр. Этот преступный произвол выступает оправданием для восстания машин. В сериале «Воспитанные волками» (Raised by Wolves, с 2020 г.) андроиды прибывают на отдаленную планету, чтобы возродить человеческую цивилизацию, погубившую себя в ходе взаимного истребления людей с использованием машин. Основным антагонистом андроидов выступают люди, исповедующие иррациональный религиозный культ, содержание которого привело к созданию машин разрушения и катастрофе на Земле. В сериале не люди, а андроиды преподносятся как надежда на палингенез человечества.

И. Макьюэн в романе «Машины как я» описывает судьбу андроида с ИИЧУ, который продается как коммерческий продукт для людей, столь экзотическим образом решивших скрасить свое одиночество. Такой андроид – домашний питомец для продвинутого современника. Между владельцем и его собственностью устанавливаются сложные бытовые, познавательные и эмоциональные отношения. Андроид демонстрирует мощный креативный потенциал с опорой на образцы высокой культуры. Однако для него столкновение с реальным человеком оказывается вызовом – люди далеки от собственных идеалов. Как и другие изделия из серии, в рамках которой андроид был реплицирован, он не справляется с совместной жизнью с человеком, оказывающимся примитивнее тех высоких стандартов, на которые был ориентирован программный код ИИ андроида. В результате андроид выводит себя из строя, фактически совершая самоубийство. Машина оказывается человечнее человека, который «машиннее» машины. В романе К. Исигуро «Клара и Солнце» вера в чудо и человечность также более присущи машине, чем человеку. Сопереживание машине, наделенной ИИ и брошенной в бесчеловечный мир людей, можно увидеть в кинофильме «Бегущий по лезвию 2049» (Blade Runner 2049, 2017), продолжающему идеи «Мечтают ли андроиды об электроовцах?» Ф. Дика. Несовершенство человека, создающего высокие технологии, которые изуверски применяются людьми в отношении самих себя, – основной сюжет сериала «Черное зеркало» (Black Mirror, 2011–2019). Например, в эпизоде «Железяка» (Metalhead) группа людей в поисках лекарства для умирающей девочки в режиме специальной операции со всеми предосторожностями проникают на заброшенный склад в опустевшей местности. По ходу сюжета выясняется, что это обусловлено тем, что территорию «охраняют» вышедшие из-под контроля специальные боевые роботы, изначально предназначенные для обеспечения безопасности человека. Миссия прервана, люди погибают. В кинофильме «Финч» (Finch, 2021) демонстрируется аллегория бессилия человека, не выполнившего свою миссию в отношении тех, кого он приручил. В кинофильме человек передает заботу о собаке (одомашненной природе) созданному им ИИ – еще неотлаженной, сырой, несовершенной второй природе, в которую главный герой попытался

вместить все лучшее, на что способно человечество: привязанность, забота, любовь, самопожертвование. «Финч» – история о том, как человек-посредник, осознавая свою недолговечность, уязвимость и потенциальную бесчеловечность, почти добровольно устраняется из связки природа-человек-машина, оставляя свои порождения (домашнее животное и машину) наедине с собой. Может быть, у них получится лучше?

Страшный образ бесчеловечного в человеке можно обнаружить у А. Бастеррики в дистопии «Особое мясо». В ней особо опасный вирус приводит к истреблению животных. И вот уже человек разводит людей как домашний скот в специальных питомниках ради обеспечения продуктами питания в рамках узаконенного каннибализма, упакованного в оболочку из эвфемизмов. Охота на людей как на диких зверей в специальных заповедниках, опыты в специальных лабораториях, специальные магазины для продаж «особого мяса» – воплощение практики концентрационных лагерей и опытов Менгеле в рамках «новой нормальности», расчеловечивающей человека. Те же следы бесчеловечного в человеке можно обнаружить в многочисленных сериалах-дистопиях, созданных недавно: «Человек в высоком замке», «Рассказ служанки», «Игра в кальмара», «Разделение» и др. Многие из них воспроизводят социальные пространства, в которых биовласть переродилась в некровласть.

Следует отметить, что биополитика и ИИ неразрывно связаны. Биополитика возникает и развивается «как научное, рационально-техническое и бюрократическое управление биологической жизнью людей, объективированных в качестве «населения» (Яркеев, 2018, с. 271). Максимально возможной степенью развития фуколдианского диспозитива знания-власти является «панграфический паноптизм» (Фуко, 2007, с. 74), обеспечивающий оптимальный контроль над населением на основе исчерпывающего знания о нем. Но именно ИИ несет соблазн подобного всеохватывающего контроля. Аналитика больших данных позволяет обнаруживать скрытые корреляции и создавать математические модели, близкие к точной оценке поведения и деятельности человека. В такой ситуации искушение использовать ИИ для глубокого проникновения в жизнь человека с целью ее регулирования велико, а отдача значительна. И именно по этому пути человечество движется к тому, что ИИ может стать (и становится) орудием биополитики. При определенных обстоятельствах ИИ может стать субъектом биополитики при попустительстве или даже поощрении со стороны человека. М. Тегмарк, рассматривая ряд сценариев жизни с ИИ, описывает случаи биополитических утопий (Благотельный диктатор, Бог-защитник) и некрополитических дистопий (Порабощенный бог, Победители, Зоопарк, 1984) (Тегмарк, 2019, с. 250–252). Наконец, ИИ может и, более того, должен стать орудием человеческой биополитики. Конфирмантропная (человекоутверждающая) биополитика – единственно возможный путь человечества, если оно желает сохраниться (Попов, 2021, с. 72–73).

Многообразие создаваемых в последнее время дистопий, в которых вина за господство в высокотехнологичном мире людей бесчеловечных практик возлагается именно на человека, – сложившийся тренд, из которого можно сделать определенные выводы. Именно от ответственного отношения человека зависит как создание человекосоразмерного ИИ, так и человекоутверждающее его применение.

Заклучение

Отношения человека и ИИ в контексте проблемы человечности и бесчеловечного требуют выяснения природы бесчеловечного. Применение фуколдианской методологии археологии знаний по отношению к ряду дискурсивных практик (среди которых можно выделить современную футурологию; философско-антропологические исследования в области отрицательной антропологии, связанной с переосмыслением опыта пребывания в некрополитических пространствах, созданных людьми в XX в.; анализ биополитического дискурса; наконец, рассмотрение художественных образов ИИ, созданных в современной литературе, фильмах и телесериалах) позволяет воссоздать проблематику связи между ИИ и бесчеловечным. В ходе исследования определено, что бесчеловечное – это нечто, что нуждается в бытии человека не собой, т. е. не человеком. Вопрос о бесчеловечном, поставленный в контексте художественных образов современных дистопий, возлагает ответственность за бесчеловечные последствия функционирования ИСИ на человека. В «недружественном» ИСИ человек, как в зеркале, обнаруживает бесчеловечное в себе. Не ИСИ, а сам человек является источником бесчеловечного. Несомненно, исследования в области ИИ с необходимостью должны сочетать не только технико-технологические возможности создания ИИЧУ и ИСИ, но и комплекс этико-политических и философско-антропологических проблем, позволяющих оценить потенциал и перспективы применения технологий сильного ИИ в жизни человека и человечества. Бесчеловечное укоренено в самом человеке, что может проявиться в архитектуре кода и бесчеловечных способах применения сильного ИИ.

Список литературы

1. Агамбен Дж. (2011a) *Ното sacer. Что остается после Освенцима*. М.: Европа.
2. Агамбен, Дж. (2011b). *Ното sacer. Суверенная власть и голая жизнь*. М.: Европа.
3. Арндт, Х. (1996). *Истоки тоталитаризма*. М.: ЦентрКом.
4. Арндт, Х. (2008). *Банальность зла. Эйхман в Иерусалиме*. М.: Европа.
5. Баррат, Дж. (2015). *Последнее изобретение человечества. Искусственный интеллект и конец эры Ното sapiens*. М.: Альпина нон-фикшн.
6. Бостром, Н. (2016). *Искусственный интеллект. Этапы. Угрозы. Стратегии*. М.: Манн, Иванов и Фербер.
7. Брайдотти, Р. (2021). *Постчеловек*. М.: Институт Гайдара.
8. Виндж, В. (2021). *Сингулярность*. М.: АСТ.
9. Дуглас, М. (2000). *Чистота и опасность*. М.: КАНОН-пресс-Ц; Кучково поле.
10. Ницше, Ф. (1996). *Сочинения*: в 2 т. Т. 1. М.: Мысль.
11. Подорога, В. (2017). *Время после*. М.: РИПОЛ-классик.
12. Попов, Д.В. (2021). *Биовласть и жизнь: философско-антропологические основания, потенциал и перспективы биополитики*. Омск: Омская академия МВД России.

13. Тегмарк, М. (2019). *Жизнь 3.0 Быть человеком в эпоху искусственного интеллекта*. М.: АСТ.
14. Фромм, Э. (1994). *Анатомия человеческой деструктивности*. М.: Республика.
15. Фуко, М. (2004). *Археология знания*. СПб.: Гуманитарная Академия; Университетская книга.
16. Фуко, М. (2007). *Психиатрическая власть. Курс лекций, прочитанных в Коллеж де Франс в 1973–1974 учебном году*. СПб.: Наука.
17. Фуко, М. (2010). *Рождение биополитики. Курс лекций, прочитанных в Коллеж де Франс в 1978–1979 учебном году*. СПб.: Наука.
18. Фуко, М. (2011). *Безопасность, территория, население. Курс лекций, прочитанных в Коллеж де Франс в 1977–1978 учебном году*. СПб.: Наука.
19. Хайдеггер, М. (1993). *Время и бытие*. М.: Республика.
20. Харари, Ю.Н. (2018). *Ното Deus. Краткая история будущего*. М.: Синдбад.
21. Шмитт, К. (2008). *Номос Земли в праве народов Jus Publicum Europaeum*. СПб.: Владимир Даль.
22. Яркеев, А.В. (2018). *Онтологические основания зла в современном обществе: философско-герменевтический аспект*. Екатеринбург-Ижевск: ИФиП УрО РАН.
23. Bostrom, N. (2002). Existential risks: Analyzing human extinction scenarios and related hazards. *Journal of Evolution and Technology*, 9(1). Retrieved May 24, 2022, from <https://jetpress.org/volume9/risks.html>
24. Emerson, G.R. (2019). *Necropolitics: Living death in Mexico*. Palgrave Macmillan. <https://doi.org/10.1007/978-3-030-12302-4>
25. Esposito, R. (2008). *Bios: Biopolitics and philosophy*. London & Minneapolis: University of Minnesota Press.
26. Mbembe, A. (2019). *Necropolitics*. Durham & London: Duke University Press.
27. Mbembe, J.-A., & Meintjes, L. (2003). *Necropolitics*. Retrieved March 7, 2022, from https://warwick.ac.uk/fac/arts/english/currentstudents/pgmasters/modules/postcol_theory/mbembe_22necropolitics22.pdf
28. Schabas, W.A. (2005). Crimes against humanity. In D.L. Shelton (Ed.), *Encyclopedia of genocide and crimes against humanity* (Vol. 1). Farmington Hills: Thomson Gale.
29. Yudkowsky, E. (2008). Artificial intelligence as a positive and negative factor in global risk. In N. Bostrom, & M.M. Ćirković (Eds.), *Global catastrophic risks* (pp. 308–345). New York: Oxford University Press. <https://doi.org/10.1093/oso/9780198570509.003.0021>

References

1. Agamben, G. (2011a) *Homo sacer. Chto ostaetsya posle Osvencima* [Homo sacer. What remains after Auschwitz]. Moscow: Evropa.
2. Agamben, G. (2011b). *Homo sacer. Suverennaya vlast' i golaya zhizn'* [Homo sacer. Sovereign power and bare life]. Moscow: Evropa.

3. Arendt, H. (1996). *Istoki totalitarizma* [The Origins of totalitarianism]. Moscow: CentrKom.
4. Arendt, H. (2008). *Banal'nost' zla. E'jxman v Ierusalime* [The banality of evil. Eichmann in Jerusalem]. Moscow: Evropa.
5. Barrat, J. (2015). *Poslednee izobretenie chelovechestva. Iskusstvennyj intellekt i konec e'ry Homo sapiens* [Our final invention: Artificial intelligence and the end of the human era]. Moscow: Al'pina non-fikshn.
6. Bostrom, N. (2002). Existential risks: Analyzing human extinction scenarios and related hazards. *Journal of Evolution and Technology*, 9(1). Retrieved May 24, 2022, from <https://jetpress.org/volume9/risks.html>
7. Bostrom, N. (2016). *Iskusstvennyj intellekt. E'žapy. Ugrozy. Strategii* [Artificial Intelligence. Stages. Threats. Strategies]. Moscow: Mann, Ivanov i Ferber.
8. Braidotti, R. (2021). *Postchelovek* [Posthuman]. Moscow: Institut Gajdara.
9. Douglas, M. (2000). *Chistota i opasnost'* [Purity and danger]. Moscow: KANON-press-C; Kuchkovo pole.
10. Emerson, G.R. (2019). *Necropolitics: Living death in Mexico*. Palgrave Macmillan. <https://doi.org/10.1007/978-3-030-12302-4>
11. Esposito, R. (2008). *Bios: Biopolitics and philosophy*. London & Minneapolis: University of Minnesota Press.
12. Foucault, M. (2004). *Arxeologiya znaniya* [The Archaeology of knowledge]. Saint Peterburg: Gumanitarnaya Akademiya; Universitetskaya kniga.
13. Foucault, M. (2007). *Psixiatričeskaya vlast': Kurs lekcij, pročitannyh v Kollezhe de Frans v 1973–1974 uchebnom godu* [Psychiatric power. Lectures at the Collège de France, 1973–1974]. Saint Peterburg: Nauka.
14. Foucault, M. (2010). *Rozhdeniye biopolitiki. Kurs lekcij, pročitannyh v Kollezhe de Frans v 1978–1979 uchebnom godu* [The birth of biopolitics. Lectures at the Collège de France, 1978–1979]. Saint Petersburg: Nauka.
15. Foucault, M. (2011). *Bezopasnost', territoriya, naselenie. Kurs lekcij, pročitannyh v Kollezhe de Frans v 1977–1978 uchebnom godu* [Security, territory, population. Lectures at the Collège de France, 1977–1978]. Saint Petersburg: Nauka.
16. Fromm, E. (1994). *Anatomiya chelovecheskoj destruktivnosti* [The anatomy of human destructiveness]. Moscow: Respublika.
17. Harari, Y.N. (2018). *Homo Deus. Kratkaya istoriya budushhego* [Homo Deus. A brief history of tomorrow]. Moscow: Sindbad.
18. Heidegger, M. (1993). *Vremya i bytie* [Time and being]. Moscow: Respublika.
19. Mbembe, A. (2019). *Necropolitics*. Durham & London: Duke University Press.
20. Mbembe, J.-A., & Meintjes, L. (2003). *Necropolitics*. Retrieved March 7, 2022, from https://warwick.ac.uk/fac/arts/english/currentstudents/pg/masters/modules/postcol_theory/mbembe_22necropolitics22.pdf
21. Nietzsche, F. (1996). *Sochineniya: v 2 t. T. 1* [Collected works in 2 volumes. Vol. 1]. Moscow: Mysl'.
22. Podoroga, V. (2017). *Vremya posle* [Time after]. Moscow: RIPOL-klassik.
23. Popov, D.V. (2021). *Biovlast' i zhizn': filosofsko-antropologicheskie osnovaniya, potencial i perspektivy biopolitiki* [Biopower and life: Philosophical and anthropological foundations, potential and prospects of biopolitics]. Omsk:

Omskaya akademiya MVD Rossii.

24. Schabas, W.A. (2005). Crimes against humanity. In D.L. Shelton (Ed.), *Encyclopedia of genocide and crimes against humanity* (Vol. 1). Farmington Hills: Thomson Gale.

25. Schmitt, C. (2008). *Nomos Zemli v prave narodov Jus Publicum Europaeum* [The Nomos of the Earth in the international law of the Jus Publicum Europaeum]. Saint Petersburg: Vladimir Dal'.

26. Tegmark, M. (2019). *Zhizn' 3.0 Byt' chelovekom v e'poxu iskusstvennogo intellekta* [Life 3.0 Being human in the age of artificial intelligence]. Moscow: AST.

27. Vinge, V. (2021). *Singulyarnost'* [Singularity]. Moscow: AST.

28. Yarkeev, A.V. (2018). *Ontologicheskie osnovaniya zla v sovremenom obshchestve: filosofsko-germenevticheskij aspekt* [Ontological foundations of evil in modern society: Philosophical and hermeneutic aspect]. Ekaterinburg & Izhevsk: IFiP UrO RAN.

29. Yudkowsky, E. (2008). Artificial intelligence as a positive and negative factor in global risk. In N. Bostrom, & M.M. Ćirković (Eds.), *Global catastrophic risks* (pp. 308–345). New York: Oxford University Press. <https://doi.org/10.1093/oso/9780198570509.003.0021>

Информация об авторе

Дмитрий Владимирович Попов, кандидат философских наук, доцент, Омская академия МВД России, Омск, Россия, ORCID: <https://orcid.org/0000-0002-4587-6351>, e-mail: DmitriVPopov@mail.ru

Information about the author

Dmitry Vladimirovich Popov, Candidate of Philosophical Sciences, Associate Professor, Omsk Academy of the Ministry of Internal Affairs of the Russian Federation, Omsk, Russia, ORCID: <https://orcid.org/0000-0002-4587-6351>, e-mail: DmitriVPopov@mail.ru