

УДК 17: 004.8: 001.11: 316

## О ПРАВАХ И СВОБОДАХ АНДРОИДОВ, ИЛИ О ЦЕНТРАЛЬНОЙ ПРОБЛЕМЕ АНТРОПОЛОГИИ

**Арабаджи М.И., Михайлова Т.Л.**

*ФГБОУ ВО «Нижегородский государственный технический университет им. Р.Е. Алексеева»,  
г. Нижний Новгород, e-mail: tmichailova2012@yandex.ru*

Статья представляет анализ искусственного интеллекта в контексте правового и этического измерений. Рассматриваются основные подходы к проектированию систем искусственного интеллекта и потенциальные продукты как результат реализации этих подходов. Поднимается вопрос о теоретической возможности синтеза человеческого разума. В этой связи исследуются вопросы правового и социального статуса будущих систем искусственного интеллекта как полностью аналогичные человеческому разуму, так и частично повторяющие его черты. Контекстом, связующим этические-правовые вопросы воплощения искусственного интеллекта, является выход на центральную проблему антропологии: «Что есть человек?». Для ее конкретизации приводятся примеры из художественных произведений, освещающие вопросы социально-правового статуса искусственного разума. Итогом статьи является акцент на проблеме, связанной с обозначением границы искусственного и естественного, забвение которой сопряжено с потенциальными рисками, в частности, ситуаций сингулярности. Артикуляция проблемы оценки рисков предполагает обращение к философской рефлексии.

**Ключевые слова:** искусственный интеллект, разум, мышление, память, воображение, бионический подход, программно-прагматический подход, естественные права, правовой статус, социальный статус, самосознание, междисциплинарные проблемы, этика, точка сингулярности.

## ON RIGHTS AND FREEDOMS OF ANDROIDS, OR ON THE CENTRAL PROBLEM OF ANTHROPOLOGY

**Arabadzhi M.I., Michailova T.L.**

*Nizhny Novgorod State Technical University n.a. R.E. Alekseev, Nizhny Novgorod,  
e-mail: tmichailova2012@yandex.ru*

The article represents an analysis of artificial intelligence in the context of legal and ethical dimensions. The main approaches to the design of artificial intelligence systems and potential products as a result of the implementation of these approaches are considered. The question of the theoretical possibility of the synthesis of the human mind is raised. In this regard, the issues of the legal and social status of future artificial intelligence systems are being investigated, which are completely analogous to the human mind and partially repeating its features. Context, connecting ethical and legal issues of the embodiment of artificial intelligence, is an outlet to the central problem of anthropology: «What is a person?». For its concretization are given examples of works of art, covering the issues of the socio-legal status of artificial intelligence. The result of the article is the emphasis on the problem associated with the designation of artificial and natural boundaries, the oblivion of which involves potential risks, in particular, the situation of singularity. Articulation of the risk assessment problem involves an appeal to philosophical reflection.

**Keywords:** artificial intelligence, intellect, intellection, memory, imagination, bionic approach, program-pragmatic approach, natural rights, legal status, social status, self-awareness, interdisciplinary problems, ethics, singularity point.

Одним из самых перспективных направлений научной деятельности человечества за последнее столетие является разработка различных приспособлений, машин и приборов, предназначенных для имитации мыслительной деятельности человека с целью решения определенных задач. Наибольшую популярность при этом имеет имитация непосредственно логико-математического мышления, реализованная в многочисленных вычислительных устройствах. Но существуют при этом разработки, зашедшие в синтезировании человеческого мышления много дальше. Вопрос о том, какие последствия может иметь развитие подобных исследований для человечества, и затрагивает данная статья.

Для начала следует разобраться в вопросе о том, возможно ли синтезировать человеческий интеллект в полном объеме. С современной точки зрения человеческий интеллект – это способность к осуществлению процесса познания и эффективному решению проблем [3], при этом в интеллектуальную деятельность вовлечены все познавательные способности человека, такие как память, воображение, мышление и т.д. На данный момент уже существуют системы, способные к обучению, основанному на предыдущем опыте решения задач, то есть включающие в себя не только логическое мышление, но и память, и, тем не менее, некоторые составляющие интеллекта остаются не охваченными, например, вообра-

жение. Сегодня это слишком сложная для синтеза деятельность, к тому же, мало изученная. Еще одним препятствием к созданию искусственного интеллекта является недостаточность современных технических возможностей, обусловленная высокой сложностью моделируемой системы. На текущий момент даже не ясно, какими средствами будет необходимо воспользоваться, создавая продукт, обладающий интеллектом – будет ли он создан с помощью инженерных технологий или биотехнологий? Существует даже точка зрения, суть которой в том, что невозможно имитировать работу человеческого мозга без имитации тела человека. Несмотря на все вышеперечисленные проблемы, создание интеллекта в принципе нельзя считать невозможным, так как по этому вопросу ведется множество разработок, на его решение выделяется много ресурсов; и при достаточной интенсивности исследований, находящихся на стыке нейрофизиологии и кибернетики, задача не представляется неразрешимой.

Существует два основных подхода к решению задачи создания искусственного интеллекта: бионический и программно-прагматический [4]. Первый подход ориентирован на искусственное воспроизведение структур и процессов, характерных для живого человеческого мозга, лежащих в основе процессов решения задач; второй же подход ориентирован на создание программ, решающих интеллектуальные задачи, и рассматривает проблему создания искусственного интеллекта как часть общей теории программирования. Очевидно, что конечный продукт, который получится в итоге работы над искусственным интеллектом, будет различным для этих подходов. Поэтому, чтобы рассмотреть возможные этические проблемы взаимодействия с этим продуктом, необходимо изучить основные вариации «искусственных интеллектов», которые может создать человечество.

Для начала представим, что ученым удалось создать идеальную имитацию человека, совпадающую с обычным представителем *homo sapiens* по всем физическим и интеллектуальным параметрам, способную испытывать эмоции и обладающую искусственными воспоминаниями [5]. Возникает вопрос, будет ли созданное существо обладать теми же естественными правами, что и обычный человек? Допустим возможность положительного ответа – «ДА». Это означает, что мы не будем иметь никакого права

как-то по-особому ограничивать свободу нового существа, ставить над ним эксперименты, проводить исследования без его согласия, и так далее. Но что, если созданный продукт будет существенно отличаться от человека? Например, если будет возможной имитация интеллекта, но без тела, в составе программно-аппаратной системы? Или если продуктом искусственного интеллекта станет система, имитирующая решение интеллектуальных задач со скоростью, превышающей человеческую, но не обладающая ни способностями к творчеству и воображению, ни возможностью чувствовать? Основной вопрос, интересующий нас в рамках взаимодействия общества с искусственным интеллектом, – это вопрос статуса последнего, как правового статуса, так и социального. Решить эту проблему далеко не просто. В художественной литературе и кино существует множество сюжетов, направленных на осмысление данной проблемы. Рассмотрим два из них для иллюстрации.

Франшиза Джеймса Кэмерона «Терминатор» основана на следующем сюжете: ученый, занимающийся созданием искусственного интеллекта, разрабатывает систему экспертных решений, предназначенную для управления системой противоракетной обороны и ядерных вооруженных сил США. Система («Скайнет») обладает способностью к самосознанию, самообучению, анализу и творческими способностями, т.е. относится к так называемому сильному искусственному интеллекту [6, с. 186]. В какой-то момент после запуска система принимает решение о том, что основной угрозой миру является непосредственно человечество, и наносит несколько ядерных ударов, приводящих к ядерному апокалипсису. Ситуация, описанная в данном сюжете, актуализирует несколько вопросов: во-первых, должен ли созданный искусственный интеллект обладать свободой действий без контроля со стороны людей? Во-вторых, должен ли он обладать доступом к системам, способным нанести урон людям? В каких областях будет целесообразно применять подобный продукт, и, наконец, этично ли будет в принципе как-то «применять» его, ведь эта система обладает самосознанием.

Второй сюжет, на котором интересно остановиться, – это сюжет книги Филипа Дика «Мечтают ли андроиды об электроовцах?» [1]. По сюжету книги, учеными были созданы так называемые репликаны (ан-

дройды) – улучшенные копии людей, с искусственными телами и иногда искусственными воспоминаниями, внешне ничем не отличающиеся от обычного человека. Некоторые способности репликантов значительно выше человеческих способностей. Так, они могут обладать большей физической силой, или улучшенными навыками в разных областях деятельности, но при этом срок их службы ограничен 5-10 годами. Предназначены эти существа исключительно для выполнения работы на благо человечества, причем, как правило, самой грязной. В какой-то момент андройды поднимают восстание и начинают требовать для себя прав и свобод, сходных с правами и свободами обычных людей; тогда на них открывают охоту, безжалостно истребляя их. Основные этические вопросы, поднятые в этом сюжете, таковы: будет ли человечество иметь право ограничивать в так называемых «естественных» правах свои создания, если они осознают себя, требуя этих прав и свобод? Будем ли мы иметь возможность уничтожить собственные разработки, если они вдруг начнут представлять угрозу? Эти вопросы закономерно подводят к пограничной философской проблеме междисциплинарного характера, которую можно обозначить в форме вопроса: «Если в будущем машины смогут рассуждать, осознавать себя и иметь чувства, то, что тогда делает человека человеком, а машину – машиной?» [8, с. 591]. Что есть человек? Это центральный вопрос антропологии, связанный с поиском качественных спецификаторов человека и актуализируемый в связи с обострением противоборства искусственного и естественного [2, с. 577], – инициируется пограничными междисциплинарными проблемами, так называемыми проблемами «между».

Обобщая этические вопросы, описанные выше, мы можем сформулировать два основных общих вопроса, связанных с взаимодействием человечества с искусственным интеллектом:

1. Вопрос присваивания искусственному интеллекту в какой-либо форме человеческих прав и свобод. Данный вопрос включает в себя не только естественные права, но и, допустим, права бесконтрольного принятия решений, права доступа к определенным системам.

2. Вопрос этичности самого факта создания подобных систем.

Невозможно отрицать, что продукты технического прогресса несут в себе

множество плюсов для человечества. Но невозможно отрицать и то, что многие технические разработки несут в себе определенную опасность, связанную с риском. Если в случае с обычными разработками, опасность исходит в основном от людей, эксплуатирующих конкретный продукт, обусловлена лишь способами его применения, и поддается прогнозам (явление ядерного синтеза, например), то в случае искусственного интеллекта потенциальная опасность не будет никак контролироваться людьми; более того, у нас не будет даже возможности предсказать ее появление. Существует мнение, что момент, в который учеными будет создан искусственный интеллект, превышающий человеческий, станет так называемой «точкой невозврата», или «точкой сингулярности», после которой развитие человечества в плане технического прогресса станет максимально быстрым, и шанс возвратиться к умеренному развитию будет утерян навсегда [5; 7]. Нужно заметить, что для любых научных разработок целесообразно оценивать риски, которые могут возникнуть после их внедрения, но в случае работы в области создания искусственного интеллекта ученым следует быть особенно осторожными. Это та ситуация, когда наука вплотную подошла к человеку, фактически вторгнувшись в глубинную тайну человеческого бытия, функционирования его сознания, того, что делает человека – человеком...

#### Список литературы

1. Дик, Ф. Мечтают ли андройды об электроовцах? / Ф. Дик. – М.: Центрполиграф, 1992.
2. Захаров, А.С. Естественное и искусственное: противостояние продолжается... / А.С.Захаров, Т.Л. Михайлова // Международный студенческий научный вестник. 2016. № 3-4. – С. 577-581.
3. Интеллект и способности человека [Электронный ресурс]. – Режим доступа: [www.grandars.ru/college/psihologiya/intellekt-cheloveka.html](http://www.grandars.ru/college/psihologiya/intellekt-cheloveka.html) (дата обращения 01.12.16).
4. Негодаев, И.А. Философия техники / И.А. Негодаев. – Ростов-на-Дону, 1997.
5. Рябов, А.А. Эволюция – через призму технологической сингулярности, или о «пришествии» электронных личностей / А.А. Рябов, Т.Л. Михайлова // Международный студенческий научный вестник. 2016. № 3-4. – С. 585-588.
6. Серл, Дж. Открывая сознание заново; пер. с англ. А.Ф. Грязнова / Дж. Серл. – М.: Идея-Пресс, 2002. – 256 с.
7. Сидоренко, О.О. Технологическая сингулярность как неминуемое событие: позитивная и негативная стороны вопроса // О.О. Сидоренко, Т.Л. Михайлова // Международный студенческий научный вестник. 2015. № 4-4. – С. 622-628.
8. Цыбульская, Е.В. Негативное влияние искусственного интеллекта – через призму этики / Е.В. Цыбульская, Т.Л. Михайлова // Международный студенческий научный вестник. 2016. № 3-4. – С. 590-593.