

PHILOSOPHICAL SCIENCES

ETHICAL SKETCHES TO THE QUESTION OF ARTIFICIAL INTELLIGENCE

Kozubaev O.,

*Doctor of philosophical science, professor
National Academy of Sciences of the Kyrgyz Republic
Institute of Philosophy named after A. Altmyshbaev
Bishkek c., Chui Avenue, 265a*

Karabukaev K.

*Doctor of Philosophy Sciences
National Academy of Sciences of the Kyrgyz Republic
Institute of Philosophy named after A. Altmyshbaev
Bishkek c., Chui Avenue, 265a*

ЭТИЧЕСКИЕ ЭСКИЗЫ К ВОПРОСУ ОБ ИСКУССТВЕННОМ ИНТЕЛЛЕКТЕ

Козубаев О.,

*доктор философских наук, профессор
Национальная академия наук Кыргызской Республики
Институт философии им. академика А.Алтымышбаева*

Карабукаев К.

*доктор философских наук
Национальная академия наук Кыргызской Республики
Институт философии им. академика А.Алтымышбаева
<https://doi.org/10.5281/zenodo.15536097>*

Abstract

The article examines the place and role of ethics in the field of artificial intelligence and traces the main ethical principles observed in the application of complex technologies and devices. In order to understand the discourse at the current stage of approaches to the ethics of the actions of smart machines, large and small, an attempt was made to specify what is included in this concept and what is its scope in relation to the practical processes of creating, using and regulating artificially capable intelligence systems.

Аннотация

В статье рассматривается вопрос о месте и роли этики в области искусственного интеллекта и прослеживаются основные этические принципы, соблюдаемые при применении сложных технологий и устройств. Для осмысления дискурса на текущем этапе подходов к этике действия умных машин, больших и малых, сделана попытка конкретизировать, что вкладывается в данное понятие и каков его охват применительно к практическим процессам создания, использования и регулирования систем искусственно способного интеллекта.

Keywords: Ethics, artificial intelligence, science, scientific and technological progress, technology, technology, philosophy.

Ключевые слова: Этика, искусственный интеллект, наука, научно-технический прогресс, техника, технологии, философия.

Этика искусственного интеллекта является частью этики науки и технологий, характерной для роботов и других искусственно интеллектуальных устройств, созданных Человеком. Тезис этики искусственного интеллекта подразумевает, прежде всего, моральных вопросов, относящихся специалистам, проектирующим и конструирующим роботы, машины, имеющие возможности интеллектуального действия. Этика технологий занимается вопросами, относительно того, как искусственно умные существа могут быть использованы в обеспечении интересов человека не во вред его интересов. Как бы эти гуманоидные роботы не нанесли непоправимый урон человеческой экзистенции. Так

как, сегодня искусственный интеллект сильно влияет на человечество, день ото дня расширяются горизонты и масштабы применения умных технических систем, то действия их разработчиков с этической точки зрения должны быть прозрачными и надежными в плане благодеяния. Следовательно, для достижения прозрачности так или иначе потребует государственное регулирование для поддержки моральной ответственности человека. Но это не однозначное положение. С одной стороны, это может ограничивать действия создателей искусственного интеллекта, что в свою очередь может привести к замедлению темпов развития инноваций. С другой стороны, очевидно, что государственное регулирование приводит к устойчивой

стабильности, способно в достаточной степени поддерживать инноваций в области искусственного интеллекта. Выше вкратце обозначена моральная ответственность создателей конструкторов и разработчиков гуманоидных машин и систем. Но в немалой степени интересен и загадочен и другая сторона этики новых технологий. Имеется ввиду, какова моральная сторона действий сверхумных роботов или искусственно интеллектуальных компьютеров, которые ведут себя морально или как бы морально, т.е. так называемых искусственных моральных агентов. Здесь уже возникают новые философские вопросы, такие как характеристики субъектности, связанные с концепцией искусственно умных систем. Известный писатель-фантаст Айзек Азимов еще в 50-х годах XX века в своей серии рассказов “Я робот” обозначивал свои три закона – для защиты людей от их собственных созданий. Какие эти законы? Первый закон абсолютно запрещает роботу вредить людям, второй – озадачивает робота подчиняться людям до тех пор, пока это не станет противоречить первому закону, третий закон обязывает робота защищать самого себя, не нарушая первых двух. Какова мораль из сказанного? Автор подчеркивает последствия неадекватного регулирования и управления в мире, движимом так называемыми передовыми технологиями. Как бы там ни было, человек не гарантирован от той опасности, как поломка робота или его парадоксальное непредвиденное поведение [1].

Следовательно, можно предположить, что ни один из установленных Айзеком Азимовым законов не может в достаточной мере предвидет все возможные тревожные обстоятельства. Поэтому в последнее время все более поставляется под сомнение идея о моральной ответственности искусственно умных устройств и систем. Комиссия, созданная в Англии в 2010 году, пересмотрела законы Айзека Азимова, и установила, что искусственный интеллект должен действовать под ответственностью либо его производителей, либо его владельца или оператора [1]. Ученый математик и писатель фантаст Вернор Видж выдвигает гипотезу о том, что может наступить момент, когда компьютеры станут умнее людей, назвав это явление сингулярностью [2]. Он считает, что это может представлять из себя частичную или полную опасность для человека. Отдельные специалисты утверждают, что в результате интеллектуальной сверхактивности, происшедшей в XXI веке искусственный интеллект наделенный способностью самосовершенствоваться может стать настолько умнее людей, что настанет время, когда человек не может помешать антигуманным действиям такой системе. В статье “Этические проблемы в продвинутом искусственном интеллекте” философ Ник Бостром допускает, что искусственный интеллект созданный человеком может достичь вымирания человечества. Автор предполагает, что технический сверхразум может достичь до степени самостоятельной инициативы и планирования. В таком случае он уже может рассматриваться как автономный субъект. Теоретически автор допускает, что сверхинтеллектуальная техника сможет привести практически к любому возможному непредвиденному исходу, в том числе

и вредному для человека. Однако ученый отвергая свои пессимистические предположения высказывает уверенность в том, что искусственный интеллект изначально призван помогать человеку решить многие сложные проблемы, прежде всего в области здравоохранения, в преодолении бедности, в охране окружающей среды, в осуществлении космических проектов, в обеспечении продовольственной достаточности и безопасности и т. д. [3].

Обращение к этике в связи с развитием информатики, автоматизации, а затем и умных машин вполне закономерно – различные автоматизированные и компьютерные системы и цифровые технологии уже в прошлом веке начали оказывать серьезное влияние на все стороны жизнедеятельности человека. При этом скорость внедрения цифровых технологий настолько высока, что общество не всегда успевает дать всестороннюю и обоснованную моральную оценку всем воздействиям и проявлениям техники в различных областях жизни, равно как и сделать однозначный прогноз по перспективам безопасного справедливого применения новейших технологий. Информационно-коммуникационные технологии не только кардинально изменили личное поведение, стиль жизни и межличностные отношения, но и восприятие и понятие самого общества, а также информацию, которая может быть собрана о человеке с его согласия или без него. Технологические изменения также затрагивают повседневные аспекты современного общества, такие как администрирование и управление, электронная коммерция, электронное здравоохранение, электронное образование, удаленная работа и электронное голосование телекоммуникации, общественно-политическая деятельность, права потребителей, интеллектуальная собственность, демократия и т.д. Отдельную значимость приобретают в наше время каналы восприятия окружающей действительности и многочисленные источники информации, которые формируют картину мира. И здесь технические средства играют все более важную роль и в ряде случаев имеют определяющее значение для воспитания нового поколения. Таким образом, вопросы этики теснейшим образом связаны и включены в более глобальную проблематику – изучение социо-и техно-природной системы, изучение человека в окружении технической реальности, процессов технологической революции и глобализации.

Углубляя поиск взаимосвязей этики и технологического развития, неизбежно следует обратиться к техноэтике, которая уже примерно к 80-м годам XX века оформилась в качестве отдельного научного и исследовательского направления, в рамках которого изучались этические аспекты развития технологий, их влияние на человека и общество. Техноэтика развивается в результате слияния целого ряда дисциплин и направлений на стыке философии и прикладной этики, которые фокусируются на науке и различных отраслевых технологиях (биотехнология, нанотехнология, информатика, информационные технологии, компьютерные науки и др.) Термин “техноэтика” был введен аргентинским философом Марио Бунге для обозначения моральной ответственности технологов и ученых в про-

цессе развития цифрового научно-технического прогресса и выявления рационально-обоснованных правил управления наукой и техническим совершенствованием. По его мнению, “технолог должен нести не только техническую, но и моральную ответственность за все, что он проектирует или выполняет: его расчеты должны быть не только оптимально эффективными, но и не вредными, они должны быть полезными, причем не только в краткосрочной, но и в долгосрочной перспективе”. В целом техноэтика рассматривает морально-нравственные аспекты применительно к человеку-специалисту в техносфере. Понятие “техносфера” можно объяснить как “техноцентрический подход – результат взаимодействия между человечеством и природой; антропоцентрический подход – органопроекции человека или объективирования взаимоотношений людей в ходе их жизнедеятельности”[4].

В условиях очередного (современного) витка развития искусственного интеллекта основные дискуссии, а также научные и исследовательские материалы в области этического аспекта данной отрасли в основном концентрируются в тематическом поле, которое условно можно обозначить рамками широкого англоязычного термина “AI Ethics”. Это обусловлено появлением целого ряда уже существующих и совершенно новых морально-нравственных вызовов, дилемм и методик, которые касаются всех этапов жизненного цикла систем искусственного интеллекта, начиная от чисто философских, мировоззренческих проблем развития технологий и заканчивая сугубо практическими, связанными с применением конкретных оборудования. При этом, с учетом масштаба и потенциала воздействия технологий на общество и окружающую среду данный дискурс затрагивает практически все значимые для человека сферы – политику и социальное развитие, экономику, науку, здравоохранение образование, вопросы права и свобод граждан, экологию, защиту интеллектуальной собственности и др.

В философской литературе можно обнаружить множество самых различных попыток объяснить и значение самого термина этика, не говоря уже о разнообразии теоретических концепций и постулатов, связанных с этой дисциплиной. Очевидно, что этика, как сложное, многозначное понятие включает в себе множество пластов и может рассматриваться с различных ракурсов, использовать различные методы познания. В мире продолжается процесс поиска этических норм или принципов, как в качестве теоретической базы, так и для профильного/прикладного применения. Несмотря на большое количество формализованных этических принципов, разработчики и пользователи систем искусственного интеллекта сталкиваются с необходимостью поиска практических механизмов и решений, соответствующих ожиданиям общества. Так, возникают ситуации, когда различные принципы входят в конфликт и противоречия друг с другом, порой вызывая сложные этические дилеммы. Этические дилеммы относятся к ситуациям, в которых любой доступный выбор ведет к нарушению какого-либо принятого этического принципа, но решение все же должно быть найдено. В этой

связи к разрешению этических дилемм может быть применен подход, сочетающий в себе несколько обозначенных выше или других теорий и учитывающий конкретные обстоятельства использования систем искусственного интеллекта. При этом определяющим фактором становится идентификация рисков и всесторонняя оценка возможного ущерба для всех участников или активов, задействованных в том или ином практическом моменте. В данном случае особое значение приобретают вопросы обеспечения справедливости.

Вопросы справедливости в области искусственного интеллекта поднимают весь спектр философских, правовых и научно-практических подходов к данному понятию. В частности, в широком смысле справедливость может быть рассмотрена прежде всего как принцип права. С такого ракурса справедливость – это “важнейшая категория классической юриспруденции, прежде всего, естественного права. Попытки ее содержательной универсализации не могут быть признаны убедительными в силу множественности социальных и личностных идентичностей”[5, с.35-50]. В целом же, на практике становится очевидным, что проблемы справедливости на различных этапах жизненного цикла систем искусственного интеллекта могут быть успешно решены лишь применительно к контекстуальным условиям, то есть применительно к анализу конкретных обстоятельств использования искусственного интеллекта для решения конкретных задач.

Этические дилеммы и вопросы морального выбора продолжают быть предметом исследований философов, они вызывают научную и профессиональную полемику, особенно в парадигме активной научно-технической и социальной модернизации, в условиях поиска новых смыслов существования для современного социума. Искусственный интеллект, как технология, способная на многих сегментах деятельности заменить человека и вырабатывать рекомендательные решения на основе большого массива данных, неизбежно будет сталкиваться с необходимостью этического сопровождения и моральной оценки рисков.

Вопрос этической актуальности проблем искусственного интеллекта еще более обостряется в последние годы в связи с возникновением нанотехнологий, когда научно-изобретательские способности человека распространяются на атомарно-молекулярный уровень. Объект наноисследований – это материалы в масштабе атомов и молекул, в переводе с греческого “нано” в прямом смысле означает “очень маленький”. Одним словом, “нанотехнология – это технология, оперирующая величинами, сопоставимыми с размерами атомов, переход от широко освоенных ныне микротехнологий к нанотехнологиям. Это качественный скачок от манипуляции веществом к манипуляции отдельными молекулами и атомами” [6]. Получается, что нанотехнология работает не с веществом, а его частицами – атомами.

Нанотехнология представляет собой относительно новую область науки. Она возникла на стыке физики, химии, биологии и технических наук. С ней связаны ожидания в самых различных

областях науки и техники. Нанотехнологии рассматриваются как основа новейшей промышленной революции. Она рассматривалась сначала исключительно с точки зрения позитивной стороны научно-технического прогресса. Однако постепенно появились опасения по поводу нанотехнологических рисков, например о токсичности наночастиц. В результате возникла необходимость этической рефлексии к возможным нежелательным последствиям применения нанотехнологий.

Начало эры учения о наноструктурах следует отнести к 1959 году, когда физик, лауреат Нобелевской премии Ричард Фейнман вознес идею о возможных путях развития физики классической и квантовой, и миникомпьютерах, субмикроскопических исследованиях на уровне атомов. Тогда он заявлял о возможности получения любой вещи манипулируя атомами. Сам термин нанотехнологии введен в 1974г. Японским ученым физиком Норио Танигучи для обозначения искусственно изобретенных изделий из атомов и молекул. Таким образом, нанотехнология – область фундаментальной и прикладной науки и техники на основе анализа и синтеза для производства и применения продукты путем манипулирования отдельными атомами и молекулами [7].

Резюмируя, можно сказать, что нанотехнология – одна из высокотехнологичных отраслей современной науки и техники, которая занимается исследованием атомов и молекул и созданием из них различного рода искусственных изделий. Достижения в этой области ведут к революции в медицине, электронике, искусственном интеллекте, промышленности и в других сферах человеческой деятельности. Другими словами, нанотехнология – это путь к созданию новой цивилизации с присущими ей набором ценностей и идеалов[8, с.59].

А на самом деле насколько безграничны интеллектуально-технические возможности современных компьютеров и роботов? Как свидетельствует Российский ученый физик А. А. Боровой в одном из своих интервью корреспонденту радио «Спутник», в начальный период ликвидации последствий аварии Чернобыльской АЭС в 1986г., из-за сверхвысокого уровня радиоактивности даже самые передовые роботы того времени, по образному выражению ученого «в буквальном смысле, сходили с ума», т.е. попросту говоря они потеряли свою адекватную работоспособность.

Вполне уместно, что в современной действительности все более значимым становится философская рефлексия социокультурных последствий развития нанотехнологий, выявления изменения социальных ценностей и смысла человеческой жизни. Поэтому важен анализ социокультурных последствий развития нанотехнологий, так как их познавательная мощь не только обладает положительным потенциалом, но и угрожает существованию человечества. В условиях достаточно милитаризованной техногенной цивилизации. В целом, нанотехнологии дают возможность смены технологических парадигм индустриальной эпохи и общей виртуализации социума, выделяющейся в изменении ценностных приоритетов, компьютеризации и

цифровизации всех сфер общества, вплоть до создания искусственного суперинтеллекта.

Согласно известным строкам А. С. Пушкина человеческий гений есть «парадоксов друг». А одним из причин возникновения не всегда желательных парадоксов как раз и является столкновение человека с социально-этическими и гуманистическими проблемами бытия. Почему и в силу каких причин эта проблема приобрела в прошлом и приобретает сейчас такое значение. Сегодня как подчеркивает И.Т.Фролов, чтобы дать ответ этому вопросу необходимо обратиться к тем социальным условиям, в которых происходит процесс функционирования и развития науки от которых существенно зависит. И тогда станет ясно, что человечество никогда не переживало столь ответственного и во многом определяющего дальнейшую будущность мировой цивилизации этапа, как сейчас. Историческое развитие подвело человека к такому качественному рубежу, когда НТП, рост влияния человека на природные процессы, реальная перспектива исчерпания не возобновляемых источников энергии, естественных ресурсов и ряд других факторов лишает человечество права на фатальную ошибку.

В наши дни масштабность и острота неблагоприятных последствий любых ошибочных решений, где-бы оно не принято, обретают глобальный характер. И любые расчеты, игнорирующие это - аморальны, антигуманны и бесчеловечны.[9, с.4-5]. Следовательно, вопрос нанотехнологий не просто вопрос технического прогресса, он имеет и этическое измерение. «Будучи новой отраслью научно-технического и одновременно фундаментального исследования с широкой сферой технического применения, нанотехнология как явление привлекает к себе не только инженеров, биологов, специалистов сугубо технического профиля, но и философов, порождая новые области философского исследования, - одной из которых является нанотехнологическая этика [10].

Непрерывный рост научно-технического прогресса всегда свидетельствует о том, что чуть ли не все полезные изобретения и научные разработки не только способствуют развитию и прогрессу, прежде всего в экономике, но и ставят человечество перед новыми, подчас трудно прогнозируемыми опасностями и рисками.

Таким образом, несмотря на то, что нанотехнологии находятся на этапе становления, уже сейчас их можно отнести к остроактуальным вопросам будущего наряду с вопросами о демографических или климатических изменениях. Помимо бесспорных положительных возможностей, сопутствующих прогрессу в промышленности, сельском хозяйстве, медицине, энергетике, охраны окружающей среды и других сферах человеческой деятельности, нарастает и множество опасностей. Поэтому этическая оценка достижений науки всегда остается актуальной и необходимой. Человек как разумное существо при творении любого нового механизма, в том числе и нанотехнологий выступает как творец, как хозяин и смотрит в будущее с техническим оптимизмом, сформулировав на будущее новые претензии в уверенности на господство. Однако проблема

заключается в том, что всякие новые научные разработки, в том числе и нанотехнологические в основном носят прагматический характер, где морально-этические максимы иногда отодвигаются на второй план. Например, молекулярная нанотехнология (МНТ) содержит в себе определенные угрозы для человеческой безопасности. О. В. Летов интересно трактует как прямых, так и косвенных угроз МНТ. По его мнению, прямая угроза связана, прежде всего, с процессом гонки вооружений. Использование МНТ в неблагоприятных целях способно нарушить мирное сосуществование. МНТ в принципе позволяет создавать «интеллектуальное» оружие массового уничтожения, такое как, например смертельный вирус. Производителями подобного оружия могут быть не только государства, но и группы людей [11]. Следовательно, не уделяя достаточного внимания этике нанотехнологий, невозможно гарантировать эффективного и гармоничного развития социума, достичь оптимистических результатов, предотвратить возможный вред, наносимый человеку, уменьшить риск нежелательных экономических последствий. Нанотехнологию следует развивать с учетом безопасности и с чувством ответственности. Это значит соблюдать такой этический принцип как «не навреди» с учетом рисков для здоровья, окружающей среды и безопасности. В рамках общей безопасности особое место занимают проблемы нравственной безопасности при создании и использовании нанотехнологий. На сегодня она явно очевидно в военной, медицинской, экологической, интеллектуальной областях. К сожалению, из мировой практики известно, что исследования по безопасности нанотехнологий существенно отстают от их разработки и практической реализации.

Технологии, когда бы они не создавались на каждом конкретном историческом отрезке времени имеют прежде всего прогрессивную направленность. Однако, как отметил Ф. Энгельс каждая вещь и каждый процесс сами по себе внутренне противоречивы [12, с.579-580]. В научно-техническом прогрессе также каждый новый этап первоначально заявляет только одни преимущества. Но затем вторые и третьи производные, т.е. отдельные конкретные последствия, могут оказаться негативными. Не надо забывать, что «в любом проявляющемся как добро явлении зло заключается уже в том, что при всей положительности оно таит в себе момент превращения в зло» [13, с.81]. Это можно иллюстрировать на примере освоения ядерной энергетики или применения геномодифицированных элементов в пищевом производстве. Значит, в процессе внедрения нанотехнологий и наноматериалов принцип предосторожности должен занимать не последнее место.

Вышеизложенное дает возможность резюмировать, что этические вопросы науки и техники постоянно сопровождаются с новыми вызовами развития общества. В связи с широкой междисциплинарностью и междисциплинарностью современной

науки на уровне нанонауки и нанотехнологий, этика науки и искусственного интеллекта охватывает достаточно широкий круг проблем и многие ее вопросы неотрывно связаны с иными контекстами этической рефлексии. Потому что, этический аспект науки тесно связан «как с развитием самой науки, так и с её взаимодействием с обществом. Так или иначе трактовка науки в значительной степени и определяет характер постановки её этических вопросов» [9]. Следовательно, все вопросы научно-технической этики крайне важны. Это связано с тем, что динамика достижений науки и техники год от года возрастает. В связи с этим актуализируется важность своевременной адекватной реакции на возможные негативные последствия научно-технического прогресса.

Список литературы:

1. А.Азимов Я, робот - Эксмо, 2019 – 250с. Электронный ресурс: <https://www.litres.ru/book/ayzek-azimov/ya-robot-435545/> Дата обращения: 25апреля 2025
2. Винджи В. Сингулярность перевод М.Левин, А.Гришечкин – АСТ, 2022 - 224с.
3. Бостром Н. Этические проблемы, связанные с искусственным интеллектом. Bostrom N.ethical issues in advanced artificial intelligence // cognitive, emotive and ethical aspects of decision making in humans and in artificial intelligence. - Vol. 2: / ed. I. Smit et al.. - Oxford U. k, 2003. - P. 12-17
4. Попкова Н.В. Техносфера как объект философского исследования / дис. ... докт. философ. наук. - М., 2008.- 420с. Режим доступа: [http:// www. Dissercat.com/content/](http://www.Dissercat.com/content/).
5. Честнов И.Л. Концепт справедливости в постклассическом правопонимании. // Журнал "Известия ВУЗов. Правоведение", 2013 год, № 2 (307) -С. 35 - 50.
6. Шамов И.А. Какие этические проблемы несут в себе нанотехнологии. Медицинская этика и деонтология- Издательство: НИЦ ИНФРА - М.: 2024 - 288с.
7. Wikipedia org/wiri/Нанотехнология.
8. Аршинов В.И., Лебедев М.В. Философские проблемы развития и применение нанотехнологий // Вопросы философии. –М.: 2008 №1. – С.59
9. Фролов И.Т., Юдин Б.Г. Этика науки. Проблемы и дискуссии. - М.: Политиздат, 1986. - 399 с
10. Грунвальд А. Этика для нанотехнологий // Электронный ресурс: https://iphrrs.ru/uplfile/root/biblio/ps/ps20_1/126-143.pdf
11. Летов О.В. Нанотехнологии и этика. (Аналитический обзор). Электронный ресурс: <https://cyberleninka.ru/article/n/2008-02-020-letov-o-v-nanotehnologii-i-etika-analiticheskiy-obzor> Дата обращения: 12мая 2025
12. Маркс К., Энгельс Ф. Соч. Т.20. - С.579-580
13. Диалектика и этика. – Алма -Ата: Наука, 1983. - 303с.