

УДК 336.22

Гибизова Юлия Романовна, студент Российского экономического университета им. Г. В. Плеханова, г. Москва, Российская Федерация

Григорьян Анна Эрастовна, студент Российского экономического университета им. Г. В. Плеханова, г. Москва, Российская Федерация

Научный руководитель: **Лозовицкая Галина Петровна**, доктор юридических наук, профессор кафедры государственно-правовых и уголовно-правовых дисциплин РЭУ им. Г. В. Плеханова, г. Москва, Российская Федерация

E-mail: Lozovitskaya.GP@rea.ru

ЭТИКА ИННОВАЦИЙ: ПРИНЦИП «НЕ НАВРЕДИ» В РАЗРАБОТКЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА И БИОТЕХНОЛОГИЙ; МОРАТОРИЙ НА ТЕХНОЛОГИИ, УГРОЖАЮЩИЕ ЧЕЛОВЕЧНОСТИ

Аннотация. В статье рассматривается одна из наиболее актуальных задач современного административного права – обеспечение этических основ разработки и применения технологий искусственного интеллекта и биоинженерии. Через призму принципа «не навреди» анализируются пределы допустимого вмешательства технологий в жизнь общества, исследуются формы ответственности ученых за последствия их научной деятельности, а также рассматривается возможность введения временных запретов на технологии, способные нанести необратимый ущерб личности.

Ключевые слова: этика инноваций, искусственный интеллект, биоинженерия, принцип «не навреди», ответственность ученого, мораторий, административное право, цифровая экономика, регулирование технологий.

Yulia Romanovna Gibizova, student, Plekhanov Russian University of Economics, Moscow, Russian Federation

Anna Erastovna Grigoryan, student, Plekhanov Russian University of Economics, Moscow, Russian Federation

Academic Supervisor: **Galina Petrovna Lozovitskaya**, Doctor of Law, Professor, Department of State, Legal, and Criminal Law Disciplines, Plekhanov Russian University of Economics, Moscow, Russian Federation

Email: Lozovitskaya.GP@rea.ru

INNOVATION ETHICS: THE "DO NO HARM" PRINCIPLE IN THE DEVELOPMENT OF ARTIFICIAL INTELLIGENCE AND BIOTECHNOLOGY; MORATORIUM ON TECHNOLOGIES THAT THREATEN HUMANITY

Abstract. The article examines one of the most pressing issues in modern administrative law: ensuring the ethical foundations for the development and application of artificial intelligence and bioengineering technologies. Through the lens of the «do no harm» principle, the article analyzes the limits of acceptable technological intervention in society, explores the forms of responsibility for the consequences of scientific activities, and examines the possibility of imposing temporary bans on technologies that could cause irreversible harm to individuals.

Keywords: ethics of innovation, artificial intelligence, bioengineering, the «do no harm» principle, the responsibility of a scientist, moratorium, administrative law, digital economy, and technology regulation.

Введение

Технологический прогресс XXI века неизбежно ставит перед государством, обществом и научным сообществом один и тот же фундаментальный вопрос: где проходит граница между допустимым развитием и угрозой, которую само это развитие создает? Взрывной рост применения систем искусственного интеллекта (далее – ИИ) и стремительное расширение биоинженерных возможностей – от редактирования генома до нейроинтерфейсов – выдвигают на передний план проблему этического и правового регулирования, которая в условиях цифровой экономики приобретает подлинно государственное измерение.

Именно государство, как публичный субъект, несет ответственность за установление правил, ограничений и процедур контроля в сфере разработки и внедрения высоких технологий. Особую значимость в этом контексте приобретает вопрос о применимости принципа «не навреди» – заимствованного из медицинской этики, но все активнее обсуждаемого применительно к технологическим инновациям – в качестве универсального стандарта при создании систем ИИ и биотехнологий. Параллельно встает вопрос о персональной ответственности ученого за социальные последствия его открытий и о возможности введения временных запретов на технологии, несущие угрозу

самой человечности. Именно взаимосвязь этих трех вопросов – принципа безопасности, ответственности и моратория – составляет предмет настоящего исследования.

Цели и задачи исследования

Целью настоящей статьи является комплексный анализ этико-правовых механизмов регулирования инновационной деятельности в сферах ИИ и биоинженерии с позиций административного права, а также обоснование необходимости системного закрепления принципа «не навреди» в нормативных документах государственного управления.

Для достижения указанной цели в работе последовательно решаются следующие задачи: рассмотреть историю формирования этических требований к технологическим инновациям и выявить ключевые точки, в которых наука была вынуждена столкнуться с вопросом об ответственности; раскрыть содержание принципа «не навреди» применительно к ИИ и биотехнологиям, опираясь на конкретные примеры; исследовать формы ответственности ученых за социальные последствия научной деятельности; проанализировать концепцию моратория как самостоятельного правового инструмента. Каждая из этих задач вытекает из предыдущей: понять, как этика технологий складывалась исторически, – значит получить основание для анализа ее современного содержания; а понять содержание – значит получить основание для разговора о механизмах ее обеспечения.

История проблемы: от медицинской этики к этике технологий

Принцип «не навреди» восходит к традиции Гиппократов и на протяжении столетий являлся моральным фундаментом медицинской практики. Однако в условиях технологической революции он вышел далеко за пределы медицины и превратился в центральный предмет дискуссий об ответственном применении новых разработок. Переломным моментом стала Асиломарская конференция 1975 года, когда ученые впервые добровольно объявили временный запрет на ряд экспериментов с рекомбинантной ДНК. С развитием информационных

технологий и формированием цифровой экономики вопрос об этике инноваций приобрел принципиально иное измерение – коммерческое давление на науку резко возросло. Как отмечают И. Ю. Еланский и В. Г. Гриб, «в условиях цифровой экономики интеллектуальная собственность становится главным инструментом для извлечения прибыли» [3, с. 47], что порождает мощные стимулы к ускоренному внедрению технологий в ущерб их надлежащей проверке. Это наблюдение важно не только само по себе, но и как контекст для понимания того, почему добровольные этические обязательства ученых нередко оказываются недостаточными и уступают место рыночной логике.

В России правовую основу для регулирования применения ИИ образует Федеральный закон от 31 июля 2020 года № 258-ФЗ «Об экспериментальных правовых режимах в сфере цифровых инноваций» [8], предусматривающий апробацию инновационных решений в тестовом режиме. Федеральный закон от 8 июля 2024 года № 169-ФЗ [7] дополнил этот механизм нормой о рассмотрении случаев причинения вреда при использовании технологий ИИ. Тем не менее действующее регулирование носит ситуационный характер: оно формирует отдельные правила для конкретных сценариев, не создавая единой этической концепции. Именно этот пробел и предстоит преодолеть.

Принцип «не навреди» как этико-правовой стандарт в сфере ИИ и биоинженерии

Закрепление принципа «не навреди» в качестве административно-правового стандарта предполагает прежде всего ответ на вопрос о его конкретном нормативном содержании. Обязанность разработчиков, государственных органов и научных организаций предвидеть и предотвращать негативные последствия еще на этапе проектирования систем – это не просто этическое пожелание, а требование, которое должно быть зафиксировано в регламентах допуска технологий к применению. В. А. Леонов, Е. В. Каштанова и А. С. Лобачева констатируют, что «этика должна быть встроена в идею о том, почему разрабатывается определенная технология, оснащенная ИИ» [5, с. 7],

поскольку никакая другая сфера не оказывает столь разрушительного влияния на привычный уклад общества. Это требование звучит просто, однако его реализация предполагает серьезную реорганизацию всего процесса разработки – от постановки задачи до выпуска продукта.

Показательным примером того, как отсутствие предварительной этической проверки приводит к конкретному вреду, служит история рекрутинговой системы компании Amazon. Алгоритм, обученный на исторических данных, отражавших гендерный дисбаланс в технологической отрасли, стал занижать оценки резюме женщин-кандидатов – инструмент, созданный для объективного отбора, воспроизводил и усиливал существующее неравенство. Этот случай стал хрестоматийным примером того, как добросовестное намерение не гарантирует безвредного результата. Следует подчеркнуть, что действующий Кодекс Российской Федерации об административных правонарушениях [4] не содержит специальных норм, предусматривающих ответственность за внедрение алгоритмических систем с дискриминационными последствиями. Данный пробел представляется существенным: его восполнение потребует введения самостоятельного состава, устанавливающего ответственность за допуск к применению систем ИИ без предварительной проверки на предвзятость в сферах трудоустройства, кредитования и государственного управления.

Переходя от ИИ к биоинженерии, важно осознать, что масштаб и природа рисков здесь принципиально иные. П. А. Сиваев указывает, что «одними из главных рисков разработки и применения технологий биоинженерии на основе искусственного интеллекта являются неэкологичность и неэтичность эксперимента, размытие границ реальности у потребителей технологий, киберпсихоз, кража данных и манипулирование ими злоумышленниками» [6, с. 127]. Биоинженерные риски нелинейны, накапливаются постепенно и могут проявиться лишь через несколько поколений, нередко нося необратимый характер. Редактирование генома по технологии CRISPR/Cas9 способно

устранить наследственное заболевание, однако непредвиденные последствия такого вмешательства могут сказаться на отдаленных потомках. По этой причине долгосрочный анализ рисков должен быть закреплен как обязательное условие проведения клинически значимых исследований, а не оставаться рекомендацией.

Особого рассмотрения заслуживают нейроинтерфейсы – технологии вживления чипов в мозг человека [9]. Изначально они создавались для восстановления двигательных функций парализованных пациентов, однако зафиксированные случаи потери чувства самоидентичности и попытки самоубийства среди испытуемых показали: вмешательство в деятельность мозга способно повлечь глубокий кризис личности. КоАП РФ в настоящее время не содержит норм, регулирующих проведение подобных экспериментов с точки зрения административной ответственности организаций-разработчиков. Между тем, именно здесь необходимо предусмотреть специальный контрольно-надзорный механизм: обязательное лицензирование нейротехнологических исследований, административную ответственность за проведение испытаний без надлежащего информированного согласия участников и за сокрытие сведений о зафиксированных побочных эффектах.

Ответственность ученых за социальные последствия открытий

Проблема ответственности ученого находится на пересечении профессиональной этики и административно-правового регулирования. В традиционной парадигме наука мыслилась как ценностно нейтральная деятельность, а ответственность за применение ее результатов возлагалась на пользователей. Однако эта позиция была поставлена под серьезное сомнение в условиях, когда сами ученые становятся предпринимателями и соучредителями технологических компаний, монетизирующих результаты исследований. Там, где наука и бизнес неразличимы, нейтральность ученого оказывается фикцией.

Для философского осмысления этой трансформации обратимся к этике ответственности Ганса Йонаса. М. Х. Бадмаева, исследуя его концепцию, формулирует ключевое требование технологической этики: «поступай так,

чтобы последствия твоих действий были совместимы с непрерывностью подлинной человеческой жизни на Земле» [1, с. 74]. Данный императив означает, что ученый обязан учитывать не только ближайшие, но и отдаленные последствия своих исследований – в том числе для будущих поколений, которые не имеют возможности выразить согласие или несогласие. Страх перед возможными катастрофическими последствиями, по Йонасу, должен быть не парализующим, а мобилизующим: он побуждает к большей осторожности при принятии решений.

Философский анализ позволяет перейти к конкретному вопросу о том, как именно в современных условиях деформируется независимость научного суждения. П. А. Сиваев показывает: там, где ученый выполняет функции наемного работника корпорации, оценка достижений строится на коммерческой полезности для организации, а научная истина превращается в «сопутствующий момент» на пути к экономической выгоде. Это порождает конфликт интересов, требующий нормативного урегулирования – обязательного раскрытия аффилированности исследователей, создания независимых этических комитетов с реальными полномочиями по приостановлению работ, а также установления личной ответственности руководителей научных проектов.

С позиций административного права ответственность реализуется в нескольких взаимосвязанных формах. Персональная ответственность предполагает санкции за нарушение установленных регламентов безопасности. Корпоративная – обязанность организации обеспечить этическую экспертизу продуктов до их выпуска. Институциональная ответственность государственных органов заключается в надлежащем контроле за выдачей разрешений на проведение исследований и их практическое применение. Без взаимодействия всех трех уровней система контроля неизбежно окажется неполной. При этом КоАП РФ на сегодняшний день не предусматривает специального состава, связанного с нарушением требований этической экспертизы в сфере ИИ и биоинженерии, – санкция носит скорее гражданско-правовой или уголовно-

правовой характер лишь в очевидных случаях. Систематическая административная превенция в данной сфере фактически отсутствует, что делает ее формирование самостоятельной законодательной задачей.

Отечественная правоприменительная практика тем не менее делает первые шаги в нужном направлении. Верховный Суд Российской Федерации разъяснил, что создание агитационных материалов с использованием вводящих в заблуждение изображений, сгенерированных нейросетью, влечет административную ответственность по ч. 1 ст. 5.12 КоАП РФ [2]. Это свидетельствует о том, что суды готовы рассматривать применение ИИ как самостоятельное основание для привлечения к ответственности. Однако такой подход остается реактивным – он фиксирует уже свершившиеся нарушения, не создавая превентивных механизмов. Именно это противоречие между реакцией и предупреждением подводит к вопросу о более решительном инструменте – моратории.

Мораторий как инструмент административно-правовой защиты человечности

Концепция моратория на опасные технологии – временного запрета на их разработку или применение до накопления достаточных данных о безопасности – имеет весомое историческое обоснование. Асиломарская конференция 1975 года наглядно показала, что мораторий может быть не только принудительным актом государства, но и формой ответственного самоограничения научного сообщества. Ключевым здесь является слово «временный»: мораторий не отрицает развитие, он создает пространство для того, чтобы это развитие не опередило нашу способность понять его последствия.

Сегодня вопрос о введении моратория наиболее остро стоит в двух направлениях. В биоинженерии – прежде всего применительно к редактированию генома в зародышевой линии человека, то есть к вмешательству, последствия которого передаются следующим поколениям. В сфере ИИ – применительно к летальным автономным системам вооружений,

способным принимать решения о применении силы без участия человека. Оба случая объединяет одна черта: речь идет о технологиях, способных необратимо изменить фундаментальные условия человеческого существования – биологические или политические.

Административно-правовое обоснование моратория опирается на принцип предосторожности, признанный в ряде международных документов: при наличии серьезной и научно обоснованной угрозы государство вправе принимать защитные меры даже в отсутствие полной научной определенности относительно степени риска. Данный принцип уже нашел отражение в Регламенте ЕС по ИИ, запрещающем системы социального скоринга и ряд приложений, признанных недопустимыми по своей природе. Принцип предосторожности – это не антинаучный консерватизм, а признание того, что некоторые ошибки исправить невозможно. В целом, на необходимость наведения порядка в сфере применения ИИ обращают внимание многие ученые [10,11,12].

Введение моратория не означает полного запрета научного поиска. Его функция состоит в создании временного пространства для формирования регуляторной базы, проведения независимых оценок рисков и достижения общественного консенсуса по ключевым этическим вопросам. Технология, которой не доверяют, обречена на отторжение – и мораторий как раз создает условия для того, чтобы это доверие было заработано, а не взято авансом. В российском контексте данный механизм мог бы быть реализован через расширение инструментария экспериментальных правовых режимов, предусмотренных Федеральным законом № 258-ФЗ. Наряду с режимом «тестирования» целесообразно предусмотреть режим «приостановления» – для технологий, в отношении которых накоплены данные о существенных рисках, однако окончательная научная оценка еще не завершена. При этом действующий КоАП РФ не содержит нормы, прямо обязывающей участников экспериментальных правовых режимов прекращать применение технологии при

выявлении признаков системного вреда до завершения административного расследования. Восполнение этого пробела – в частности, путем введения обязанности уведомления надзорного органа и временной приостановки деятельности по его требованию с установлением административной санкции за уклонение – представляется одним из приоритетных направлений развития законодательства в данной сфере.

Выводы

Проведенное исследование позволяет констатировать, что принцип «не навреди», ответственность ученых за социальные последствия открытий и концепция моратория образуют взаимосвязанную систему этико-правовых ориентиров, которая нуждается в последовательном воплощении в нормах административного права. Сложившаяся в России ситуационная модель регулирования ИИ должна быть дополнена системными механизмами: обязательной предварительной этической экспертизой высокорисковых технологий, институтом персональной и корпоративной ответственности за их последствия, а также процедурой временного приостановления применения технологий, создающих угрозу необратимых изменений в биологической или социальной природе человека. Действующий КоАП РФ не содержит специальных составов, охватывающих эти сферы, – что само по себе является весомым аргументом в пользу разработки самостоятельного законодательного блока, регулирующего административную ответственность в сфере этики технологических инноваций.

Литература

1. Бадмаева М.Х. Этика искусственного интеллекта: принцип ответственности Ганса Йонаса // Вестник Бурятского государственного университета. Философия. 2022. Вып. 1. С. 67–79.
2. Кодекс Российской Федерации об административных правонарушениях от 30.12.2001 № 195-ФЗ (ред. от 28.01.2025) // Собрание законодательства РФ. 2002. № 1 (ч. 1). Ст. 1.

3. Еланский И.Ю., Гриб В.Г. Правовое регулирование инноваций в условиях цифровой экономики // Юридическая наука. 2023. № 11. С. 47–49.
4. Леонов В.А., Каштанова Е.В., Лобачева А.С. Этика искусственного интеллекта: проблемы и инициативы в социальной сфере // Управление персоналом и интеллектуальными ресурсами в России. 2021. № 2 (53). С. 5–12.
5. Сиваев П.А. Искусственный интеллект и биоинженерия: этические проблемы взаимодействия // Гуманитарные ведомости ТГПУ им. Л. Н. Толстого. 2024. № 2 (50). С. 120–129.
6. Федеральный закон от 31.07.2020 № 258-ФЗ «Об экспериментальных правовых режимах в сфере цифровых инноваций в Российской Федерации» // Собрание законодательства РФ. 2020. № 31. Ст. 5017.
7. Федеральный закон от 08.07.2024 № 169-ФЗ «О внесении изменений в Федеральный закон «Об экспериментальных правовых режимах в сфере цифровых инноваций в Российской Федерации»» // Официальный интернет-портал правовой информации. URL: <http://pravo.gov.ru> (дата обращения: 20.02.2025).
8. Верховный Суд Российской Федерации. Разъяснение об административной ответственности за использование нейросетей при создании агитационных материалов (2025) // Гарант-Уфа. URL: <https://garant-ufa.ru> (дата обращения: 20.02.2025).
9. Лозовицкая, Г. П. О некоторых аспектах использования искусственного интеллекта и вживления микрочипов в мозг человека / Г. П. Лозовицкая, Н. Н. Лапупина // Вестник Восточно-Сибирской Открытой Академии. – 2023. – № 49(49). – EDN SDZCYT.
10. Лозовицкая, Г. П. Безопасность детей и подростков в киберпространстве / Г. П. Лозовицкая // Теория и практика судебной экспертизы: международный опыт, проблемы, перспективы (к 20-летию образования Московского университета МВД России имени В.Я. Кикотя) : Сборник научных трудов Международного форума, Москва, 25 марта 2022 года / Сост. В.В. Бушуев. –

Москва: Московский университет Министерства внутренних дел Российской Федерации им. В.Я. Кикотя, 2022. – С. 194-202. – EDN QWFSZZ.

11. Лозовицкая, Г. П. Искусственный интеллект: правовые проблемы использования / Г. П. Лозовицкая, Л. В. Маринкина // Стратегическое развитие системы МВД России: состояние, тенденции, перспективы : Сборник статей Международной научно-практической конференции, Москва, 23 октября 2020 года / Под общей редакцией И. Г. Чистобородова, А. Л. Ситковского, В. О. Лапина. – Москва: Академия управления Министерства внутренних дел Российской Федерации, 2020. – С. 517-521. – EDN JOAWVF.

12. Socio-Psychological Practices and the Quality of Education: Reconstruction of Traditional Technologies in the Context of Digitalization / E. E. Rumyantseva, S. V. Nedbayeva, D. N. Nedbayev [et al.] // II International Scientific Forum on Sustainable Development and Innovation (WFSDI 2023) : Conference Proceedings, Porto, Portugal, 27–28 апреля 2023 года. – Ekaterinburg: Institute of Digital Economics and Law LLC, 2024. – P. 157-161. – EDN GRTMCL.