



Известия высших учебных заведений. Прикладная нелинейная динамика. 2026. Т. 34, № 1
Izvestiya Vysshikh Uchebnykh Zavedeniy. Applied Nonlinear Dynamics. 2026;34(1)

Обзорная статья
УДК 004.8

DOI: 10.18500/0869-6632-003196
EDN: JXIIOD

Козволюция нейротехнологий и ИИ: этические вызовы и подходы к регуляции

А. В. Шендерюк-Жидков^{1,2}, *В. А. Максименко*³, *А. Е. Храмов*^{4,5} ✉

¹Совет Федерации Федерального Собрания Российской Федерации, Москва, Россия

²Балтийский федеральный университет им. И. Канта, Калининград, Россия

³Автономная некоммерческая организация «Неймарк», Нижний Новгород, Россия

⁴Российский экономический университет имени Г. В. Плеханова, Москва, Россия

⁵ФГАУ «Цифровые промышленные технологии», Москва, Россия

E-mail: a.shenderyuk@yandex.ru, maximenkov1@gmail.com, ✉hramov.ae@rea.ru

Поступила в редакцию 8.07.2025, принята к публикации 1.08.2025,

опубликована онлайн 12.09.2025, опубликована 30.01.2026

Аннотация. Цель настоящего исследования — проанализировать этические вызовы, возникающие на стыке нейротехнологий и искусственного интеллекта (ИИ), а также предложить подходы к их регулированию, обеспечивающие ответственное развитие этих технологий. Особое внимание уделено вопросам автономии личности, конфиденциальности данных, социальной справедливости и предотвращению манипуляций сознанием. **Методы.** В работе использован междисциплинарный подход, включающий анализ научной литературы, нормативных документов и позиций религиозных институтов. Проведено сравнение рисков, связанных с ИИ и нейротехнологиями, с акцентом на их коэволюцию. **Результаты.** Впервые показано, что нейротехнологии, в отличие от ИИ, создают уникальные риски, такие как прямое воздействие на психику, угрозы идентичности и когнитивной свободе. Выявлены пробелы в регулировании, включая отсутствие специализированных законов о нейроданных. Предложены адаптированные этические рамки, объединяющие принципы прозрачности, подотчетности и защиты прав человека. **Заключение.** Сформулированы рекомендации по регулированию, включая запрет на манипуляцию сознанием, обязательную маркировку контента ИИ и приоритет человеческого контроля над технологиями. Подчеркнута необходимость международного сотрудничества и междисциплинарного диалога для минимизации рисков и обеспечения устойчивого развития нейротехнологий и ИИ в интересах общества.

Ключевые слова: искусственный интеллект, нейротехнологии, объяснимость, правовые нормы регулирования, этические дилеммы.

Для цитирования: Шендерюк-Жидков А. В., Максименко В. А., Храмов А. Е. Козволюция нейротехнологий и ИИ: этические вызовы и подходы к регуляции // Известия вузов. ПНД. 2026. Т. 34, № 1. С. 116–160. DOI: 10.18500/0869-6632-003196. EDN: JXIIOD

Статья опубликована на условиях Creative Commons Attribution License (CC-BY 4.0).

Co-evolution of neurotechnology and AI: ethical challenges and regulatory approaches

A. V. Shenderyuk-Zhidkov^{1,2}, V. A. Maksimenko³, A. E. Hramov^{4,5} ✉

¹Federation Council of the Russian Federation, Moscow, Russia

²Immanuel Kant Baltic Federal University, Kaliningrad, Russia

³Autonomous Nonprofit Organization “Neimark”, Nizhny Novgorod, Russia

⁴Plekhanov Russian University of Economics, Moscow, Russia

⁵Federal State Autonomous Institution “Digital Industrial Technologies”, Moscow, Russia

E-mail: a.shenderyuk@yandex.ru, maximenkovl@gmail.com, ✉ hramov.ae@rea.ru

Received 8.07.2025, accepted 1.08.2025, available online 12.09.2025, published 30.01.2026

Abstract. Purpose of this study is to analyze the ethical challenges at the intersection of neurotechnology and artificial intelligence (AI), and propose regulatory approaches to ensure their responsible development. Special focus is given to personal autonomy, data privacy, social justice, and prevention of mind manipulation. *Methods.* The research employs an interdisciplinary approach, including analysis of scientific literature, regulatory frameworks, and positions of religious institutions. Risks associated with AI and neurotechnologies are compared, emphasizing their co-evolution. *Results.* Neurotechnologies, unlike AI, pose unique risks such as direct mental interference and threats to identity. Regulatory gaps, including the lack of laws on neurodata, are identified. Adapted ethical frameworks combining transparency, accountability, and human rights protection are proposed. *Conclusion.* Recommendations include bans on mind manipulation, mandatory AI content labeling, and human oversight priority. International collaboration and interdisciplinary dialogue are emphasized to mitigate risks and promote sustainable development of these technologies.

Keywords: artificial intelligence, neurotechnology, explainability, legal regulatory standards, ethical dilemmas.

For citation: Shenderyuk-Zhidkov AV, Maksimenko VA, Hramov AE. Co-evolution of neurotechnology and AI: ethical challenges and regulatory approaches. *Izvestiya VUZ. Applied Nonlinear Dynamics.* 2026;34(1):116–160. DOI: 10.18500/0869-6632-003196

This is an open access article distributed under the terms of Creative Commons Attribution License (CC-BY 4.0).

Политика и политики должны опережать технический прогресс и находить своевременные решения прежде, чем технологии породят новую реальность.

Д. А. Медведев. Из интервью газете «Жэньминь Жибао», 13 декабря 2015

Что касается современных технологий, то они, безусловно, угрожают человечеству, но проклинать их не следует, ибо без них будет ещё хуже.

Станислав Лем. Речь на церемонии присвоения ему степени почётного доктора университета Билефельда, 13 ноября 2003

Несмотря на все достижения технологий, что у нас есть, мы остаемся фундаментально ограничены своим биологическим «хардом» и не можем помочь «нашей Вселенной полностью раскрыть свой потенциал и окончательно пробудиться». Но есть способ преодоления этих оков «харда», и имя ему AGI (Общий искусственный интеллект).

Макс Тегмарк. «Жизнь 3.0. Быть человеком в эпоху искусственного интеллекта»

Введение

Бурное развитие нейротехнологий и искусственного интеллекта (ИИ) открывает беспрецедентные перспективы для улучшения жизни человека — от персонализации лечения заболеваний и восстановления утраченных функций до повышения когнитивных способностей [1–3], оптимизации процессов принятия решений [4] и автоматизации рутинных повторяющихся задач [5]. Генеративный ИИ и большие языковые модели уже сделали революцию в нашем восприятии новых информационных технологий, удалив барьеры на общение и взаимодействие с ИИ на естественном языке [6]. Алгоритмы машинного обучения и нейронные сети вторгаются во все сферы

деятельности, позволяя сделать скачок в автоматизации производства [7], улучшить здравоохранение [8, 9], трансформировать образование [10, 11], дать новые инструменты развития науки и технологий [12, 13]. Такие инструменты нейротехнологий, как нейровизуализация активности мозга [14], нейроинтерфейсы [15], интегрированные с биологическими обратными связями [16, 17], различные типы нейромодуляции, такие как транскраниальная магнитная [18, 19] и электрическая стимуляция [20], глубокая стимуляция [21] мозга и т. д., уже доказали свой потенциал в медицине [22], реабилитации [23] и даже в индустрии развлечений [24, 25].

Объединение нейротехнологий и ИИ в рамках одной статьи обусловлено их глубоким взаимовлиянием и схожестью вызовов, которые они ставят перед обществом. Несмотря на то, что эти направления возникли на стыке различных научных дисциплин — нейробиологии, математики и информационных технологий, — их объединяет способность существенно трансформировать человека и социум, особенно в сферах интеллектуального и когнитивного развития, что ранее считалось недоступным для машин, а теперь активно осваивается технологиями.

Например, в 1997 году суперкомпьютер Deep Blue победил чемпиона мира по шахматам Гарри Каспарова [26, 27], а в 2015 году нейросети на соревнованиях ImageNet Large Scale Visual Recognition Challenge превзошли человека в задачах классификации изображений [28]. В 2016 году AlphaGo обыграла лучшего игрока в Го, а в 2017 году AlphaGo Zero, обучаясь самостоятельно, достигла уровня, превосходящего все существующие системы игры в Го [29]. Параллельно нейротехнологии, хотя и развиваются медленнее, также демонстрируют значительные успехи. Методы нейровизуализации, такие как фМРТ [30], ЭЭГ [31], МЭГ [32], фБИКС [33], позволяют визуализировать и анализировать когнитивную или патологическую активность мозга, а нейроинтерфейсы и методы нейромодуляции — влиять на когнитивные процессы. Например, технологии ТМС и нейроинтерфейсы позволяют передавать данные в мозг, минуя традиционные органы чувств [34, 35], и повышать производительность человека при выполнении задач [36, 37]. Стартапы, такие как Neuralink и Synchron, привлекают значительные инвестиции, разрабатывая технологии для взаимодействия с мозгом. Neuralink фокусируется на роботизированной имплантации электродов [38, 39], а Synchron — на стентродах, вводимых через кровеносные сосуды мозга [40]. Эти разработки подчеркивают потенциал нейротехнологий и ИИ как прорывных технологий XXI века, способных усилить естественный интеллект и интегрировать его с искусственным.

Однако такие технологии несут и серьезные риски. Как отмечает Макс Тегмарк в книге «Жизнь 3.0. Быть человеком в эпоху искусственного интеллекта» (см. эпиграф к статье), биологическая природа человека ограничивает его возможности, что порождает стремление к «улучшению» через синергию нейротехнологий и ИИ. Это может привести к вмешательству в когнитивные способности и личность человека, а также к созданию иллюзорных миров, где контент, генерируемый ИИ, становится неотличимым от реальности. Современные тенденции, такие как распространение «умной среды» и гибридизация физической и цифровой реальности [41, 42], развитие которых невозможно без ИИ и нейротехнологий, подчеркивают необходимость всестороннего анализа этических, социальных и юридических вопросов, связанных с этими технологиями. Только опережающее регулирование и междисциплинарное сотрудничество позволят минимизировать риски и обеспечить ответственное использование нейротехнологий и ИИ в интересах общества.

22 апреля 2024 года генеральный директор ЮНЕСКО Одрэ Азуле назначила 24 эксперта для разработки первого глобального стандарта по этическим принципам развития нейротехнологий. В состав группы вошли специалисты из различных регионов мира, представляющие нейронауки, психологию, нейротехнологии, ИИ, интерфейсы «мозг-компьютер», этику и биоэтику. Азуле подчеркнула, что нейротехнологии, несмотря на их потенциал в решении проблем здравоохранения, могут угрожать правам человека, свободе мысли и неприкосновенности частной жизни. Она отметила необходимость создания международных этических рамок, аналогичных тем, что были разработаны для ИИ в 2021 году, подчеркнув, что «без нейроправа не может

быть нейроданных»¹. Однако нейроправа пока нет, а нейроданные активно регистрируются, коллекционируются в датасеты, исследуются и все чаще используются для принятия критических для человека решений.

Итак, стремительное внедрение нейротехнологий и ИИ вызывает серьезные этические опасения, связанные с автономией, конфиденциальностью, справедливостью и ответственностью. Например, использование нейроинтерфейсов и нейромодуляции для изменения поведения или когнитивных функций ставит под сомнение границы личной свободы, а сбор нейроданных повышает риски утечки конфиденциальной информации и ее использования для манипуляции или дискриминации [43, 44]. Кроме того, алгоритмы ИИ, принимающие решения, влияющие на жизнь людей, требуют четкого определения ответственности за ошибки и предвзятость [45–47], а также борьбы с фейковым контентом [48–50], который уже приводит к судебным ошибкам [51, 52]. Особую проблему представляет использование ИИ, включая большие языковые модели, такие как ChatGPT, в образовании. Эти технологии создают новые вызовы, такие как генерация текстов, программ и даже ложных данных, что требует пересмотра образовательных подходов. Вместо запретов необходимо интегрировать ИИ как инструмент, поддерживающий, но не заменяющий интеллектуальный труд студентов [53, 54].

Неоднозначные социальные последствия развития нейротехнологий и ИИ подчеркивают необходимость опережающего обсуждения и разработки нормативных рамок, обеспечивающих их ответственное применение. Отсутствие таких стандартов может привести к злоупотреблениям, углублению социального неравенства и новым формам дискриминации. Согласование этических принципов с технологическим прогрессом становится насущной задачей для обеспечения благополучия и справедливости в обществе [55].

Хотя этические проблемы ИИ [56–58] и нейротехнологий [59–61] часто изучаются отдельно, а их пересечение в методах и приложениях обсуждается [62, 63], данная статья фокусируется на коэволюции этих технологий. Мы ставим целью исследовать этические вызовы, возникающие на стыке нейротехнологий и ИИ, и предложить подходы для их согласования, обеспечивающие ответственное и социально ориентированное совместное развитие этих технологий.

1. Современное состояние развития нейротехнологий и ИИ

1.1. Нейротехнологии. Нейротехнологии — это быстро развивающаяся область, посвященная пониманию устройства мозга и созданию технологий, которые с ним взаимодействуют. Нейротехнологии включают в себя любые технологии, которые оказывают влияние на наше понимание мозга и различных аспектов сознания и мыслительной деятельности, а также технологии, которые позволяют визуализировать активность мозга, так называемая нейровизуализация. Если говорить о конкретных нейротехнологиях, которые мы рассматриваем в статье, то это в первую очередь нейроинтерфейсы, или, как их еще называют, интерфейсы «мозг-компьютер» (ИМК), технологии нейромодуляции, методы нейровизуализации (электро- и магнитоэнцефалография, функциональная магнитно-резонансная томография, функциональная ближняя инфракрасная спектроскопия, позитронно-эмиссионная томография, инвазивные методы регистрации активности мозга).

Хотя история нейротехнологий насчитывает более века, своего расцвета они достигли лишь в конце XX века. Ключевым фактором, обусловившим этот прогресс, стало развитие и внедрение в практику технологий нейровизуализации, которые позволили наблюдать активность мозга *in vivo* в процессе экспериментальных исследований [64, 65]. Нейротехнологии оказывают все большее влияние на общество, хотя их повсеместное распространение зачастую остается незамеченным. От разработки фармакологических препаратов до диагностики патологий мозга нейротехнологии

¹Этические аспекты нейротехнологий: ЮНЕСКО назначила международную группу экспертов для разработки рекомендаций // UNESCO: официальный сайт. 2021. URL: <https://www.unesco.org/ru/articles/eticheskie-aspekty-neyrotekhnologiy-yunesko-naznachila-mezhdunarodnuyu-gruppu-ekspertov-dlya> (дата обращения: 28.01.2025).

оказывают прямое или косвенное воздействие на значительную часть населения развитых стран. К числу примеров относятся медикаментозные и функциональные методы лечения депрессии, инсомнии, синдрома дефицита внимания и гиперактивности (СДВГ), тревожных расстройств. В клинической практике нейротехнологии способствуют улучшению двигательной координации у пациентов, перенесших инсульт, снижению угасания когнитивных функций с возрастом, избавлению от эпилептических приступов, облегчению симптомов двигательных расстройств (болезни Паркинсона, болезни Хантингтона, бокового амиотрофического склероза) и даже снижению интенсивности фантомной боли [66–71].

По мере дальнейшего развития нейротехнологии внесут существенный вклад в улучшение самочувствия и качества жизни людей, страдающих параличом, неврологическими расстройствами и психическими заболеваниями. Одновременно нейротехнологии открывают перспективы для более глубокого понимания и контроля когнитивных и поведенческих функций, влияющих на личность и образ жизни. Уже сегодня существуют технологии, которые стремятся воздействовать на мозговые процессы. Примерами являются когнитивные тренировочные программы, направленные на улучшение различных аспектов когнитивного функционирования [72, 73], нейропомощники, управляющие индивидуальной траекторией обучения студентов с учетом особенностей организации когнитивных процессов в мозге, измеряемых с использованием методов нейровизуализации [74, 75] и т. д.

На переднем крае развития современных нейротехнологий стоят нейроинтерфейсы или ИМК — аппаратно-программные комплексы для функциональной взаимосвязи человека и машины, то есть для прямого соединения вычислительных или других цифровых интеллектуальных систем управления с мозгом. В отличие от традиционных устройств управления, таких как клавиатуры, мыши, джойстики и т. д., ИМК регистрирует активность мозга и преобразует эти сигналы в команды для управления внешним цифровым устройством. ИМК — одна из наиболее быстро развивающихся тем в различных областях науки и техники, включая инженерию, физику, нейронауки, медицину, высокотехнологичные отрасли, коммуникации и робототехнику [15]. Кроме того, нейроинтерфейсы представляют особый интерес для реабилитации и улучшения качества жизни людей с ограниченными возможностями [71, 76–78]. Области применения ИМК включают, в частности, [15]:

- диагностику и контроль патологической активности мозга и нейродегенеративных заболеваний;
- реабилитацию людей после повреждений мозга, например, восстановление двигательных навыков после инсульта;
- анализ и тренировку устойчивости человека к специфическим стрессовым воздействиям;
- улучшение самоконтроля и качества психофизиологического состояния;
- управление роботизированными устройствами, в том числе экзоскелетами, для расширения возможностей человека;
- обеспечение социального взаимодействия за счет того, что социальные приложения могут точно оценивать и передавать эмоции человека;
- помощь частично или полностью парализованным людям во взаимодействии с различными внешними устройствами, например, технология нейрочата, позволяющая людям с ограниченными возможностями общаться с собой и другими людьми;
- игровую индустрию.

ИМК могут включать в себя имплантированные в мозг устройства для съема активности небольших нейронных ансамблей из интересующей нас области мозга, например, моторной коры при необходимости управления манипулятором в случае парализованного пациента [79]. Такие ИМК носят название инвазивных, позволяя ценой хирургического вмешательства обеспечить очень точное представление и классификацию паттернов активности мозга. Для имплантируемых систем характерна проблема биосовместимости — с течением времени имплантированные устройства отторгаются мозгом путем «зарастания» электродов глиальными клетками, так что

регистрация активности нейронов становится невозможной. Неинвазивные ИМК предполагают регистрацию нейронной активности без нарушения целостности покровов мозга, например, используя сенсоры, размещенные над поверхностью головы, то есть они совершенно безопасны с точки зрения медицины [80]. Однако качество расшифровки паттернов активности мозга в этом случае существенно ниже, поэтому многие исследователи считают, что будущее за имплантируемыми системами, и, следовательно, надо прилагать усилия к повышению биосовместимости за счет использования новых материалов, а также повышения безопасности соответствующих хирургических операций [81, 82].

ИМК часто совмещаются с различными типами воздействия на мозг, выступая в качестве систем мониторинга и управления стимуляцией в ответ на те или иные состояния мозга. Такие целенаправленные воздействия на определенную область мозга стимулов различной природы называют нейромодуляцией. Это могут быть неинвазивные воздействия, например, популярная в последнее время транскраниальная магнитная стимуляция (ТМС), то есть прямая стимуляция нейронов коры мозга импульсами магнитного поля, вызывающая активацию или ингибирование нейронов в зависимости от параметров воздействия [83]. Более широким потенциалом применения обладают имплантируемые системы. Разработано несколько типов имплантатов, доступных для клинического применения при лечении болезни Паркинсона, которые используются для электростимуляции в парализованных участках мозга [84]. Имплантируемые системы также разрабатываются для предотвращения эпилептических приступов при фармакорезистентных формах эпилепсии [85]. Нейромодуляция может включать в себя и доставку препарата с использованием устройств, вживляемых непосредственно в субдуральное пространство (интратекальная доставка лекарств) [86]. Новые приложения включают в себя целенаправленное внедрение генов или регуляторов генов и света (оптогенетика), и в начале 2010-х годов такие технологии были впервые продемонстрированы в животных моделях (мыши и обезьяны). Они позволяют модулировать активность популяций нейронов с помощью лазерного излучения [87].

1.2. Искусственный интеллект. Можно давать самые различные определения ИИ, но надо наиболее строго определить ИИ как область математики и ИТ, занимающуюся созданием систем, способных выполнять задачи, требующие человеческого интеллекта, такие как обучение, логическое рассуждение, распознавание образов, принятие решений и обработка естественного языка. Эти системы основаны на алгоритмах и моделях (например, машинном обучении, нейронных сетях), которые позволяют им адаптироваться к данным и улучшать свою производительность без явного программирования для каждой конкретной задачи. В настоящее время ИИ перестал быть исключительно научной областью и активно внедряется в различные сферы жизни — от медицины и финансов до беспилотного транспорта и автоматизированного производства.

Ключевым фактором развития ИИ в последние десятилетия является прогресс в области машинного обучения (ML), особенно в области глубокого обучения. Глубокое обучение, использующее многослойные нейронные сети с огромным числом оптимизируемых параметров, продемонстрировало выдающиеся результаты в таких задачах, как распознавание образов, обработка естественного языка и игра в сложные игры [88]. Создание такой архитектуры глубоких нейронных сетей, как трансформер [89], дало значительный толчок развитию моделей генеративного ИИ, способных генерировать текст, изображения и музыку [90]. Это привело к прогрессу сразу в нескольких областях автоматической интеллектуальной обработки информации.

Во-первых, прогресс в компьютерном зрении позволил ИИ воспринимать и анализировать изображения и видео с высокой точностью. ИИ используется для распознавания лиц, объектов, обнаружения аномалий, анализа медицинских изображений и управления беспилотными транспортными средствами [91–95]. Развитие сверточных нейронных сетей и других алгоритмов глубокого обучения значительно улучшило способность ИИ анализировать визуальные данные [96].

Во-вторых, современный ИИ находит применение в автономных системах, включая беспилотные автомобили, дроны и роботы [97–100]. Разработка интеллектуальных алгоритмов

планирования, навигации и распознавания объектов позволяет создавать автономные системы, способные действовать в сложных динамичных условиях. Однако обеспечение безопасности и надежности автономных систем остается важной и до сих пор не решенной задачей [101, 102].

В-третьих, развитие технологий обработки естественного языка (NLP) позволило ИИ понимать, интерпретировать и генерировать человеческий язык с большей точностью и нюансами [103]. Успехи в области NLP нашли применение в различных областях, включая машинный перевод, создание чат-ботов, анализ тональности текста и информационный поиск. Генеративные модели, основанные на архитектуре «трансформер», такие как ChatGPT, Google Gemini, DeepSeek или GigaChat, демонстрируют способность генерировать связный и контекстуально релевантный текст, что открывает новые возможности для автоматизации коммуникации и доступа к информации, автоматизировать разработку программного обеспечения, собирать информацию из различных источников, эффективно решая задачу реферирования материалов.

Будущее развитие ИИ, вероятно, будет связано с дальнейшим усовершенствованием алгоритмов машинного обучения, разработкой более эффективных архитектур нейронных сетей, исследованием новых парадигм обучения (например, обучения с подкреплением) и интеграцией ИИ в различные области человеческой деятельности. Усилия также направлены на создание более интерпретируемых и прозрачных моделей ИИ, способных объяснить свои решения, что является важным условием для доверия и широкого распространения технологий ИИ, например, в бизнесе или медицине [104–106].

1.3. Синергия и коэволюция нейротехнологий и ИИ. Нейротехнологии и ИИ становятся все более взаимосвязанными, их интеграция открывает обширные возможности создания инновационных технологий, способных расширить наше понимание мозга и улучшить качество биомедицинских технологий по лечению и профилактике различных неврологических заболеваний [1, 107, 108]. Нейротехнологии, обеспечивающие инструментарий для регистрации, стимуляции и модуляции нейронной активности, в сочетании с аналитическими и прогностическими возможностями ИИ формируют мощный синергетический эффект, который выходит за рамки возможностей каждой из этих областей по отдельности [109].

Одним из ключевых направлений, демонстрирующих эту синергию, является разработка ИМК, управляемых ИИ. ИМК могут быть интегрированы с алгоритмами ML для декодирования когнитивных состояний, двигательных намерений и эмоциональных реакций [15]. Это позволяет создавать более интуитивные и адаптивные системы управления внешними устройствами, протезами и коммуникационными системами, которые превосходят существующие аналоги по точности и эффективности. Алгоритмы ML способны обучаться на нейронных данных, выявлять паттерны и предсказывать результаты, позволяя персонализировать работу ИМК и максимально адаптировать их к потребностям конкретного пользователя с использованием специальных методов машинного обучения [15]. ИИ уже повсеместно взаимодействует с нашими нервными системами, влияя, усиливая и изменяя наше поведение и когнитивные процессы.

Другим важным аспектом синергии ИИ и нейротехнологий является использование алгоритмов ИИ для анализа нейроданных, таких как фМРТ и ЭЭГ, что позволяет исследователям выявлять связи между активностью мозга и когнитивными функциями [110–112], а также разрабатывать новые подходы к диагностике и лечению неврологических и психических заболеваний [113–115]. ИИ позволяет обнаружить ранние признаки нейродегенеративных заболеваний или предсказать эффективность различных терапевтических вмешательств на основе анализа нейроданных [113, 116, 117].

Успехи в области понимания функционирования мозга, особенно связанные с обработкой сенсорной информации и принятия решений, влияют на развитие технологий ИИ. Например, хорошо известные сверточные нейронные сети, зарекомендовавшие себя в области анализа изображений, во многом были созданы благодаря исследованиям обработки визуальной информации в зрительной коре мозга [96, 118]. В конце 2024 года исследователи из Google Research представили новый подход к обработке длинных последовательностей данных, вдохновленный работой

человеческой памяти [119]. Исследователи разработали новую архитектуру Titans с тремя типами памяти: (1) краткосрочная память (механизм внимания) — для точной обработки текущего контекста; (2) долговременная память — для хранения важной информации из прошлого; (3) постоянная память — для хранения общих знаний о задаче. Система учится запоминать важную информацию так же, как это делает человеческий мозг — фокусируясь на неожиданных и значимых событиях. При этом используется умный механизм «забывания», позволяющий освобождать память от неактуальной информации.

Взаимопроникновение нейротехнологий и ИИ влечет за собой большие этические и правовые последствия. Вызывает беспокойство отсутствие надежной нормативной основы для регулирования в этой пограничной области. Для понимания глубины проблемы достаточно привести два примера.

Во-первых, компании могут использовать данные об активности мозга, полученные с помощью нейротехнологических устройств, в маркетинговых целях. Обнаруживая биомаркеры активности мозга, связанные с нашими предпочтениями и антипатиями, эти компании могут влиять на поведение клиентов для получения максимальной прибыли. В частности, развитие языковых моделей (англ. Large Language Models, LLM) позволяет создавать цифровые двойники потребителей на базе LLM-агентов с учётом их демографических данных и доступной персональной информации, что даёт возможность достаточно точно прогнозировать поведение и предпочтения отдельных групп. Это поднимает тревожные вопросы о системах слежения, маркетинговых тактиках и политическом влиянии на наши частные мысли и эмоции, что в конечном итоге угрожает как основам общества, так и индивидуальной свободе каждого гражданина [120, 121].

Во-вторых, внедрение нейротехнологий усугубляет существующее социальное неравенство. Это особенно проявляется в двух ключевых аспектах. В социальных науках распространён так называемый WEIRD-эффект (Western, Educated, Industrialized, Rich, Democratic): порядка 80% всех исследований человеческого поведения базируются на выборке преимущественно образованных, индустриализированных, обеспеченных жителей западных демократий, что составляет всего 12% населения Земли. Информации о предпочтениях, психофизиологических особенностях и социальном опыте остальных 88% попросту не хватает, или она отсутствует вовсе. Поскольку именно эти данные ложатся в основу обучения крупных моделей искусственного интеллекта, создаётся ситуация, когда ИИ отражает исключительно мировоззрение небольшого процента людей, что приводит к системному неучастию и дискриминации большинства. Сами тренировочные данные часто содержат скрытые дискриминационные факторы. Например, алгоритм, предназначенный для показа рекламных вакансий в областях науки, технологий, инженерии и математики (STEM), изначально разрабатывался как гендерно-нейтральный, однако на практике женщины видят такие объявления реже, чем мужчины: признаки, используемые моделью, коррелируют с полом и распределены неравномерно. В результате алгоритм непреднамеренно воспроизводит гендерные стереотипы [122]. Другой пример — «редлайнинг»: при оценке заявок на кредит в качестве фактора принимается почтовый индекс, который сам по себе кажется нейтральным, но на деле коррелирует с расой из-за исторического и демографического состава районов. Это приводит к косвенной расовой дискриминации. В литературе по машинному обучению такую ситуацию называют «косвенная дискриминация». Она возникает, когда алгоритм опирается на атрибуты, которые прямо не содержат дискриминирующих признаков, но статистически связаны с ними.

Важным аспектом социальной дискриминации является тот факт, что доступ к передовым технологиям ИИ и нейротехнологиям — например, к методам когнитивной реабилитации — неизбежно будет преимущественно у состоятельных и образованных групп населения. Это способно ещё больше углубить разрыв между богатыми и бедными на международном, государственном и локальном уровнях [43, 123].

Синергия ИИ и нейротехнологий представляет собой мощную силу, способную улучшить наше понимание мозга и предоставить новые инструменты для персонализированного лечения,

реабилитации, повышения когнитивных возможностей и, как следствие, улучшения качества жизни. Однако необходимо осознавать этические и социальные последствия такого технологического прогресса и сформулировать те этические проблемы и дилеммы, которые ставит перед нами развитие таких сквозных технологий, как нейротехнологии и ИИ. Проведем в следующем разделе систематический обзор этических дилемм, которые ставит перед нами развитие этих технологий.

2. Этические проблемы и дилеммы, возникающие при коэволюции технологий

Итак, развитие нейротехнологий и ИИ представляет собой не только технологическую революцию, но и источник глубоких этических и социальных дилемм, которые требуют тщательного анализа и проработки [55]. Сложность этических проблем, связанных с этими технологиями, затрагивающими непосредственно когнитивную и интеллектуальную сторону нашей деятельности, обусловлена их многогранностью и способностью влиять на различные сферы человеческой жизни, включая принятие решений, занятость, конфиденциальность, социальные отношения, улучшение биологической природы человека и даже само понимание человеческой идентичности. Изучение этических принципов и последствий, связанных с разработкой, развертыванием и использованием нейротехнологий (и связанных с ними исследований в области нейронауки и нейроданных), обычно называют нейроэтикой, относительно зарождающейся, но растущей областью исследований, возникшей в конце 1990-х и начале 2000-х годов из медицинской и биоэтики [124]. Нейроэтика критически относится к предположениям и намерениям, лежащим в основе нейротехнологических и нейронаучных открытий. Она также обеспокоена вопросами о влиянии нейротехнологий на самопонимание человека и о влиянии изменений в этом фундаментальном понимании на нашу биологию, нашу психологию и наше общество.

2.1. Предвзятость, справедливость и доступность. Одной из ключевых этических проблем ИИ является предвзятость алгоритмов [125], которые могут воспроизводить и усиливать существующие социальные неравенства и дискриминацию [126, 127]. ИИ-системы обучаются на данных, содержащих исторические предрассудки и стереотипы, что приводит к дискриминационным практикам в таких областях, как трудоустройство, кредитование и правосудие [128–131]. Обеспечение справедливости алгоритмов требует тщательного отбора данных, разработки прозрачных методов обучения и постоянного мониторинга. Неравномерный доступ к нейротехнологиям и ИИ усугубляет социальное неравенство, создавая новые формы исключения. Этот «технологический разрыв» (англ. digital divide) способствует перераспределению ресурсов в пользу привилегированных групп, усиливая социальную несправедливость [132]. Например, нейроинтерфейсы и нейромодуляция, способные улучшать когнитивные и физические способности, могут стать инструментом создания «нейрократического» общества, где элита обладает не только экономическим, но и когнитивным превосходством. Люди без доступа к таким технологиям рискуют столкнуться с социальной изоляцией и стигматизацией и будут восприниматься как «отсталые» или «неполноценные». Использование ИИ в найме и оценке персонала также вызывает опасения. Алгоритмы, обученные на данных, отражающих предрассудки, могут дискриминировать по признакам расы, пола или возраста [127]. Нейротехнологии, применяемые для мониторинга состояния работников, создают новые формы контроля и давления, а также дискриминации в отношении тех, чьи нейрофизиологические характеристики не соответствуют «идеальным» стандартам. Еще одной проблемой является неравный доступ к цифровому здравоохранению [9]. ИИ и нейротехнологии активно внедряются в диагностику, лечение и реабилитацию, но их доступность ограничена социально-экономическими факторами, такими как низкий доход, отсутствие страховки или проживание в удаленных регионах. Это может привести к ухудшению здоровья и качества жизни для тех, кто не может воспользоваться передовыми медицинскими услугами.

2.2. Прозрачность и интерпретируемость. Стремительное развитие ИИ порождает фундаментальную этическую дилемму, связанную с балансом между эффективностью и интерпретируемостью решений ИИ. С одной стороны, сложные модели ИИ, такие как глубокие нейронные сети, демонстрируют высокую производительность в задачах распознавания образов, прогнозирования и других областях. С другой стороны, модели ИИ часто подобны «черному ящику», что затрудняет понимание и объяснение принимаемых ими решений, особенно в критически важных сферах, таких как медицина, правосудие и финансы [133–135].

Это противоречие обусловлено стремлением разработчиков к повышению точности моделей, что достигается за счет их усложнения. Однако чем сложнее модель, тем труднее выявить факторы и логику, лежащие в основе ее решений. Это создает серьезные этические и социальные проблемы, включая:

- Невозможность оспаривания решений: если решение ИИ нельзя объяснить, люди лишаются возможности оспорить его, даже если оно кажется несправедливым или ошибочным, что подрывает принципы справедливости и равенства.
- Отсутствие ответственности: если алгоритмы работают как «черные ящики», сложно определить, кто несет ответственность за ошибки или вред, причиненный их решениями.
- Потеря доверия: отсутствие прозрачности снижает доверие к технологиям, что затрудняет их широкое внедрение и принятие.

Для решения этих проблем активно развиваются методы построения объяснимого ИИ (Explainable AI, XAI), которые направлены на создание моделей, способных обосновывать свои решения и стать более прозрачными для человека. Это может быть достигнуто за счет использования более простых и интерпретируемых моделей, разработки методов визуализации и анализа сложных систем [136, 137], а также применения методов постобработки для изучения поведения ИИ как «черного ящика» для выявления важных факторов, влияющих на решения [138].

Среди современных методов XAI можно выделить несколько ключевых подходов:

- *LIME (Local Interpretable Model-agnostic Explanations)* — аппроксимирует сложную модель локально линейными интерпретируемыми моделями для объяснения отдельных предсказаний [105].
- *SHAP (SHapley Additive exPlanations)* — использует теорию игр для распределения вклада каждого признака в итоговое решение модели [139]. Эти подходы находят применение и на стыке ИИ и нейротехнологий [140].
- *Механизмы внимания* — визуализируют веса внимания в нейронных сетях, показывая, на какие части входных данных модель обращает больше внимания. Данный подход находит большое применение в анализе медицинских систем ИИ [141, 142].
- *Градиентные методы* (например, Grad-CAM) — выделяют важные области входных данных на основе градиентов модели [143].
- *Правила и деревья решений* — предоставляют интерпретируемые логические правила, лежащие в основе работы модели [144].

Эти методы позволяют не только понять логику работы сложных моделей, но и выявить потенциальные смещения в данных или неожиданные зависимости, что особенно важно в критически важных областях, таких как медицина или финансы [145, 146].

Аналогичные проблемы прозрачности и интерпретируемости возникают в нейротехнологиях, особенно при использовании нейроданных и их анализе с помощью ИИ. Отсутствие прозрачности в этой области может привести к аналогичным этическим последствиям, включая невозможность оспаривания решений, отсутствие ответственности и потерю доверия.

2.3. Конфиденциальность и безопасность данных. ИИ-системы, основанные на анализе больших данных, обрабатывают значительные объемы персональной информации, что создает угрозы для конфиденциальности и безопасности. Утечки данных, несанкционированный доступ и

злоупотребление информацией могут иметь серьезные последствия для отдельных лиц и общества [147, 148]. Для минимизации этих рисков необходимы надежные механизмы защиты данных, включая методы анонимизации и псевдонимизации, что является ключевым условием этичного использования ИИ.

Нейротехнологии, особенно в сочетании с ИИ, также сталкиваются с проблемами конфиденциальности. Данные нейровизуализации, отражающие когнитивные и эмоциональные процессы, представляют собой особо чувствительную информацию. Риски включают несанкционированный доступ, использование нейроданных для дискриминации (например, при найме или страховании), манипулирование поведением и нарушение приватности [43, 109]. Отсутствие четких нормативных рамок и механизмов защиты увеличивает вероятность злоупотреблений, что угрожает автономии и благополучию человека.

2.4. Принятие решений без участия человека и контроль за ним. Одной из ключевых этических дилемм, связанных с развитием ИИ, является конфликт между автономией систем и необходимостью сохранения человеческого контроля. По мере роста сложности ИИ-систем, способных принимать решения без прямого вмешательства человека, возникает вопрос о допустимых пределах их автономии и механизмах обеспечения ответственности [149, 150].

С одной стороны, автономные системы, такие как беспилотные автомобили, медицинские диагностические платформы или финансовые алгоритмы, способны действовать быстрее и эффективнее, чем человек, что открывает значительные преимущества. С другой стороны, утрата контроля над ИИ создает риски, включая несоответствие действий систем человеческим ценностям, злоупотребление технологиями и трудности в установлении ответственности за ошибки.

Таким образом, поиск баланса между автономией ИИ и сохранением человеческого контроля становится критически важным для предотвращения негативных последствий и защиты прав и свобод людей в условиях растущей автоматизации.

2.5. Автономия и самоопределение. Развитие ИИ и нейротехнологий создает угрозы для автономии и самоопределения человека, включая возможность манипулирования сознанием и изменения идентичности [60]. Например, технологии ИИ, такие как дипфейки (фото, видео, аудио), могут использоваться для манипуляции общественным мнением, нанесения репутационного ущерба или склонения людей к незаконным действиям [48–52]. Нейротехнологии, такие как ТМС и глубокая стимуляция мозга, способны влиять на нейронную активность, что в перспективе может затрагивать когнитивные и эмоциональные процессы, лежащие в основе принятия решений. В сочетании с ИИ, анализирующим нейроданные, эти технологии могут использоваться для скрытого воздействия на выбор и поведение людей, подрывая их свободу воли.

Риски манипуляции сознанием и изменения идентичности угрожают личной автономии и достоинству. Использование нейротехнологий и ИИ для изменения убеждений, ценностей или мотиваций без согласия человека может привести к утрате самоопределения и превращению индивида в объект внешнего контроля. Кроме того, возможность «улучшения» когнитивных способностей создает социальное давление, вынуждая людей подвергаться таким процедурам, чтобы не отставать от других, что ограничивает их свободу выбора.

В 2012 году Центр стратегического анализа при Правительстве Франции опубликовал доклад «Мозг и законодательство: анализ появления нейроправа» [151], где рассмотрены этические вопросы применения нейровизуализации для диагностики личностных особенностей. Один из авторов доклада подчеркивает [152], что различия в уровнях эмпатии, интеллекта, импульсивности и агрессии между людьми создают вызов для правовой системы, основанной на презумпции равенства. Однако утверждение о равенстве применительно к нервным системам является ложным. Достижения нейронаук дадут возможность понять поведение людей в длительном континууме, а не через упрощенные категории, которые сейчас используются. Это требует пересмотра подходов к оценке этичности нейротехнологий.

Таким образом, при обсуждении автономии и самоопределения необходимо учитывать биологические, генетические и когнитивные особенности человека. Также требуется разработка

технологий и процедур, обеспечивающих защиту нейроданных от несанкционированного доступа, изменения или уничтожения.

2.6. Потеря рабочих мест и стимулирование экономики. Влияние ИИ на занятость и экономику представляет собой этическую дилемму, связанную с сокращением рабочих мест из-за автоматизации и созданием новых возможностей за счет повышения производительности и инноваций за счет внедрения технологий ИИ [153, 154]. С одной стороны, автоматизация на основе ИИ может привести к сокращению рабочих мест, особенно в секторах с рутинными задачами, что усугубит неравенство в доходах и потребует масштабной переквалификации рабочей силы. Темпы внедрения ИИ вызывают опасения, что люди и организации не успеют адаптироваться к изменениям, что может привести к негативным социальным последствиям.

С другой стороны, ИИ способен стимулировать экономический рост за счет повышения эффективности и создания новых отраслей, связанных с разработкой и обслуживанием ИИ-систем. Это может компенсировать потери рабочих мест в традиционных секторах. Однако успешное использование ИИ зависит от адаптации образовательных систем, политик социальной защиты и способности стран эффективно внедрять эти технологии.

2.7. Благополучие и безопасность. Развитие ИИ и нейротехнологий требует оценки их влияния на благополучие и безопасность пользователей. Под благополучием понимается состояние физического и психического удовлетворения (включая здоровье, безопасность, счастье и комфорт отдельных лиц и/или сообществ), достигаемое за счет предотвращения вреда, травм или необоснованных рисков при разработке и внедрении технологий. Это включает психологическую безопасность, а также социальные и экологические аспекты, такие как сохранение культуры и предотвращение экологических катастроф.

Важно отметить, что рассматриваемые проблемы не являются взаимоисключающими и в значительной степени связаны. Например, в соответствии с современными этическими нормами, закрепленными в Хельсинкской декларации Всемирной медицинской ассоциации «Рекомендации для врачей, занимающихся биомедицинскими исследованиями с участием людей», принятой 18-й Всемирной медицинской ассамблеей (последний раз пересмотренной в Эдинбурге в 2000 г.), применение нейротехнологий возможно только после получения информированного согласия на сбор нейроданных, которое предполагает обеспечение психической конфиденциальности, автономии и самостоятельности человека, безопасности данных. Благополучие может быть достигнуто только при условии решения этических проблем, связанных с прозрачностью, ответственностью и защитой прав пользователей.

2.8. Экзистенциальные риски. Хотя ИИ имеет много применений, он все еще очень далек от достижения форм интеллекта, близких к людям. Однако тот факт, что эта технология в основном неизвестна широкой общественности, вызывает опасения по поводу возможности контролировать ее и согласовывать с общечеловеческими ценностями, если она достигнет более высокой формы интеллекта [155]. В табл. 1 перечислены этические проблемы и как они связаны с возможностями ИИ.

Хотя в настоящее время маловероятны, но потенциально существуют экзистенциальные риски, связанные с развитием сверхинтеллекта — ИИ, превосходящего человеческий интеллект [156, 157]. Исследователи и философы активно обсуждают, как предотвратить негативные последствия и обеспечить использование сверхинтеллекта в интересах человечества. Стивен Хокинг, один из наиболее известных исследователей, предупреждал, что создание ИИ, превосходящего человека, может привести к потере контроля: *«Такой разум возьмет инициативу на себя и станет сам себя совершенствовать со все возрастающей скоростью. Люди, ограниченные медленной биологической эволюцией, не смогут конкурировать с машинами»*².

²С. Хокинг: Искусственный интеллект может уничтожить человечество // BBC Russian : служба новостей. 2014. URL: https://www.bbc.com/russian/science/2014/12/141202_hawking_ai_danger (дата обращения: 27.01.2025).

Таблица 1. Основные вопросы и соответствующие им проблемы этики ИИ

Table 1. Key questions and related problems of AI ethics

Вопросы развития ИИ	Подходы к разрешению этических противоречий
ИИ нуждается в данных	Конфиденциальность данных и управление ими
ИИ является «черным ящиком»	Объяснимость и прозрачность технологии
ИИ может принимать или рекомендовать решения	Справедливость и согласование ценностей
ИИ основан на больших данных, статистике и всегда имеет некоторый процент ошибок	Защита от дискриминации со стороны ИИ
ИИ очень распространен и динамичен	Подотчетность ИИ
Хорошее или плохое использование ИИ	Большие негативные последствия злоупотребления технологией ИИ
	Быстрая трансформация рынка труда, экономики и общества
	Цели устойчивого развития ООН и этические принципы развития ИИ ЮНЕСКО
	Автономные устройства и массовый контроль
	Дипфейки, создаваемые ИИ

Из рассмотренных этических дилемм вытекают риски, которые уже сегодня могут иметь серьезные последствия в различных сферах. Специалисты компании Яндекс выделяют четыре группы рисков, связанных с ИИ: социальные, юридические, этические и технологические [158]. Например, на этапе сбора и подготовки данных для обучения моделей ИИ выделяются следующие риски:

- *Юридические*: (1) использование данных из запрещенных источников; (2) нарушение авторских прав; (3) несоответствие законодательным требованиям.
- *Технологические*: (4) ограничения на использование зарубежных систем и инфраструктуры.
- *Этические*: (5) включение предвзятых данных в обучающие массивы.

Таким образом, исследователи все чаще указывают на необходимость регулирования ИИ и нейротехнологий для минимизации рисков. В следующем разделе статьи рассматриваются подходы к разработке и эксплуатации систем ИИ, основанные на этических принципах, ориентированных на человека и соблюдение его прав.

3. Разработка руководящих принципов регулирования при разработке и использовании ИИ и нейротехнологий

Проанализируем ключевые аспекты разработки и внедрения руководящих принципов для регулирования ИИ и нейротехнологий. Мы остановимся на различных подходах к определению принципов регулирования, включая этические, правовые и социально-экономические аспекты. Необходимо отметить, что правовая база, регулирующая разработку и использование ИИ и нейротехнологий, находится на стадии формирования основных принципов и подходов. Большинство нормативных актов в области ИИ используют комбинацию различных стратегий, ориентированных на уровень риска, связанного с конкретными технологиями. Однако единый подход и методология в юридической практике пока отсутствуют, а правовые документы содержат размытые формулировки, не охватывающие весь жизненный цикл разработки и применения систем ИИ и нейротехнологий.

В отношении нейротехнологий ситуация еще более критична — специальные законы или правовые акты в большинстве стран отсутствуют. Регулирование осуществляется через существующие законы о медицине, защите персональных данных, правах человека и интеллектуальной собственности. Некоторые страны разрабатывают этические рекомендации, но они редко закрепляются на законодательном уровне.

3.1. Подход, основанный на принципах. В рамках данного подхода всем заинтересованным сторонам предполагается набор фундаментальных предложений (руководящих принципов), которым соответствуют все процессы по разработке и использованию систем ИИ. Законы об ИИ и законопроекты об ИИ разных стран явно и неявно воспроизводят принципы, закрепленные в документах ЮНЕСКО. Это касается Закона Перу № 31814 от 2023 года «Закон, способствующий использованию искусственного интеллекта в интересах экономического и социального развития страны»: в перуанском законодательстве закреплены такие принципы, как «стандарты безопасности, основанные на оценке риска», «многосторонний подход», «этическая разработка для ответственного искусственного интеллекта» и «конфиденциальность». Также одним из самых ярких примеров использования данного подхода является созданная многими государствами так называемая «Белая книга», являющаяся сборником этических принципов и рекомендаций к работе ИИ (см. разделы 4.3 и 4.4).

Регулирование ИИ, основанное исключительно на принципах, не налагает конкретных обязательств или ограничений на государственные органы, частные организации или людей, и они не имеют последствий за несоблюдение.

3.2. Подход, основанный на стандартах. Подход, основанный на стандартах, делегирует (полностью или частично) регулирующие полномочия государства органам по установлению стандартов, которые могут быть государственными, частными или гибридными организациями. При таком подходе регулирование означает, что организации, устанавливающие стандарты, разрабатывают технические стандарты, которые регулируют реализацию обязательных правил. Такой подход предполагает и поощряет отраслевые организации прямо или косвенно участвовать в разработке технических стандартов, регулирующих процессы и деятельность, связанные с ИИ и нейротехнологиями.

Например, в пункте 121 Закона ЕС об ИИ — нормативного акта Европейского союза об искусственном интеллекте, принятого Европейским парламентом 13 марта 2024 года и одобренного Советом ЕС 21 мая 2024 года, — говорится, что «стандартизация должна играть ключевую роль по предоставлению технических решений поставщикам для обеспечения соблюдения настоящего Регламента в соответствии с современным уровнем техники, для продвижения инноваций, а также конкурентоспособности и роста единого рынка»³.

3.3. Гибкий и экспериментальный подход на основе нормативных «песочниц». Данный подход предполагает создание гибких схем регулирования, таких как нормативные «песочницы» и другие экспериментальные площадки, которые позволяют организациям тестировать новые бизнес-модели, методы, инфраструктуру и инструменты в условиях более гибкого регулирования при сопровождении государственных органов. Такие гибкие подходы были изначально предложены для таких областей экономики, как телекоммуникации и финансы, а также для сквозных технологий, например, защиты данных и конфиденциальности. В последнее время такой подход также изучает возможность регулирования ИИ.

Следующие пять целей, которые статья 57 (9) Закона ЕС об ИИ устанавливает для регуляторных «песочниц» ИИ: «(1) повышение правовой определенности для достижения нормативного соответствия настоящему Регламенту или, где это применимо, другому применимому

³World's first major law for artificial intelligence gets final EU green light // CNBC. 2024. URL: <https://www.cnbc.com/2024/05/21/worlds-first-major-law-for-artificial-intelligence-gets-final-eu-green-light.html> (дата обращения: 27.01.2025).

законодательству Союза и национальному законодательству; (2) поддержка обмена передовым опытом посредством сотрудничества с органами, участвующими в регуляторной «песочнице» ИИ; (3) содействие инновациям и конкурентоспособности и содействие развитию экосистемы ИИ; (4) содействие нормативному обучению на основе фактических данных; (5) содействие и ускорение доступа к рынку Союза для систем ИИ, в частности, предоставляемых малыми и средними предприятиями, включая стартапы». Согласно статье 57 Закона ЕС об ИИ, каждое государство-член Евросоюза на национальном уровне должно обеспечить создание компетентными органами нормативных «песочниц» для ИИ.

3.4. Создание среды, способствующей разработке и использованию ответственных, этических и соответствующих правам человека систем ИИ. Регулирующие инструменты могут быть направлены на создание среды, которая поощряет разработку и использование ответственных, этических и соответствующих правам человека систем ИИ. Законы и законопроекты об ИИ могут включать положения, направленные на содействие созданию возможности, связанные с человеческим капиталом, технологиями, инфраструктурой и институциональным контекстом.

В этом ключе, например, ЮНЕСКО разработала методологию оценки готовности (Readiness assessment methodology, RAM)⁴, которая направлена на то, чтобы помочь «странам понять, где они находятся по степени готовности внедрять ИИ этично и ответственно для всех своих граждан, делая акцент на том, какие институциональные и нормативные изменения необходимы». Предложенная методология помогает определить стране сильные стороны и пробелы по пяти направлениям: (1) юридическому, (2) социальному и культурному, (3) научно-образовательному, (4) экономическому, (5) технологическому и инфраструктурному.

3.5. Адаптация существующих законов. Данный подход предполагает внесение поправок в отраслевые правила (например, здравоохранение, финансы, образование, правосудие) и трансверсальные правила (например, уголовные кодексы, государственные закупки, законы о защите данных, трудовое законодательство) для постепенного улучшения существующей нормативно-правовой базы в соответствии с текущим уровнем развития технологий.

Некоторые юрисдикции предпочли адаптировать правила, специфичные для конкретных секторов (например, здравоохранение, финансы, образование, правосудие) и трансверсальные правила (например, уголовные кодексы, государственные закупки, данные законы о защите, трудовое законодательство) вместо создания специализированного законодательства, регулирующего развитие ИИ. Одно из потенциальных преимуществ этого подхода заключается в том, что он позволяет законодателям обсуждать и вносить постепенные улучшения в нормативную базу структуры, основанной на том, что они узнают о последствиях использования технологий ИИ.

Примером адаптации трансверсальных правил является законопроект Колумбии № 225/2024 - Сенат, который внесет изменения в Уголовный кодекс, чтобы ужесточить наказание за преступление выдачи себя за другое лицо, когда системы ИИ используются для создания глубоких подделок (deepfakes) или иного цифрового мошенничества, наносящего ущерб чести, репутации или безопасности жертв. Более того, в Аргентине законопроект предлагает внести поправки в три статьи Закона № 25467 от 2021 года («Закон о науке, технологиях и инновациях»), чтобы изложить принципы и ценности для «этичной эксплуатации» систем ИИ, установить обязательства регистрировать системы ИИ, предоставляя компетентному органу полномочия останавливать разработку систем ИИ, когда это нарушает принципы и ценности и позволяет кому-либо подвергаться угрозе или вреду от такой системы, чтобы сообщить об этом.

3.6. Подход, основанный на оценке рисков. Данный подход предполагает установление обязательства и требования в соответствии с оценкой рисков, связанных с развертыванием и

⁴Readiness assessment methodology // UNESCO: официальный сайт. URL: <https://www.unesco.org/ethics-ai/en/ram> (дата обращения: 14.01.2025).

использованием определенных инструментов ИИ в конкретные контексты. Такой подход может быть эффективно реализован в различных секторах, включая экологию, налоги, безопасность пищевых продуктов, стихийные бедствия и защиту прав потребителей. В рамках него необходимо установить приоритеты и типы рисков, а также выбрать систему вмешательств, связанных с разработкой и использованием систем ИИ, которая будет адаптивно зависеть от степени риска.

Ярким примером такого подхода к регулированию является Закон ЕС об ИИ (см. разделе 4.2), который устанавливает обязательства, основанные на различных уровнях риска: неприемлемый, высокий, системный, ограниченный и минимальный. Согласно статье 3 (1а), «риск» понимается как «совокупность вероятности возникновения вреда и серьезности этого вреда». Практики ИИ, которые относятся к категории «неприемлемый риск» подпадают под запрет. Примером такой практики является «использование систем удаленной биометрической идентификации в режиме реального времени в общедоступных местах в целях обеспечения правопорядка», если только это не является строго необходимым для достижения целей, связанных с определенными уголовными расследованиями и предотвращением конкретных, например, террористических угроз. Более того, системам ИИ, которые попадают в категорию «высокий риск», должны соответствовать конкретные обязанности, связанные с системами управления рисками, управлением данными, технической документацией, ведением учета, прозрачностью, человеческим надзором, точностью, надежностью и кибербезопасностью (глава 2 Раздела III), а поставщики систем ИИ с высоким уровнем риска также должны соблюдать ряд установленных требований.

3.7. Подход, основанный на правах. Подход, основанный на правах, ориентирован на обеспечение защиты прав и свобод человека и общества в целом при регулировании ИИ. Этот подход подчеркивает, что регулирование должно основываться не только на экономических и технологических аспектах, но и на необходимости защиты гражданских прав, достижения социальных целей, обеспечения справедливого распределения общественных благ и равного доступа к технологиям для всех категорий граждан на всех этапах жизненного цикла систем ИИ.

Ярким примером реализации подхода, основанного на правах, являются положения Европейского регламента по защите данных (General Data Protection Regulation, GDPR⁵), регулирующие автоматизированную обработку персональных данных, включая профилирование [159]. Согласно GDPR, субъект данных имеет право не подвергаться решениям, принятым исключительно на основе автоматизированной обработки, если такие решения имеют для него юридические последствия или существенно влияют на его права. Эти обязательства не зависят от уровня риска и применяются во всех случаях использования подобных систем.

Гражданские организации Евросоюза настаивали на том, чтобы законодательство в области ИИ базировалось на подходе, основанном на правах, а не на оценке рисков. Их позиция заключалась в том, что права человека являются неотъемлемыми и должны соблюдаться независимо от уровня риска, связанного с внешними факторами. Такой подход обеспечивает приоритет защиты фундаментальных свобод и достоинства личности в условиях развития технологий.

3.8. Подход, основанный на ответственности. Данный подход предполагает установление четких механизмов распределения ответственности и введения санкций в случаях, когда применение ИИ-технологий приводит к негативным последствиям. В рамках этого подхода требуется определить конкретные субъекты, которые несут ответственность за разработку, внедрение и эксплуатацию ИИ-систем. Это может включать в себя разработчиков алгоритмов, поставщиков ИИ-решений, операторов систем и в некоторых случаях конечных пользователей. Разграничение ответственности должно основываться на анализе причинно-следственных связей между действиями субъектов и возникшими негативными последствиями, такими как негативные инциденты (например, наезд беспилотного автомобиля на пешехода), нарушения прав человека (например, незаконный отказ от приема на работу), дискриминация или нанесение экономического ущерба.

⁵General Data Protection Regulation (GDPR). URL: <https://gdpr-info.eu> (дата обращения: 27.01.2025).

Санкции за проблемное использование ИИ должны быть соразмерны тяжести последствий и могут включать в себя как административные меры (штрафы, приостановление деятельности), так и гражданско-правовую ответственность (компенсация ущерба). Важно отметить, что подход к ответственности может быть не только репрессивным, но и стимулирующим: поощрять ответственное использование ИИ, в частности, внедрение этических принципов и обеспечение прозрачности алгоритмов.

4. Анализ существующей правовой базы регулирования ИИ и нейротехнологий

Правовая база, регулирующая разработку и использование ИИ и нейротехнологий, находится на стадии формирования основных принципов и подходов. Большинство нормативных актов в области ИИ используют комбинацию различных стратегий, ориентированных на уровень риска, связанного с конкретными технологиями. Однако единый подход и методология в юридической практике пока отсутствуют, а правовые документы содержат размытые формулировки, не охватывающие весь жизненный цикл разработки и применения систем ИИ и нейротехнологий.

В отношении нейротехнологий специальные законы или правовые акты в большинстве стран отсутствуют. Регулирование осуществляется через существующие законы о медицине, защите персональных данных, правах человека и интеллектуальной собственности. Некоторые страны разрабатывают этические рекомендации, но они редко закрепляются на законодательном уровне.

Остановимся кратко на мировом опыте по регулированию ИИ и нейротехнологий.

4.1. Китайская народная республика. 15 августа 2023 года в Китае вступила в силу Директива по регулированию ИИ, устанавливающая следующие правила:

- маркировка контента, созданного с помощью ИИ;
- использование только легальных данных для обучения моделей;
- создание механизмов для рассмотрения публичных жалоб на ИИ-услуги и контент.

Китай также разработал детальные требования к качеству, оценке и актуальности данных для обучения ИИ.

4.2. Европейский Союз. Закон о регулировании ИИ, вступающий в силу в 2025 году, разделяет системы ИИ на группы по уровню риска:

- **Неприемлемый риск:** запрещены системы, воздействующие на подсознание, системы социального рейтинга, разделяющие людей на категории по их поведению, социально-экономическому положению и прочим личным характеристикам, и распознавания в реальном времени (за исключением правоохранительных органов с разрешения суда).
- **Высокий риск:** системы, применяемые в лицензируемых товарах и услугах (игрушки, автомобили, медицинская аппаратура), управлении критической инфраструктурой, образовании, работе с персоналом и правоохранительных органах. Такие системы подлежат обязательной оценке и регистрации в специальной базе данных ЕС.
- **Низкий риск:** системы, генерирующие контент (например, ChatGPT), должны маркировать контент, предотвращать создание незаконного контента и раскрывать данные, защищенные авторским правом, на которых обучались модели. От всех прочих систем требуется лишь быть достаточно открытыми — пользователь должен знать, что имеет дело с ИИ, и иметь достаточно информации, чтобы принять обоснованное решение, пользоваться ли им дальше.

Евросоюз является одним из лидеров в области регулирования нейротехнологий. В 2020 году была опубликована Белая книга по ИИ, в которой заложены принципы этичного и безопасного развития технологий, включая нейротехнологии. Также действуют общие положения уже упомянутого в предыдущем разделе Европейского регламента по защите данных (General Data Protection Regulation) относительно защиты нейроданных.

Впервые специальный раздел законодательства, посвященный регулированию нейротехнологий, появился во Франции, в частности, было ограничено коммерческое применение технологий нейровизуализации [160]. В 2011 г. в кн. I «О лицах» Гражданского кодекса Франции была включена гл. IV, содержащая лишь одну статью — ст. 16–14 «Использование методов визуализации мозга», в которой зафиксировано, что «методы визуализации мозга могут использоваться только в медицинских или научных целях или в качестве части судебной экспертизы». В 2021 году во Франции принята новая редакция закона о биоэтике, в котором присутствует ряд положений, касающихся регулирования нейротехнологий, в частности, регулируется использование всех методов записи мозговой активности); вводится понятие нейромодуляции; а также вводится ст. L. 1151-4: «Действия, процедуры, техники, методы и оборудование, предназначенные для изменения мозговой деятельности, представляющие серьезную опасность или предполагаемую серьезную опасность для здоровья человека, могут быть запрещены указом после заключения Высшего органа здравоохранения».

4.3. Соединенные Штаты Америки. 4 октября 2022 года президент Джо Байден представил Билль о безопасном и надежном ИИ, включающий следующие направления⁶:

- создание безопасных и эффективных систем;
- защита от дискриминации со стороны алгоритмов;
- обеспечение конфиденциальности данных;
- объяснение принципов работы систем пользователям;
- возможность выбора человека в качестве альтернативы машине.

В июле 2023 года крупнейшие технологические компании (Amazon, Google, Meta, Microsoft и др.) подписали соглашение о саморегулировании, обязавшись разрабатывать системы водяных знаков для идентификации ИИ-контента, публично сообщать о возможностях и ограничениях своих систем, а также исследовать социальные риски, связанные с ИИ.

В США в 2023 году Белый дом и 7 крупнейших американских технологических компаний — Amazon, Anthropic, Google, Inflection, Meta, Microsoft и OpenAI — заключили соглашение «Обеспечение безопасного, защищенного и заслуживающего доверия ИИ», где взяли на себя обязательство о саморегулировании для управления рисками, связанными с ИИ⁷. Среди ключевых мер: разработка системы водяных знаков для идентификации контента, созданного с помощью ИИ, что является частью борьбы с дезинформацией и другими рисками, связанными с этой быстро развивающейся технологией; обязательство публично сообщать о возможностях, ограничениях и областях использования своих систем ИИ; расставлять приоритеты в исследованиях социальных рисков, связанных с ИИ, включая предвзятость, дискриминацию и проблемы с конфиденциальностью; разрабатывать системы ИИ для решения общественных проблем, начиная от профилактики рака и заканчивая изменения климата. В конце 2023 года еще восемь IT-компаний подписались на эти добровольные обязательства.

4.4. Российская Федерация. В Российской Федерации регулирование ИИ находится на этапе формирования нормативно-правовой базы, что отражает глобальные тренды в области технологического развития и этических стандартов. В 2020 году была утверждена Национальная стратегия развития искусственного интеллекта до 2030 года, которая определяет основные направления развития ИИ, включая создание правовых и этических рамок. В рамках этой стратегии акцент делается на обеспечение безопасности, защиту персональных данных и предотвращение дискриминации при использовании ИИ-технологий. Кроме того, в 2021 году был принят Федеральный закон № 331-ФЗ, который вносит изменения в Гражданский кодекс РФ и регулирует

⁶Code civil. URL: <https://www.legifrance.gouv.fr/affichCode.do?cidTexte=LEGITEXT000006070721> (дата обращения: 27.01.2025).

⁷The White House. FACT SHEET: Biden-Harris Administration Secures Voluntary Commitments from Leading Artificial Intelligence Companies to Manage the Risks Posed by AI. 2023. URL: <https://www.whitehouse.gov> (дата обращения: 27.01.2025).

вопросы ответственности за действия автономных систем, включая ИИ. Однако на данный момент правовое регулирование остается фрагментарным и требует дальнейшей детализации, особенно в вопросах определения ответственности за решения, принимаемые ИИ. Необходимо отметить также, что в настоящее время происходит значительная работа по формированию стандартов в области использования систем ИИ, и здесь одним из лидеров является здравоохранение, где активно развивается правовое и техническое регулирование ИИ (см., подробнее [9, глава 4]).

Одним из ключевых аспектов регулирования ИИ в России является этическая составляющая. В 2021 году был разработан Кодекс этики в области ИИ, который устанавливает принципы разработки и использования ИИ, такие как прозрачность, справедливость и уважение прав человека. Этот документ, хотя и не имеет обязательной юридической силы, служит ориентиром для разработчиков и компаний, работающих в сфере ИИ. Тем не менее отсутствие четких механизмов контроля и санкций за нарушение этических норм остается проблемой.

5. Этические принципы регулирования ИИ и нейротехнологий с позиции Церкви

Из вышесказанного очевидно, что развитие ИИ и нейротехнологий ставит перед обществом сложные этические вопросы, которые не остаются без внимания религиозных институтов. Церковь на протяжении веков играла ключевую роль в формировании моральных и этических норм общества. Ее позиция по вопросам, связанным с технологиями, имеет значительный вес, так как она основывается на глубоких философских и теологических традициях, которые рассматривают человека как высшую ценность. Включение позиции Церкви в дискуссию об ИИ и нейротехнологиях позволяет учесть духовно-нравственные аспекты, которые часто упускаются в технических и научных обсуждениях. Церковь подчеркивает, что технологии не должны нарушать достоинство человека, которое рассматривается как божественный дар. Это особенно важно в контексте нейротехнологий, которые могут влиять на сознание и идентичность личности. Кроме того, обсуждение позиции Церкви в статье будет способствовать диалогу между научным и религиозным сообществами, что позволит найти баланс между инновациями и сохранением человеческих ценностей в условиях, когда современные технологии начинают затрагивать вопросы, которые традиционно находились в сфере морали и философии.

В данном разделе кратко рассмотрим позиции Русской Православной Церкви (РПЦ) и Римско-католической Церкви (РКЦ) в отношении ИИ и нейротехнологий, а также их подходы к этическому регулированию. Особое внимание уделяется документам и заявлениям, которые формируют основу для диалога между религией и технологиями

5.1. Этические принципы Церкви в отношении ИИ и нейротехнологий. Церковь подчеркивает, что технологии должны служить человеку, а не подчинять его себе, и предлагает ряд ключевых этических принципов для регулирования ИИ и нейротехнологий. Во-первых, она настаивает на приоритете человеческого достоинства, которое рассматривается как божественный дар. Это означает, что технологии не должны использоваться для манипуляции сознанием, вмешательства в личную свободу или нарушения автономии личности. Вместо этого они должны быть направлены на улучшение качества жизни, избегая рисков дегуманизации. Во-вторых, Церковь подчеркивает ответственность разработчиков и пользователей технологий за их последствия. Особое внимание уделяется нейротехнологиям, которые могут существенно сильнее влиять на сознание и идентичность человека, чем технологии ИИ. Для предотвращения злоупотреблений Церковь призывает к созданию механизмов прозрачности и подотчетности новых технологий перед обществом.

Еще одним важным принципом является защита приватности и данных. Церковь выступает за строгое регулирование сбора и использования персональной информации, особенно в контексте

нейротехнологий, чтобы предотвратить утечки данных и их использование в манипулятивных целях. Кроме того, Церковь предупреждает об опасности дегуманизации, когда технологии заменяют человеческие отношения, эмоции и духовный опыт. Она настаивает на том, что технологии должны дополнять, а не подменять человеческую природу. Наконец, Церковь призывает к справедливости и равенству, подчеркивая, что технологии должны быть доступны всем слоям общества и не должны усугублять социальное неравенство или дискриминацию. Эти принципы формируют основу для этического регулирования технологий, направленного на сохранение человеческих ценностей в эпоху цифрового прогресса.

5.2. Позиция Патриарха Кирилла и Русской Православной Церкви. Патриарх Кирилл неоднократно высказывался о рисках, связанных с развитием ИИ и нейротехнологий. В своих обращениях он подчеркивает, что технологии не должны ставить под угрозу духовную природу человека. В частности, Патриарх предупреждает об опасности «цифрового рабства», когда человек становится зависимым от технологий, теряя свою свободу и духовную автономию. «Человеческая мысль, техническая цивилизация сегодня достигли такого уровня, когда, внедряя цифровые технологии, можно обеспечить тотальный контроль над человеческой личностью. Не просто наблюдение за человеком, но управление человеческим поведением», — сказал он в рождественском интервью телеканалу «Россия-1» в 2021 г.⁸

На XXVIII Международных Рождественских образовательных чтениях в 2020 году Патриарх Кирилл отметил, что «технологии должны служить человеку, а не превращать его в объект манипуляций»⁹. Он также подчеркнул необходимость разработки этических норм, которые бы ограничивали использование ИИ в сферах, затрагивающих человеческую свободу и достоинство. Патриарх призвал к созданию международных стандартов регулирования ИИ, которые бы учитывали не только технические, но и духовно-нравственные аспекты.

РПЦ поддерживает идею запрета на использование технологий для манипуляции сознанием и изменения идентичности человека. В своих документах Церковь подчеркивает, что технологии должны быть направлены на укрепление, а не разрушение человеческой природы.

5.3. Позиция Католической Церкви. Ватикан, как и РПЦ, активно участвует в обсуждении этических вопросов, связанных с ИИ и нейротехнологиями. Папа Франциск неоднократно подчеркивал, что технологии должны служить общему благу и не нарушать принципы человеческого достоинства. Основные принципы регулирования ИИ, которые провозглашает Ватикан, можно суммировать как:

- Человеческий контроль над технологиями: Ватикан настаивает на том, что ИИ должен оставаться под контролем человека, а решения, затрагивающие жизнь людей, не должны приниматься автономными системами.
- Защита приватности и данных: Католическая Церковь призывает к строгому регулированию сбора и использования персональных данных, особенно в контексте нейротехнологий.
- Справедливость и инклюзивность: технологии должны быть доступны всем, а их использование не должно приводить к социальному неравенству.
- Духовное измерение: Ватикан подчеркивает, что технологии не должны заменять духовный опыт и человеческие отношения. ИИ и нейротехнологии должны использоваться для укрепления, а не разрушения человеческих связей.

Особое внимание в документах Ватикана уделяется принципу «Antiqua et nova» («Древнее и новое»), который подчеркивает необходимость сочетания традиционных ценностей с новыми

⁸Патриарх Кирилл предостерег от цифрового рабства // Интерфакс : информ. агенство России : сайт. 2021. URL: <https://www.interfax.ru/russia/744269> (дата обращения: 8.02.2025).

⁹Патриарх Кирилл. Выступление на XXVIII Международных Рождественских образовательных чтениях // Отдел внешних церковных связей Московского Патриархата : официальный сайт. 2020. URL: <https://mospat.ru> (дата обращения: 8.02.2025).

технологическими достижениями. Этот принцип предполагает, что технологии должны развиваться в гармонии с человеческой природой и духовными ценностями, а не противоречить им. По инициативе Папской академии защиты жизни и фонда RenAissance в 2020 г. был разработан документ Rome Call — «Римский призыв к этике в области искусственного интеллекта»¹⁰, в котором отражены вышеуказанные принципы.

5.4. Сходство и различие в позициях регулирования ИИ и нейротехнологий Русской Православной Церкви и Римско-католической Церкви. Хотя РПЦ и Ватикан сходятся в основных принципах регулирования ИИ и нейротехнологий, их подходы имеют свои особенности, обусловленные как богословскими традициями, так и культурно-историческим контекстом. Обе Церкви подчеркивают, что технологии должны служить человеку, а не подчинять его себе, однако акценты в их позициях несколько различаются.

Патриарх Кирилл в своих выступлениях делает особый упор на духовные риски, связанные с развитием технологий. Ватикан, в свою очередь, также уделяет внимание духовным аспектам, но делает больший акцент на социальной справедливости и общем благе. Папа Франциск в своих выступлениях и документах, таких как Послание Святого отца Франциска участникам собрания «Этика ИИ за мир»¹¹, подчеркивает, что технологии должны служить всем людям, особенно наиболее уязвимым слоям общества. В документе «Antiqua et nova» Ватикан призывает к гармоничному сочетанию традиционных ценностей с новыми технологическими достижениями, чтобы избежать дегуманизации¹². Например, Папа Франциск отмечает, что ИИ должен использоваться для укрепления человеческих связей, а не для их разрушения.

РПЦ поддерживает создание международных стандартов регулирования ИИ, но с учетом культурных и религиозных особенностей разных стран. Патриарх Кирилл подчеркивает, что универсальные нормы должны быть адаптированы к местным условиям, чтобы учитывать разнообразие культурных и духовных традиций. Например, в своих выступлениях он отмечает, что западные подходы к регулированию технологий не всегда учитывают специфику православного мировоззрения, которое делает акцент на духовной свободе и ответственности перед Богом. Ватикан, напротив, выступает за универсальные этические нормы, которые бы применялись во всем мире. Папа Франциск призывает к международному сотрудничеству в области регулирования ИИ, чтобы предотвратить злоупотребления и обеспечить соблюдение этических принципов. В документе «Этика искусственного интеллекта» Ватикан предлагает создать глобальные стандарты, которые бы защищали человеческое достоинство и обеспечивали справедливость. Например, Ватикан поддерживает инициативы, направленные на запрет использования ИИ в военных целях или для массовой слежки.

В отношении нейротехнологий Патриарх Кирилл более категоричен, предупреждая об опасности манипуляции сознанием. Он считает, что технологии, способные влиять на мозговую активность, представляют угрозу для свободы воли и духовной автономии человека. В своих выступлениях Патриарх призывает к строгому регулированию нейротехнологий, чтобы предотвратить их использование в манипулятивных целях. Например, в слове Патриарха о научно-техническом прогрессе на встрече с учеными во Всероссийском научно-исследовательском институте экспериментальной физики в г. Сарове 31 июля 2019 г. он задается вопросом: «*Не приведут ли разработки нейроинтерфейса к глубинному коллапсу наших глубинных представлений о границах человеческого?*»¹³.

¹⁰The Call for AI Ethics // RenAissance: фонд: официальный сайт. Ватикан, 2020. URL: <https://www.romecall.org/the-call/> (дата обращения: 8.02.2025).

¹¹Message of the Holy Father to the participants in the «AI Ethics for Peace» gathering // The Holy See: официальный сайт. Ватикан, 2024. URL: <https://www.vatican.va/content/francesco/en/messages/pont-messages/2024/documents/20240710-messaggio-ai-ethics-forpeace.html> (дата обращения: 8.02.2025).

¹²ANTIQUA ET NOVA. Note on the Relationship Between Artificial Intelligence and Human Intelligence // The Holy See: официальный сайт. Ватикан, 2025. URL: https://www.vatican.va/roman_curia/congregations/cfaith/documents/rc_ddf_doc_20250128_antiqua-et-nova_en.html (дата обращения: 8.02.2025).

¹³Слово Патриарха о научно-техническом прогрессе // Православие.ru: российский православный информационный

Ватикан, хотя и разделяет эти опасения, больше фокусируется на защите приватности и данных. В документе «Этика искусственного интеллекта» подчеркивается, что сбор и использование данных, особенно в контексте нейротехнологий, должны быть строго регламентированы. Ватикан призывает к созданию правовых рамок, которые бы защищали приватность и предотвращали использование данных в манипулятивных целях. Например, Папа Франциск поддерживает запрет на использование нейротехнологий для слежки или контроля над людьми.

Однако в своем последнем выступлении в Госдуме лидер Русской Православной Церкви говорит уже об экзистенциальной опасности, заявляя, что ИИ «опаснее, чем ядерная энергия», «и еще не известно, какой может быть роль искусственного интеллекта в приближении этой апокалиптической картины».

Таким образом, хотя РПЦ и РКЦ сходятся в основных принципах, их подходы к регулированию ИИ и нейротехнологий имеют свои особенности. РПЦ делает акцент на духовных рисках и культурной специфике, в то время как Ватикан фокусируется на социальной справедливости и универсальных нормах. Оба подхода дополняют друг друга, предлагая комплексное видение этического регулирования технологий с гуманитарных христианских позиций.

6. Согласование этических норм при коэволюции ИИ и нейротехнологий

6.1. Осознание рисков использования ИИ и нейротехнологий. Анализ этических проблем и дилемм, рассмотренных в разделе 3, позволяет сделать вывод, что ИИ и нейротехнологии ставят схожие этические вызовы, особенно в вопросах согласования ценностей и норм при их разработке и использовании. В ряде случаев решения, разработанные для ИИ, могут быть адаптированы для нейротехнологий. Однако нейротехнологии сопряжены с уникальными рисками, которые могут не охватываться существующими нормативными актами или подходами, применяемыми в ИИ.

Например, вопросы психической конфиденциальности, идентичности, автономии и самостоятельности человека становятся более сложными в контексте нейротехнологий. В отличие от ИИ, который влияет на человека косвенно, нейротехнологии могут как напрямую собирать нейроданные, так и записывать новую информацию в нервную систему, причем потенциально незаметно для пользователя [161, 162]. Большинство нейронных сигналов бессознательны и не поддаются нашему осознанию или контролю, что затрудняет точное определение типов данных, собираемых или модулируемых нейротехнологиями. Это создает сложности в понимании того, какими данными мы делимся, и повышает риски непреднамеренного предоставления нейроданных. Это справедливо и для большинства данных о нашем поведении или наших предпочтениях, привычках и т. д., к которым сегодня имеют доступ системы ИИ, но мы редко обращаем на это значительное внимание, потому что до сих пор существует презумпция приватности собственного разума, который кажется недоступным для других людей или технологий. Однако с развитием нейротехнологий это предположение может перестать быть уверенностью. Более того, тот факт, что некоторые нейротехнологии могут как напрямую изменять текущую нейронную активность и напрямую передавать данные в нервную систему в режиме реального времени, поднимает вопросы о том, как мы можем лучше защитить и обеспечить телесную/психическую автономию и способность принимать решения. Это включает в себя возможность целенаправленного изменения целостности наших психических процессов, включая наши представления об идентичности.

Нейротехнологии, способные напрямую влиять на нейронную активность, поднимают вопросы о защите телесной и психической автономии. Возможность изменения психических процессов, включая идентичность, эмоции и воспоминания, ставит сложные вопросы о свободе воли, когнитивной свободе и самосознании. Эти вызовы выходят за рамки тех, с которыми

интернет-портал. Москва, 2019. URL: <http://www.pravoslavie.ru/122890.html> (дата обращения: 8.02.2025).

сталкивается ИИ, хотя аналогичные проблемы обсуждались в биоэтике, например, в контексте трансплантации органов и фармацевтических усовершенствований.

Справедливость, являющаяся ключевым принципом этики ИИ, также актуальна для нейротехнологий. Однако прямое воздействие на мозг и доступ к скрытым данным (например, бессознательным предубеждениям и подавленным воспоминаниям) и потенциальное расширение наших когнитивных и физических способностей повышают риски неравенства и дискриминации. Определение «нормативных» нейроданных и «желательных» результатов может привести к предвзятости в отношении людей с неврологическими различиями или инвалидностью, создавая новые социальные и культурные дилеммы.

Многие из описанных проблем усугубляются при взаимодействии нейротехнологий с ИИ. Например, вопросы конфиденциальности и деидентификации данных, актуальные для ИИ, становятся более сложными в случае нейроданных, которые могут использоваться для точной идентификации личности [159, 163]. Это выходит за рамки традиционных персональных данных, используемых для идентификации личности, и требует новых подходов к защите информации.

Таким образом, этические вызовы, связанные с нейротехнологиями, представляются более острыми и многогранными, чем в случае ИИ, что требует разработки специализированных нормативных и технологических решений.

6.2. Пути решения этических проблем, связанных с совместным использованием ИИ и нейротехнологий. Для эффективного решения проблем, возникающих на стыке ИИ и нейротехнологий, требуется комплексный подход, учитывающий как их общие черты, так и специфические вызовы. В табл. 2 мы, опираясь на результаты работы [55], подвели итог сравнения этических проблем между нейротехнологиями и ИИ, которые подробно были описаны выше.

Таблица 2. Сравнение этических проблем и рисков ИИ и нейротехнологий
Table 2. Comparing the ethical issues and risks of AI and neurotechnology

Этическая проблема и/или риск	Различие между ИИ и нейротехнологиями	Обсуждение
Проблемы с данными и датасетами	Значительные	Данные нейровизуализаций обрабатываются с использованием ИИ, как и любые другие данные, но могут запросить дополнительные требования из-за своей сложности и специфичности форматов, конфиденциальности и отсутствия стандартных согласованных моделей управления и распространения
Объяснимость и доверие	Значительные	В отличие от ИИ, который воздействует на нашу когнитивную сферу опосредованно, нам меньше известно на сегодняшний день о непосредственном и отложенном воздействии нейротехнологий на нашу когнитивную сферу
Подотчетность	Небольшие	Ограниченный и специфический (в основном в области медицинской диагностики) опыт в интерпретации нейроданных, плохая повторяемость научных результатов
Справедливость и доступность	В настоящее время — небольшие. В будущем могут стать значительными.	В отличие от ИИ нейротехнологии в перспективе могут оказывать более значительное и глубокое влияние на человека. Например, воздействия на мозг, о которых мы не знаем и которые не контролируем, могут привести к дискриминации; потенциальная возможность расширения способностей за пределы того, что в настоящее время доступно человеку, может привести к усугублению неравенства.

Окончание табл. 2 на стр. 139

Профилирование и манипуляции	Небольшие	В отличие от манипуляций, которые может породить генеративный ИИ, нейротехнологии для этих целей могут расширить набор дополнительных данных личного характера (данные об особенностях нервной системы и личности), которые человек не может контролировать и скрывать
Контроль и согласование ценностей	Одинаковые	—
Ментальная конфиденциальность	Значительные	ИИ имеет доступ только к внешним прокси-данным мыслей и поведения, а не к нейроданным, которые могут быть прямыми репрезентативными показателями психических явлений.
Автономия личности и агентность	В настоящее время — небольшие. В будущем могут стать значительными.	Нейротехнологии могут «внедряться» в нашу нервную систему и напрямую влиять на наш разум и способности, часто в режиме реального времени.
Человеческая идентичность	В настоящее время — небольшие. В будущем могут стать значительными.	ИИ имеет доступ только к внешним носителям идентичности — текстам, данным, особенностям поведения (не к нейроданным) — и может лишь косвенно влиять на идентичность человека, тогда как нейротехнологии способны влиять на нее напрямую и делать это способами, которые могут быть контролируемыми или легко измеримыми, а могут и нет.
Прозрачность	Одинаковые	—
Безопасность	Одинаковые	В будущем могут стать значительными. Нейроданные могут иметь дело с более личной или конфиденциальной информацией; вероятно, в будущем появится возможность получить большой объем информации о личности человека на основе нейроданных
Благополучие	Значительные	Постоянная физическая и психологическая безопасность могут быть важными факторами для некоторых инвазивных нейротехнологических устройств (имплантов); кроме того, некоторые нейротехнологические возможности зависят от химических веществ, которые могут не поддаваться биологическому разложению или со временем накапливаться в организме, нанося вред здоровью.
Влияние на общество	Одинаковые	В будущем могут стать значительными. При объединении ИИ и нейротехнологий появляется больше возможностей для практического применения, что может в перспективе значительно повлиять на сообщества и социум в целом

Конечно, надо иметь в виду, что данные, представленные в таблице, не являются исчерпывающими, так как по мере развития сквозных технологий могут появляться дополнительные факторы и различия.

Определив проблемы, связанные с совместным использованием ИИ и нейротехнологий, мы можем начать определение путей их решения. Как уже говорилось, некоторые проблемы при комбинировании систем ИИ и нейротехнологий потенциально могут стать значительно более сложными по сравнению только с ИИ, поэтому для их решения нам могут понадобиться модифицированные или новые стратегии снижения рисков или их предотвращения, как чисто

технические, так и социальные, политические, институциональные и экономические подходы. Однако на этом пути не нужно начинать с нуля. За последнее время было проделано много фундаментальной работы, чтобы начать решать многие вопросы этики ИИ, включая разработку методологий определения проблем, лучших практик, этических принципов и руководящих указаний, создание технических решений, образовательных программ, а также формирование рамок управления, международных стандартов и законодательных актов. Уроки, извлеченные из этих усилий, должны быть использованы для ускорения процесса решения проблем нейротехнологий. Международное сообщество нейроэтиков создало более 20 этических руководств, принципов и лучших практик, скомпилированных Институтом нейроэтики (IoNx)¹⁴, которые могут стать основой этой работы.

Первым шагом является четкое определение взаимосвязи между общими и специфическими проблемами, представленными в табл. 2, после чего вторым шагом является адаптация существующих этических рамок и мер по регулированию ИИ к более широкому кругу проблем, связанных с коэволюцией ИИ и нейротехнологий. Для полного понимания последствий нейротехнологий необходим междисциплинарный подход, включающий экспертов из нейронауки, нейроэтики, социологии, антропологии, медицины, права, бизнеса и других областей. Важно привлекать к обсуждению не только специалистов, но и представителей сообществ, которые могут быть затронуты ИИ и нейротехнологиями, включая тех, кто потенциально может пострадать от новых технологий. К последним, например, могут быть отнесены представители малых народностей, данные которых могут не попадать в датасеты, формируемые для обучения медицинского ИИ, и которые могут пострадать из-за неверного диагноза из-за проблем несбалансированности данных [164]. Их потребности должны рассматриваться как важный фактор для оценки этических последствий коэволюции ИИ и нейротехнологий и создания более совершенных и безопасных технологий.

Заключение

ИИ представляет собой одну из наиболее значимых технологий современности, активно интегрирующуюся в различные сферы человеческой жизни. В сочетании с нейротехнологиями ИИ открывает новые перспективы в области профессиональной деятельности, досуга, здравоохранения и других социально значимых направлений. Однако, как и любая технология, обладающая значительным потенциалом, ИИ требует ответственного подхода к своему развитию и применению. Коммерциализация ИИ и нейротехнологий, а также использование данных, лежащих в основе их функционирования, происходят с беспрецедентной скоростью и масштабами. В отсутствие адекватной политики и регуляторных механизмов это может привести к возникновению серьезных угроз для человечества, даже без достижения уровня «сверхинтеллекта».

Технологии, оказывающие непосредственное влияние на когнитивную сферу человека, к которым относятся как ИИ, так и нейротехнологии, должны развиваться в соответствии с темпами осмысления и интеграции человеческих ценностей и общественных интересов. Это необходимо для того, чтобы технологический прогресс приносил пользу не только отдельным индивидам, но и обществу в целом. Учитывая, что рассматриваемые технологии находятся на стадии активного развития и их прогресс, вероятно, продолжится в ближайшем будущем, существует возможность извлечь уроки из прошлого опыта, заранее идентифицировать потенциальные риски и разработать превентивные меры. К ним относятся технические, правовые, социальные и образовательные решения, направленные на минимизацию негативных последствий до их возникновения. Такой подход позволит обеспечить устойчивое и этически обоснованное развитие технологий, способствуя их гармоничному внедрению в общественную жизнь.

¹⁴Institute of Neuroethics (IoNx): сайт. URL: <https://instituteofneuroethics.org> (дата обращения: 4.02.2025).

Уже на текущем этапе развития становится очевидным, что ключевым принципом регулирования ИИ и нейротехнологий, включая различные методы нейровизуализации, должен стать принцип *in dubio pro homo* (*в сомнении — в пользу человека*). Этот принцип предполагает:

- защиту человека от дискриминации, обусловленной алгоритмическими решениями;
- обеспечение права выбора человека в качестве альтернативы автоматизированным системам;
- приоритет личной автономии и самоопределения индивида;
- гарантии защиты персональных данных, включая нейроданные.

В контексте генеративного ИИ и нейротехнологий, связанных с воздействием на мозг (в частности, методов нейромодуляции), необходимо законодательное закрепление следующих требований:

- обязательная маркировка контента, созданного с использованием ИИ;
- запрет на создание противозаконного контента;
- запрет на разработку систем, воздействующих на подсознание человека и нарушающих его базовые свободы и автономию.

В данной статье мы постарались обратить внимание на козволюцию ИИ и нейротехнологий, а также на потенциальные точки их взаимодействия, что особенно актуально для специалистов, занимающихся вопросами этики технологий. Наша цель заключается в иницировании широкой дискуссии и определении путей решения этических проблем, возникающих на стыке этих технологий.

Путем выявления ключевых этических вызовов, связанных с нейротехнологиями, их сопоставления с этическими проблемами ИИ, а также анализа областей, где существующие инициативы и инструменты в области этики ИИ могут быть достаточными или, напротив, требуют новых подходов, мы надеемся стимулировать междисциплинарное сотрудничество. Такое взаимодействие между учеными, представителями бизнеса и политическими деятелями позволит разработать своевременные и конкретные меры, направленные на минимизацию негативного воздействия новых технологий и обеспечение их ответственного использования в интересах общества.

Список литературы

1. *Hassabis D., Kumaran D., Summerfield C., Botvinick M.* Neuroscience-inspired artificial intelligence // *Neuron*. 2017. Vol. 95, no. 2. P. 245–258. DOI: 10.1016/j.neuron.2017.06.011.
2. *Karpov O. E., Pitsik E. N., Kurkin S. A., Maksimenko V. A., Gusev A. V., Shusharina N. N., Hramov A. E.* Analysis of publication activity and research trends in the field of AI medical applications: Network approach // *Int. J. Environ. Res. Public Health*. 2023. Vol. 20, no. 7. P. 5335. DOI: 10.3390/ijerph20075335.
3. *Nyholm S.* Artificial intelligence and human enhancement: Can AI technologies make us more (artificially) intelligent? // *Camb. Q. Healthc. Ethics*. 2024. Vol. 33, no. 1. P. 76–88. DOI: 10.1017/S0963180123000464.
4. *Duan Y. Q., Edwards J. S., Dwivedi Y. K.* Artificial intelligence for decision making in the era of Big Data – evolution, challenges and research agenda // *International Journal of Information Management*. 2019. Vol. 48. P. 63–71. DOI: 10.1016/j.ijinfomgt.2019.01.021.
5. *Autor D. H.* Why are there still so many jobs? The history and future of workplace automation // *Journal of Economic Perspectives*. 2015. Vol. 29, no. 3. P. 3–30. DOI: 10.1257/jep.29.3.3.
6. *Brown T., Mann B., Ryder N., Subbiah M., Kaplan J. D., Dhariwal Prafulla, Neelakantan A., Shyam P., Sastry G., Askell A., Agarwal S., Herbert-Voss A., Krueger G., Henighan T., Child R., Ramesh A., Ziegler D., Wu J., Winter C., Hesse Chris, Chen M., Sigler E., Litwin M., Gray S., Chess B., Clark J., Berner Ch., McCandlish S., Radford A., Sutskever I., Amodei D.* Language models are few-shot learners // In: Larochelle H., Ranzato M., Hadsell R., Balcan M. F., Lin H. (eds.) *Advances in Neural Information Processing Systems*. Vol. 33. P. 1877–1901. DOI: 10.48550/arXiv.2005.14165.

7. Wang L., Törngren M., Onori M. Current status and advancement of cyber-physical systems in manufacturing // *Journal of Manufacturing Systems*. 2015. Vol. 37. P. 517–527. DOI: 10.1016/j.jmsy.2015.04.008.
8. Topol E. J. High-performance medicine: the convergence of human and artificial intelligence // *Nature Medicine*. 2019. Vol. 25, no. 1. P. 44–56. DOI: 10.1038/s41591-018-0300-7.
9. Карнов О. Э., Храмов А. Е. Информационные технологии, вычислительные системы и искусственный интеллект в медицине. М.: ДПК Пресс, 2022. 480 с.
10. Федоров А. А., Куркин С. А., Храмова М. В., Храмов А. Е. Нейротехнологии и искусственный интеллект как ключевые факторы кастомизации жизненно-образовательного маршрута // *Информатика и образование*. 2023. Т. 38, № 3. С. 5–15. DOI: 10.32517/0234-0453-2023-38-3-5-15.
11. Holmes W., Tuomi I. State of the art and practice in AI in education // *European Journal of Education*. 2022. Vol. 57, no. 4. P. 542–570. DOI: 10.1111/ejed.12533.
12. Carleo G., Cirac I., Cranmer K., Daudet L., Schuld M., Tishby N., Vogt-Maranto L., Zdeborová L. Machine learning and the physical sciences // *Rev Mod Phys*. 2019. Vol. 91, no. 4. P. 045002. DOI: 10.1103/RevModPhys.91.045002.
13. Silver D., Hubert T., Schrittwieser J., Antonoglou I., Lai M., Guez A., Lanctot M., Sifre L., Kumaran D., Graepel T., Lillicrap T., Simonyan K., Hassabis D. A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play // *Science*. 2018. Vol. 362, no. 6419. P. 1140–1144. DOI: 10.1126/science.aar6404.
14. Bandettini P. A. What’s new in neuroimaging methods? // *Ann NY Acad Sci*. 2009. Vol. 1156. P. 260–293. DOI: 10.1111/j.1749-6632.2009.04420.x.
15. Храмов А. Е., Максименко В. А., Писарчик А. Н. Physical principles of brain-computer interfaces and their applications for rehabilitation, robotics and control of human brain states // *Phys Rep*. 2021. Vol. 918. P. 1–133. DOI: 10.1016/j.physrep.2021.03.002.
16. Marzbani H., Marateb H. R., Mansourian M. Neurofeedback: a comprehensive review on system design, methodology and clinical applications // *Basic Clin. Neurosci*. 2016. Vol. 7, no. 2. P. 143–158. DOI: 10.15412/J.BCN.03070208.
17. Mihara M., Miyai I., Hattori N., Hatakenaka M., Yagura H., Kawano T., Okibayashi M., Danjo N., Ishikawa A., Inoue Y., Kubota K. Neurofeedback using real-time near-infrared spectroscopy enhances motor imagery related cortical activation // *PLoS ONE*. 2012. Vol. 7, no. 3. P. e32234. DOI: 10.1371/journal.pone.0032234.
18. Hallett M. Transcranial magnetic stimulation: a primer // *Neuron*. 2007. Vol. 55, no. 2. P. 187–199. DOI: 10.1016/j.neuron.2007.06.026.
19. Rossi S., Hallett M., Rossini P., Pascual-Leone A., *The Safety of TMS Consensus Group*. Safety, ethical considerations, and application guidelines for the use of transcranial magnetic stimulation in clinical practice and research // *Clin Neurophysiol*. 2009. Vol. 120, no. 12. P. 2008–2039. DOI: 10.1016/j.clinph.2009.08.016.
20. Brunoni A., Nitsche M., Bolognini N., Bikson M., Wagner T., Merabet L., Edwards D., Valero-Cabre A., Rotenberg A., Pascual-Leone A., Ferrucci R., Priori A., Boggio P., Fregni F. Clinical research with transcranial direct current stimulation (tDCS): Challenges and future directions // *Brain Stimul*. 2012. Vol. 5, no. 3. P. 175–195. DOI: 10.1016/j.brs.2011.03.002.
21. Benabid A., Chabardes S., Mitrofanis J., Pollak P. Deep brain stimulation of the subthalamic nucleus for the treatment of Parkinson’s disease // *Lancet Neurol*. 2009. Vol. 8, no. 1. P. 67–81. DOI: 10.1016/S1474-4422(08)70291-6.
22. Cometa A., Falasconi A., Biasizzo M., Carpaneto J., Horn A., Mazzoni A., Micera S. Clinical neuroscience and neurotechnology: An amazing symbiosis // *iScience*. 2022. Vol. 25, no. 10. P. 105124. DOI: 10.1016/j.isci.2022.105124.
23. Daly J. J., Wolpaw J. R. Brain-computer interfaces in neurological rehabilitation // *Lancet Neurol*. 2008. Vol. 7, no. 11. P. 1032–1043. DOI: 10.1016/S1474-4422(08)70223-0.

24. *Plass-Oude Bos D., Reuderink B., van de Laar B., Gürkök H., Mühl C., Poel M., Nijholt A., Heylen D.* Brain-Computer Interfacing and Games // In: Tan D., Nijholt A. (eds) Brain-Computer Interfaces. Human-Computer Interaction Series. London: Springer, 2010. P. 149–178. DOI: 10.1007/978-1-84996-272-8_10.
25. *Marshall D., Coyle D., Wilson S., Callaghan M.* Games, gameplay, and BCI: The state of the art // IEEE Transactions on Computational Intelligence and AI in Games. 2013. Vol. 5, no. 2. P. 164–176. DOI: 10.1109/tciaig.2013.2263555.
26. *Campbell M., Hoane A. J., Hsu Fh.* Deep Blue // Artificial Intelligence. 2002. Vol. 134, no. 1–2. P. 57–83. DOI: 10.1016/S0004-3702(01)00129-1.
27. *Hsu F. H., Anantharaman T., Campbell M., Nowatzyk A.* A grandmaster chess machine // Scientific American. 1990. Vol. 263, no. 4. P. 44–51. <https://www.jstor.org/stable/24997060>.
28. *He K., Zhang X., Ren S., Sun J.* Delving deep into rectifiers: Surpassing human-level performance on imagenet classification // In: Proceedings of the IEEE International Conference on Computer Vision (ICCV). 2015. p. 1026–1034. DOI: 10.1109/ICCV.2015.123.
29. *Silver D., Schrittwieser J., Simonyan K., Antonoglou I., Huang A., Guez A., Hubert T., Baker L., Lai M., Bolton A., Chen Y., Lillicrap T., Hui F., Sifre L., van den Driessche G., Graepel Th., Hassabis D.* Mastering the game of go without human knowledge // Nature. 2017. Vol. 550, no. 7676. P. 354–359. DOI: 10.1038/nature24270.
30. *Logothetis N. K.* What we can do and what we cannot do with fMRI // Nature. 2008. Vol. 453, no. 7197. P. 869–878. DOI: 10.1038/nature06976.
31. *Luck S. J.* An Introduction to the Event-Related Potential Technique. Massachusetts: MIT Press, 2014. 388 p.
32. *Hämäläinen M., Hari R., Ilmoniemi R. J., Knuutila J., Lounasmaa O. V.* Magnetoencephalography-theory, instrumentation, and applications to noninvasive studies of the working human brain // Rev. Mod. Phys. 1993. Vol. 65, no. 2. P. 413–497. DOI: 10.1103/RevModPhys.65.413.
33. *Ferrari M., Quaresima V.* A brief review on the history of human functional near-infrared spectroscopy (fNIRS) development and fields of application // NeuroImage. 2012. Vol. 63, no. 2. P. 921–935. DOI: 10.1016/j.neuroimage.2012.03.049.
34. *Grau C., Ginhoux R., Riera A., Nguyen T. L., Chauvat H., Berg M., Amengual J. L., Pascual-Leone A., Ruffini G.* Conscious brain-to-brain communication in humans using non-invasive technologies // PLoS ONE. 2014. Vol. 9, no. 8. P. e105225. DOI: 10.1371/journal.pone.0105225.
35. *Nam C. S., Traylor Z., Chen M., Jiang X., Feng W., Chhatbar P. Y.* Direct communication between brains: A systematic PRISMA review of brain-to-brain interface // Front Neurobot. 2021. Vol. 15. P. 656943. DOI: 10.3389/fnbot.2021.656943.
36. *Maksimenko V. A., Hramov A. E., Frolov N. S., Lüttjohann A., Nedaivozov V. O., Grubov V. V., Runnova A. E., Makarov V. V., Kurths J., Pisarchik A. N.* Increasing human performance by sharing cognitive load using brain-to-brain interface // Front Neurosci. 2018. Vol. 12. P. 949. DOI: 10.3389/fnins.2018.00949.
37. *Kurkin S., Gordleeva S., Savosenkov A., Grigorev N., Smirnov N., Grubov V. V., Udoratina A., Maksimenko V., Kazantsev V., Hramov A. E.* Transcranial Magnetic Stimulation of the Dorsolateral Prefrontal Cortex Increases Posterior Theta Rhythm and Reduces Latency of Motor Imagery // Sensors. 2023. Vol. 23, no. 10. P. 4661. DOI: 10.3390/s23104661.
38. *Musk E.* An integrated brain-machine interface platform with thousands of channels // J. Med. Internet Res. 2019. Vol. 21, no. 10. P. e16194. DOI: 10.2196/16194.
39. *Pisarchik A. N., Maksimenko V. A., Hramov A. E.* From novel technology to novel applications: Comment on “An Integrated Brain-Machine Interface Platform With Thousands of Channels” by Elon Musk and Neuralink // J. Med. Internet. Res. 2019. Vol. 21, no. 10. P. e16356. DOI: 10.2196/16356.
40. *Opie N. L., John S. E., Rind G. S., Ronayne S. M., Wong Y. T., Gerboni G., Yoo P. E., Lovell T., Scordas T., Wilson S. L., Dornom A., Vale T., O’Brien T. J., Grayden D. B., May C. N., Oxley Th. J.*

- Focal stimulation of the sheep motor cortex with a chronically implanted minimally invasive electrode array mounted on an endovascular stent // *Nat. Biomed. Eng.* 2018. Vol. 2, no. 12. P. 907–914. DOI: 10.1038/s41551-018-0321-z.
41. *Филипова И. А.* Нейротехнологии: развитие, применение на практике и правовое регулирование // *Вестник СПбГУ. Право.* 2021. Т. 12, № 3. С. 502–521. DOI: 10.21638/spbu14.2021.302.
 42. *Lecomte P.* Umwelt as the foundation of an ethics of smart environments // *Humanit Soc Sci Commun.* 2023. Vol. 10. P. 925. DOI: 10.1057/s41599-023-02356-9.
 43. *Ienca M., Andorno R.* Towards new human rights in the age of neuroscience and neurotechnology // *Life Sci Soc Policy.* 2017. Vol. 13, no. 1. P. 5. DOI: 10.1186/s40504-017-0050-1.
 44. *Kellmeyer P.* Big brain data: On the responsible use of brain data from clinical and consumer-directed neurotechnological devices // *Neuroethics.* 2018. Vol. 14. P. 83–98. DOI: 10.1007/s12152-018-9371-x.
 45. *Mittelstadt B. D., Allo P., Taddeo M., Wachter S., Floridi L.* The ethics of algorithms: Mapping the debate // *Big Data and Society.* 2016. Vol. 3, no. 2. P. 2053951716679679. DOI: 10.1177/2053951716679679.
 46. *Floridi L., Cows J., Beltrametti M., Chatila R., Chazerand P., Dignum V., Luetge C., Madelin R., Pagallo U., Rossi F., Schafer B., Valcke P., Vayena E.* AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations // *Minds Mach.* 2018. Vol. 28, no. 4. P. 689–707. DOI: 10.1007/s11023-018-9482-5.
 47. *Calo R.* Artificial Intelligence Policy: A Primer and Roadmap // *UCD L Rev.* 2017. Vol. 51. P. 399.
 48. *Battista D.* Political communication in the age of artificial intelligence: an overview of deepfakes and their implications // *Society Register.* 2024. Vol. 8, no. 2. P. 7–24. DOI: 10.14746/sr.2024.8.2.01.
 49. *Gambin A. F., Yazidi A., Vasilakos A., Haugerud H., Djenouri Y.* Deepfakes: current and future trends // *Artif Intell Rev.* 2024. Vol. 57, no. 3. P. 64. DOI: 10.1007/s10462-023-10679-x.
 50. *Chesney R., Citron D.* Deepfakes and the new disinformation war: The coming age of post-truth geopolitics // *Foreign Affairs.* 2019. Vol. 98, no. 1. P. 147–155.
 51. *Maras M. H., Alexandrou A.* Determining authenticity of video evidence in the age of artificial intelligence and in the wake of Deepfake videos // *The International Journal of Evidence and Proof.* 2019. Vol. 23, no. 3. P. 255–262. DOI: 10.1177/1365712718807226.
 52. *Citron D. K., Chesney R.* Deep fakes: A looming challenge for privacy, democracy, and national security // *California Law Review.* 2019. Vol. 107, no. 6. P. 1753–1819.
 53. *Memarian B., Doleck T.* ChatGPT in education: Methods, potentials and limitations // *Computers in Human Behavior: Artificial Humans.* 2023. Vol. 1, no. 6. P. 100022. DOI: 10.1016/j.chbah.2023.100022.
 54. *Adeshola I., Adepoju A. P.* The opportunities and challenges of ChatGPT in education // *Interactive Learning Environments.* 2024. Vol. 32, no. 10. P. 6159–6172. DOI: 10.1080/10494820.2023.2253858.
 55. *Berger S., Rossi F.* AI and neurotechnology: learning from AI ethics to address an expanded ethics landscape // *Communications of the ACM.* 2023. Vol. 66, no. 3. P. 58–68. DOI: 10.1145/3529088.
 56. *Huang C., Zhang Z., Mao B., Yao X.* An overview of artificial intelligence ethics // *IEEE Transactions on Artificial Intelligence.* 2022. Vol. 4, no. 4. P. 799–819. DOI: 10.1109/TAI.2022.3194503.
 57. *Bostrom N.* Ethical issues in advanced artificial intelligence // In: *Machine Ethics and Robot Ethics.* London: Routledge, 2020. P. 69–75. DOI: 10.4324/9781003074991-7.
 58. *Mittelstadt B.* Principles alone cannot guarantee ethical AI // *Nat Mach Intell.* 2019. Vol. 1, no. 11. P. 501–507. DOI: 10.1038/s42256-019-0114-4.
 59. *Ruiz S., Valera L., Ramos P., Sitaram R.* Neurorights in the Constitution: from neurotechnology to ethics and politics // *Philosophical Transactions B.* 2024. Vol. 379, no. 1915. P. 20230098. DOI: 10.1098/rstb.2023.0098.

60. *Robinson J. T., Rommelfanger K. S., Anikeeva P. O., Etienne A., French J., Gelinis J., Grover P., Picard R.* Building a culture of responsible neurotech: Neuroethics as socio-technical challenges // *Neuron*. 2022. Vol. 110, no. 13. P. 2057–2062. DOI: 10.1016/j.neuron.2022.05.005.
61. *Illes J.* Neuroethics: Anticipating the Future. NY: Oxford University Press, 2017.
62. *Sample M., Racine E.* Pragmatism for a digital society: The (in)significance of artificial intelligence and neural technology // In: Friedrich O., Wolkenstein A., Bublitz C., Jox R. J., Racine E., editors. *Clinical Neurotechnology meets Artificial Intelligence. Advances in Neuroethics*. Cham: Springer, 2021. p. 81–100. DOI: 10.1007/978-3-030-64590-8_7.
63. *Savage N.* How AI and neuroscience drive each other forwards // *Nature*. 2019. Vol. 571, no. 7766. P. S15–S17. DOI: 10.1038/d41586-019-02212-4.
64. *Raichle M. E.* A brief history of human brain mapping // *Trends Neurosci*. 2009. Vol. 32, no. 2. P. 118–126. DOI: 10.1016/j.tins.2008.11.001.
65. *Friston K. J.* Modalities, modes, and models in functional neuroimaging // *Science*. 2009. Vol. 326, no. 5951. P. 399–403. DOI: 10.1126/science.1174521.
66. *Lefaucheur J. P., Aleman A., Baeken C., Benninger D. H., Brunelin J., Di Lazzaro V., Filipović S. R., Grefkes C., Hasan A., Hummel F. C., Jääskeläinen S. K., Langguth B., Leocani L., Londero A., Nardone R., Nguyen J.-P., Nyffeler Th., Oliveira-Maia A. J., Oliviero A., Padberg F., Palm U., Paulus W., Poulet E., Quartarone A., Rachid F., Rektorová I., Rossi S., Sahlsten H., Schecklmann M., Szekely D., Ziemann U.* Evidence-based guidelines on the therapeutic use of repetitive transcranial magnetic stimulation (rTMS): An update (2014–2018) // *Clin Neurophysiol*. 2020. Vol. 131, no. 5. P. 474–528. DOI: 10.1016/j.clinph.2019.11.002.
67. *Benabid A. L., Chabardes S., Mitrofanis J., Pollak P.* Deep brain stimulation of the subthalamic nucleus for the treatment of Parkinson's disease // *Lancet Neurol*. 2009. Vol. 8, no. 1. P. 67–81. DOI: 10.1016/S1474-4422(08)70291-6.
68. *Fisher R., Salanova V., Witt T., Worth R., Henry T., Gross R., Oommen K., Osorio I., Nazzaro J., Labar D., Kaplitt M., Sperling M., Sandok E., Neal J., Handforth A., Stern J., DeSalles A., Chung S., Shetter A., Bergen D., Bakay R., Henderson J., French J., Baltuch G., Rosenfeld W., Youkilis A., Marks W., Garcia P., Barbaro N., Fountain N., Bazil C., Goodman R., McKhann G., Krishnamurthy K. B., Papavassiliou S., Epstein Ch., Pollard J., Tonder L., Grebin J., Coffey R., Graves N., the SANTE Study Group.* Electrical stimulation of the anterior nucleus of thalamus for treatment of refractory epilepsy // *Epilepsia*. 2010. Vol. 51, no. 5. P. 899–908. DOI: 10.1111/j.1528-1167.2010.02536.x.
69. *Bonomo R., Elia A. E., Bonomo G., Romito L. M., Mariotti C., Devigili G., Cilia R., Giossi R., Eleopra R.* Deep brain stimulation in Huntington's disease: a literature review // *Neurological Sciences*. 2021. Vol. 42, no. 11. P. 4447–4457. DOI: 10.1007/s10072-021-05527-1.
70. *Ramachandran V. S., Rogers-Ramachandran D.* Phantom limbs and neural plasticity // *Arch. Neurol*. 2000. Vol. 57, no. 3. P. 317–320. DOI: 10.1001/archneur.57.3.317.
71. *Khorev V., Kurkin S., Badarin A., Antipov V., Pitsik E., Andreev A., Grubov V., Drapkina O., Kiselev A., Hramov A.* Review on the use of brain computer interface rehabilitation methods for treating mental and neurological conditions // *J. Integr. Neurosci*. 2024. Vol. 23, no. 7. P. 125. DOI: 10.31083/j.jin2307125.
72. *Kuo M. F., Nitsche M. A.* Effects of transcranial electrical stimulation on cognition // *Clin. EEG Neurosci*. 2012. Vol. 43, no. 3. P. 192–199. DOI: 10.1177/1550059412444975.
73. *Anguera J. A., Boccanfuso J., Rintoul J. L., Al-Hashimi O., Faraji F., Janowich J., Kong E., Larraburo Y., Rolle C., Johnston E., Gazzaley A.* Video game training enhances cognitive control in older adults // *Nature*. 2013. Vol. 501, no. 7465. P. 97–101. DOI: 10.1038/nature12486.
74. *Grubov V. V., Khramova M. V., Goman S., Badarin A. A., Kurkin S. A., Andrikov D. A., Pitsik E., Antipov V., Petushok E., Brusinskii N., Bukina T., Fedorov A. A., Hramov A. E.* Open-loop neuroadaptive system for enhancing student's cognitive abilities in learning // *IEEE Access*. 2024. Vol. 12. P. 49034–49049. DOI: 10.1109/access.2024.3383847.
75. *Букина Т. В., Храмова М. В., Куркин С. А., Андриков Д. А., Гоман С. С., Дедков А. Е., Хра-*

- мов А. Е. Нейрообразовательный программный рекомендательный сервис как инструмент персонализации образовательного процесса // Информатика и образование. 2024. Т. 39, № 5. С. 50–62. DOI: 10.32517/0234-0453-2024-39-5-50-62.
76. Donati A. R., Shokur S., Morya E., Campos D. S., Moiola R. C., Gitti C. M., Augusto P. B., Tripodi S., Pires C. G., Pereira G. A., Brasil F. L., Gallo S., Lin A. A., Takigami A. K., Aratanha M. A., Joshi S., Bleuler H., Cheng G., Rudolph A., Nicoletis M. A. L. Long-term training with a brain-machine interface-based gait protocol induces partial neurological recovery in paraplegic patients // *Sci. Rep.* 2016. Vol. 6. P. 30383. DOI: 10.1038/srep30383.
 77. Poppe C., Elger B. S. Brain–computer interfaces, completely locked-in state in neurodegenerative diseases, and end-of-life decisions // *Journal of Bioethical Inquiry*. 2024. Vol. 21, no. 1. P. 19–27. DOI: 10.1007/s11673-023-10256-5.
 78. Kübler A., Neumann N. Brain-computer interfaces—the key for the conscious brain locked into a paralyzed body // *Prog. Brain Res.* 2005. Vol. 150. P. 513–525. DOI: 10.1016/S0079-6123(05)50035-9.
 79. Lebedev M. A., Nicoletis M. A. Brain-machine interfaces: past, present and future // *Trends Neurosci.* 2006. Vol. 29, no. 9. P. 536–546. DOI: 10.1016/j.tins.2006.07.004.
 80. Veena N., Anitha N. A review of non-invasive BCI devices. *Int J Biomed Eng Technol.* 2020. Vol. 34, no. 3. P. 205–233. DOI: 10.1504/IJBET.2020.111471.
 81. Grill W. M., Norman S. E., Bellamkonda R. V. Implanted neural interfaces: biochallenges and engineered solutions // *Annu Rev Biomed Eng.* 2009. Vol. 11. P. 1–24. DOI: 10.1146/annurev-bioeng-061008-124927.
 82. Salatino J. W., Ludwig K. A., Kozai T. D. Y., Purcell E. K. Glial responses to implanted electrodes in the brain // *Nat. Biomed. Eng.* 2017. Vol. 1, no. 11. P. 862–877. DOI: 10.1038/s41551-017-0154-1.
 83. Rossi S., Hallett M., Rossini P. M., Pascual-Leone A., *The Safety of TMS Consensus Group*. Safety, ethical considerations, and application guidelines for the use of transcranial magnetic stimulation in clinical practice and research // *Clin. Neurophysiol.* 2009. Vol. 120, no. 12. P. 2008–2039. DOI: 10.1016/j.clinph.2009.08.016.
 84. Kringelbach M. L., Jenkinson N., Owen S. L., Aziz T. Z. Translational principles of deep brain stimulation // *Nature Reviews Neuroscience*. 2007. Vol. 8, no. 8. P. 623–635. DOI: 10.1038/nrn2196.
 85. Bergey G. K., Morrell M. J., Mizrahi E. M., Goldman A., King-Stephens D., Nair D., Srinivasan S., Jobst B., Gross R. E., Shields D. C., Barkley G., Salanova V., Olejniczak P., Cole A., Cash S. S., Noe K., Wharen R., Worrell G., Murro A. M., Edwards J., Duchowny M., Spencer D., Smith M., Geller E., Gwinn R., Skidmore Ch., Eisenschenk S., Berg M., Heck Ch., Van Ness P., Fountain N., Rutecki P., Massey A., O'Donovan C., Labar D., Duckrow R. B., Hirsch L. J., Courtney T., Sun F. T., Seale C. G. Long-term treatment with responsive brain stimulation in adults with refractory partial seizures // *Neurology*. 2015. Vol. 84, no. 8. P. 810–817. DOI: 10.1212/WNL.0000000000001280.
 86. Deer T. R., Pope J. E., Hayek S. M., Bux A., Buchser E., Eldabe S., De Andrés J. A., Erdek M., Patin D., Grider J. S., Doleys D. M., Jacobs M. S., Yaksh T. L., Poree L., Wallace M. S., Prager J., Rauck R., DeLeon O., Diwan S., Falowski S. M., Gazelka H. M., Kim Ph., Leong M., Levy R. M., McDowell G., McRoberts P., Naidu R., Narouze S., Perruchoud Ch., Rosen S. M., Rosenberg W. S., Saulino M., Staats P., Stearns L. J., Willis D., Krames E., Huntoon M., Mekhail N. The Polyanalgesic Consensus Conference (PACC): Recommendations on Intrathecal Drug Infusion Systems Best Practices and Guidelines // *Neuromodulation*. 2017. Vol. 20, no. 2. P. 96–132. DOI: 10.1111/ner.12538.
 87. Deisseroth K. Optogenetics: 10 years of microbial opsins in neuroscience // *Nat. Neurosci.* 2015. Vol. 18, no. 9. P. 1213–1225. DOI: 10.1038/nn.4091.
 88. LeCun Y., Bengio Y., Hinton G. Deep learning // *Nature*. 2015. Vol. 521, no. 7553. P. 436–444. DOI: 10.1038/nature14539.
 89. Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Jones L., Gomez A. N., Kaiser L., Polosukhin I. Attention is all you need // In: Guyon I., Luxburg U. V., Bengio S., Wallach H., Fergus R., Vishwanathan S., Garnett R., editors. *Advances in Neural Information Processing Systems*. vol. 30. Curran Associates, Inc., 2017. P. 5998–6008.

90. Goodfellow I. J., Pouget-Abadie J., Mirza M., Xu B., Warde-Farley D., Ozair S., Courville A., Bengio Y. Generative adversarial nets // In: Proceedings of the 28th International Conference on Neural Information Processing Systems - Volume 2. NIPS'14. Cambridge: MIT Press, 2014. P. 2672–2680.
91. Krizhevsky A., Sutskever I., Hinton G. E. ImageNet classification with deep convolutional neural networks // In: Proceedings of the 26th International Conference on Neural Information Processing Systems - Volume 1. NIPS'12. Red Hook, NY, USA: Curran Associates Inc., 2012. P. 1097–1105.
92. He K., Zhang X., Ren S., Sun J. Deep Residual Learning for Image Recognition // In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016. P. 770–778. DOI: 10.1109/CVPR.2016.90.
93. Qi C. R., Su H., Mo K., Guibas L. J. PointNet: Deep learning on point sets for 3D classification and segmentation // In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017. P. 77–85. DOI: 10.1109/CVPR.2017.16.
94. Girshick R., Donahue J., Darrell T., Malik J. Rich feature hierarchies for accurate object detection and semantic segmentation // In: 2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2014. P. 580–587. DOI: 10.1109/CVPR.2014.81.
95. Esteva A., Kuprel B., Novoa R. A., Ko J., Swetter S. M., Blau H. M., Thrun S. Dermatologist-level classification of skin cancer with deep neural networks // Nature. 2017. Vol. 542, no. 7639. P. 115–118. DOI: 10.1038/nature21056.
96. LeCun Y., Bottou L., Bengio Y., Haffner P. Gradient-based learning applied to document recognition // Proceedings of the IEEE. 1998. Vol. 86, no. 11. P. 2278–2324. DOI: 10.1109/5.726791.
97. Chen C., Seff A., Kornhauser A., Xiao J. DeepDriving: Learning affordance for direct perception in autonomous driving // In: 2015 IEEE International Conference on Computer Vision (ICCV), 2015. p. 2722–2730. DOI: 10.1109/ICCV.2015.312.
98. Loquercio A., Maqueda A. I., del Blanco C. R., Scaramuzza D. DroNet: Learning to fly by driving // IEEE Robotics and Automation Letters. 2018. Vol. 3, no. 2. P. 1088–1095. DOI: 10.1109/LRA.2018.2795643.
99. Pinto L., Gupta A. Supersizing self-supervision: Learning to grasp from 50K tries and 700 robot hours // In: 2016 IEEE International Conference on Robotics and Automation (ICRA). 2016. p. 3406–3413. DOI: 10.1109/ICRA.2016.7487517.
100. Kober J., Bagnell J. A., Peters J. Reinforcement learning in robotics: A survey // The International Journal of Robotics Research. 2013. Vol. 32, no. 11. P. 1238–1274. DOI: 10.1177/0278364913495721.
101. Koopman P., Wagner M. Autonomous vehicle safety: An interdisciplinary challenge // IEEE Intelligent Transportation Systems Magazine. 2017. Vol. 9, no. 1. P. 90–96. DOI: 10.1109/MITS.2016.2583491.
102. Leveson N. G. Engineering a Safer World: Systems Thinking Applied to Safety. Cambridge: MIT Press, 2011. 560 p. DOI: 10.7551/mitpress/8179.001.0001.
103. Jurafsky D., Martin J. H. Speech and Language Processing. Harlow: Pearson Education, 2020.
104. Goodman B., Flaxman S. European Union regulations on algorithmic decision-making and a “right to explanation” // AI Magazine. 2017. Vol. 38, no. 3. P. 50–57.
105. Ribeiro M. T., Singh S., Guestrin C. “Why Should I Trust You?”: Explaining the predictions of any classifier // In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. KDD '16. New York, NY, USA: Association for Computing Machinery, 2016. p. 1135–1144. DOI: 10.1145/2939672.2939778.
106. Caruana R., Lou Y., Gehrke J., Koch P., Sturm M., Elhadad N. Intelligible models for healthcare: Predicting pneumonia risk and hospital 30-day readmission // In: Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. KDD '15. New York, NY, USA: Association for Computing Machinery, 2015. p. 1721–1730. DOI: 10.1145/2783258.2788613.
107. Bassett D. S., Khambhati A. N. A network engineering perspective on probing and perturbing

- cognition with neurofeedback // *Ann. NY Acad. Sci.* 2017. Vol. 1396, no. 1. P. 126–143. DOI: 10.1111/nyas.13338.
108. *Jiang F., Jiang Y., Zhi H., Dong Y., Li H., Ma S., Wang Y.* Artificial intelligence in healthcare: past, present and future // *Stroke Vasc Neurol.* 2017. Vol. 2, no. 4. P. 230–243. DOI: 10.1136/svn-2017-000101.
 109. *Yuste R., Goering S., Arcas B. A. Y., Bi G., Carmena J. M., Carter A., Fins J. J., Friesen P., Gallant J., Huggins J. E., Illes J., Kellmeyer P., Klein E., Marblestone A., Mitchell C., Parens E., Pham M., Rubel A., Sadato N., Sullivan L. S., Teicher M., Wasserman D., Wexler A., Whittaker M., Wolpaw J.* Four ethical priorities for neurotechnologies and AI // *Nature.* 2017. Vol. 551, no. 7679. P. 159–163. DOI: 10.1038/551159a.
 110. *Plis S. M., Hjelm D. R., Salakhutdinov R., Allen E. A., Bockholt H. J., Long J. D., Johnson H. J., Paulsen J. S., Turner J. A., Calhoun V. D.* Deep learning for neuroimaging: a validation study // *Front. Neurosci.* 2014. Vol. 8. P. 229. DOI: 10.3389/fnins.2014.00229.
 111. *Pisarchik A. N., Maksimenko V. A., Andreev A. V., Frolov N. S., Makarov V. V., Zhuravlev M. O., Runnova A. E., Hramov A. E.* Coherent resonance in the distributed cortical network during sensory information processing // *Sci. Rep.* 2019. Vol. 9, no. 1. P. 18325. DOI: 10.1038/s41598-019-54577-1.
 112. *Дедков А. Е., Андриков Д. А., Храмов А. Е.* Обзор способов измерения когнитивной нагрузки мозга и методов машинного обучения для их идентификации на основе данных ЭЭГ // *Врач и информационные технологии.* 2024. № 3. С. 20–31. DOI: 10.25881/18110193_2024_3_20.
 113. *Shatte A. B. R., Hutchinson D. M., Teague S. J.* Machine learning in mental health: A scoping review of methods and applications // *Psychol. Med.* 2019. Vol. 49, no. 9. P. 1426–1448. DOI: 10.1017/S0033291719000151.
 114. *Grubov V. V., Nazarikov S. I., Kurkin S. A., Utyashev N. P., Andrikov D. A., Karpov O. E., Hramov A. E.* Two-stage approach with combination of outlier detection method and deep learning enhances automatic epileptic seizure detection // *IEEE Access.* 2024. Vol. 12. P. 122168–122182. DOI: 10.1109/access.2024.3453039.
 115. *Pitsik E. N., Maximenko V. A., Kurkin S. A., Sergeev A. P., Stoyanov D., Paunova R., Kandilarova S., Simeonova D., Hramov A. E.* The topology of fMRI-based networks defines the performance of a graph neural network for the classification of patients with major depressive disorder // *Chaos, Solitons and Fractals.* 2023. Vol. 167. P. 113041. DOI: 10.1016/j.chaos.2022.113041.
 116. *Zhang S., Zhao H., Wang W., Wang Z., Luo X., Hramov A., Kurths J.* Edge-centric effective connection network based on multi-modal MRI for the diagnosis of Alzheimer’s disease // *Neurocomputing.* 2023. Vol. 552. P. 126512. DOI: 10.1016/j.neucom.2023.126512.
 117. *Park K. H., Sun S., Lim Y. H., Park H. R., Lee J. M., Park K., Jeon B., Park H.-P., Kim H. Ch., Paek S. H.* Clinical outcome prediction from analysis of microelectrode recordings using deep learning in subthalamic deep brain stimulation for Parkinson’s disease // *PLoS ONE.* 2021. Vol. 16, no. 1. P. e0244133. DOI: 10.1371/journal.pone.0244133.
 118. *Hubel D. H., Wiesel T. N.* Receptive fields, binocular interaction and functional architecture in the cat’s visual cortex // *J. Physiol.* 1962. Vol. 160, no. 1. P. 106–154. DOI: 10.1113/jphysiol.1962.sp006837.
 119. *Behrouz A., Zhong P., Mirrokni V.* Titans: Learning to Memorize at Test Time. arXiv:250100663 arXiv Preprint. 2024. DOI: 10.48550/arXiv.2501.00663.
 120. *Stanton S. J., Sinnott-Armstrong W., Huettel S. A.* Neuromarketing: Ethical implications of its use and potential misuse // *J. Bus. Ethics.* 2017. Vol. 144. P. 799–811. DOI: 10.1007/s10551-016-3059-0.
 121. *Luna-Nevarez C.* Neuromarketing, ethics, and regulation: An exploratory analysis of consumer opinions and sentiment on blogs and social media // *J. Consum. Policy.* 2021. Vol. 44, no. 4. P. 559–583. DOI: 10.1007/s10603-021-09496-y.
 122. *Maron D. F.* Science career ads are disproportionately seen by men. *Scientific American* [Electronic

- resource] Available from: <https://wwwscientificamericancom/article/science-career-ads-are-disproportionately-seen-by-men>. 2018.
123. *Goering S., Yuste R.* On the necessity of ethical guidelines for novel neurotechnologies // *Cell*. 2016. Vol. 167, no. 4. P. 882–885. DOI: 10.1016/j.cell.2016.10.029.
 124. *Leefmann J., Levallois C., Hildt E.* Neuroethics 1995–2012. A bibliometric analysis of the guiding themes of an emerging research field // *Front. Hum. Neurosci.* 2016. Vol. 10. P. 336. DOI: 10.3389/fnhum.2016.00336.
 125. *Mehrabi N., Morstatter F., Saxena N., Lerman K., Galstyan A.* A survey on bias and fairness in machine learning // *ACM Computing Surveys*. 2021. Vol. 54, no. 6. P. 1–35. DOI: 10.1145/3457607.
 126. *Barocas S., Selbst A.D.* Big Data’s Disparate Impact // *California Law Review*. 2016. Vol. 104, no. 3. P. 671–732. DOI: 10.2139/ssrn.2477899.
 127. *O’Neil C.* Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. New York: Crown Publishing Group, 2016.
 128. *Caliskan A., Bryson J.J., Narayanan A.* Semantics derived automatically from language corpora contain human-like biases // *Science*. 2017. Vol. 356, no. 6334. P. 183–186. DOI: 10.1126/science.aal4230.
 129. *Raghavan M., Barocas S., Kleinberg J., Levy K.* Mitigating bias in algorithmic hiring: evaluating claims and practices // In: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency. FAT* ’20. New York, NY, USA: Association for Computing Machinery, 2020. p. 469–481. DOI: 10.1145/3351095.3372828.
 130. *Bartlett R., Morse A., Stanton R., Wallace N.* Consumer-lending discrimination in the FinTech era // *Journal of Financial Economics*. 2022. Vol. 143. P. 30–56. DOI: 10.1016/j.jfineco.2021.05.047.
 131. *Angwin J., Larson J., Mattu S., Kirchner L.* Machine bias: There’s software used across the country to predict future criminals. And it’s biased against blacks. ProPublica [Electronic resource] Available from: <https://wwwpropublicaorg/article/machine-bias-risk-assessments-in-criminal-sentencing>. 2016.
 132. *Van Deursen A.J.A.M., Van Dijk J.P.M.* The digital divide shifts to differences in usage // *New Media & Society*. 2014. Vol. 16, no. 3. P. 507–526. DOI: 10.1177/1461444813487959.
 133. *Yadav B.R.* The Ethics of Understanding: Exploring Moral Implications of Explainable AI // *International Journal of Science and Research (IJSR)*. 2024. Vol. 13, no. 6. P. 1–7. DOI: 10.21275/SR24529122811.
 134. *Kundu S.* AI in medicine must be explainable // *Nat. Med.* 2021. Vol. 27, no. 8. P. 1328. DOI: 10.1038/s41591-021-01461-z.
 135. *McDermid J.A., Jia Y., Porter Z., Habli I.* Artificial intelligence explainability: the technical and ethical dimensions // *Philos. Trans. A. Math. Phys. Eng. Sci.* 2021. Vol. 379, no. 2207. P. 20200363. DOI: 10.1098/rsta.2020.0363.
 136. *Angelov P.P., Soares E.A., Jiang R., Arnold N.I., Atkinson P.M.* Explainable artificial intelligence: an analytical review // *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*. 2021. Vol. 11, no. 5. P. e1424. DOI: 10.1002/widm.1424.
 137. *Dwivedi R., Dave D., Naik H., Singhal S., Omer R., Patel P., Qian B., Wen Z., Shah T., Morgan G., Ranjan R.* Explainable AI (XAI): Core ideas, techniques, and solutions // *ACM Computing Surveys*. 2023. Vol. 55, no. 9. P. 835. DOI: 10.1145/3561048.
 138. *Minh D., Wang H.X., Li Y.F., Nguyen T.N.* Explainable artificial intelligence: a comprehensive review // *Artif. Intell. Rev.* 2022. Vol. 55. P. 3503–3568. DOI: 10.1007/s10462-021-10088-y.
 139. *Ponce-Bobadilla A.V., Schmitt V., Maier C.S., Mensing S., Stodtmann S.* Practical guide to SHAP analysis: Explaining supervised machine learning model predictions in drug development // *Clin. Transl. Sci.* 2024. Vol. 17, no. 11. P. e70056. DOI: 10.1111/cts.70056.
 140. *Islam R., Andreev A.V., Shusharina N.N., Hramov A.E.* Explainable machine learning methods for classification of brain states during visual perception // *Mathematics*. 2022. Vol. 10, no. 15. P. 2819. DOI: 10.3390/math10152819.

141. *Bhati D., Neha F., Amiruzzaman M.* A survey on explainable artificial intelligence (xai) techniques for visualizing deep learning models in medical imaging // *Journal of Imaging*. 2024. Vol. 10, no. 10. P. 239. DOI: 10.3390/jimaging10100239.
142. *Lai T.* Interpretable medical imagery diagnosis with self-attentive transformers: a review of explainable AI for health care // *BioMedInformatics*. 2024. Vol. 4, no. 1. P. 113–126. DOI: 10.3390/biomedinformatics4010008.
143. *Catherine S. B. A.* Unlocking the Black Box: Exploring the Future of Explainable AI in Real-World Applications // In: *Multidisciplinary Research in Arts, Science & Commerce*. Vol. 25. The Hill Publication, 2025. P. 37–38.
144. *Hassija V., Chamola V., Mahapatra A., Singal A., Goel D., Huang K., Scardapane S., Spinelli I., Mahmud M., Hussain A.* Interpreting black-box models: a review on explainable artificial intelligence // *Cogn. Comput.* 2024. Vol. 16. P. 45–74. DOI: 10.1007/s12559-023-10179-8.
145. *Карпов О. Е., Андриков Д. А., Максименко В. А., Храмов А. Е.* Прозрачный искусственный интеллект для медицины // *Врач и информационные технологии*. 2022. № 2. С. 4–11. DOI: 10.25881/18110193_2022_2_4.
146. *Černevičienė J., Kabašinskas A.* Explainable artificial intelligence (XAI) in finance: a systematic literature review // *Artif. Intell. Rev.* 2024. Vol. 57. P. 216. DOI: 10.1007/s10462-024-10854-8.
147. *Craig T.* *Privacy and Big Data: The Players, Regulators, and Stakeholders*. Sebastopol: O’Reilly Media, Inc., 2011. 91 p.
148. *Curzon J., Kosa T. A., Akalu R., El-Khatib K.* Privacy and artificial intelligence // *IEEE Transactions on Artificial Intelligence*. 2021. Vol. 2, no. 2. P. 96–108. DOI: 10.1109/TAI.2021.3088084.
149. *Chesterman S.* Artificial intelligence and the problem of autonomy // *Notre Dame Journal on Emerging Technologies*. 2020. Vol. 1. P. 210–250.
150. *Novelli C., Taddeo M., Floridi L.* Accountability in artificial intelligence: what it is and how it works // *AI and Soc.* 2024. Vol. 39. P. 1871–1882. DOI: 10.1007/s00146-023-01635-y.
151. *Le cerveau et la loi: analyse de l’émergence du neurodroit / Oullier O., editor.* Paris: Centre d’analyse stratégique, 2012. Available from: http://archives.strategie.gouv.fr/cas/system/files/cas-dqs_dt-neurodroit_11septembreduit_0.pdf.
152. *Eagleman D. M.* Pourquoi les sciences du cerveau peuvent éclairer le droit // In: *Oullier O., editor. Le cerveau et la loi: analyse de l’émergence du neurodroit*. Paris: Centre d’analyse stratégique, 2012. P. 33–52.
153. *Ernst E., Merola R., Samaan D.* Economics of artificial intelligence: Implications for the future of work // *IZA Journal of Labor Policy*. 2019. Vol. 9, no. 1. P. 4. DOI: 10.2478/izajolp-2019-0004.
154. *McGaughey E.* Will robots automate your job away? Full employment, basic income and economic democracy // *Industrial Law Journal*. 2022. Vol. 51, no. 3. P. 511–559. DOI: 10.1093/indlaw/dwab010.
155. *Russell S.* *Human Compatible: Artificial Intelligence and the Problem of Control*. New York: Viking, 2019. 352 p.
156. *Torres P.* Superintelligence and the future of governance: On prioritizing the control problem at the end of history // In: *Artificial Intelligence Safety and Security*. Chapman and Hall/CRC, 2018. P. 357–374. DOI: 10.1201/9781351251389-24.
157. *Totschnig W.* The problem of superintelligence: Political, not technological // *AI and Soc.* 2019. Vol. 34. P. 907–920. DOI: 10.1007/s00146-017-0753-0.
158. Яндекс: Подход к оценке рисков : электронный документ. Режим доступа : для зарегистрированных пользователей.
159. *da Silva Castanheira J., Orozco Perez H. D., Misic B., Baillet S.* Brief segments of neurophysiological activity enable individual differentiation // *Nat. Commun.* 2021. Vol. 12. P. 5713. DOI:10.1038/s41467-021-25895-8.
160. *Шваб К.* Технологии четвертой промышленной революции. М.: Эксмо, 2018. 317 с.
161. *Ienca M.* On artificial intelligence and manipulation // *Topoi*. 2023. Vol. 42, no. 3. P. 833–842. DOI: 10.1007/s11245-023-09940-3.

162. *Lemsieh H., Abarar I.* Artificial intelligence: A technological tool to manipulate the psychology and behavior of consumers: Theoretical research // *International Journal of Accounting, Finance, Auditing, Management and Economics*. 2024. Vol. 5, no. 6. P. 432–449. DOI: 10.5281/zenodo.12186136.
163. Патент № 2653239 С1 Российская Федерация, МПК G06K 9/66, A61B 5/0476, H04L 9/32. Способ идентификации человека по ЭЭГ-отклику на неоднозначные изображения: № 2017115459: заявл. 02.05.2017: опубл. 07.05.2018 / А. Н. Писарчик, В. Ю. Мусатов, А. Е. Руннова [и др.]; заявитель Федеральное государственное бюджетное образовательное учреждение высшего образования «Саратовский государственный технический университет имени Гагарина Ю. А.» (СГТУ имени Гагарина Ю. А.).
164. *Panch T., Mattie H., Celi L. A.* The “inconvenient truth” about AI in healthcare // *NPJ Digit Med*. 2019, no. 2. P. 77. DOI: 10.1038/s41746-019-0155-4.

References

1. Hassabis D, Kumaran D, Summerfield C, Botvinick M. Neuroscience-inspired artificial intelligence. *Neuron*. 2017;95(2):245–258. DOI: 10.1016/j.neuron.2017.06.011.
2. Karpov OE, Pitsik EN, Kurkin SA, Maksimenko VA, Gusev AV, Shusharina NN, Hramov AE. Analysis of publication activity and research trends in the field of AI medical applications: Network approach. *Int. J. Environ. Res. Public Health*. 2023;20(7):5335. DOI: 10.3390/ijerph20075335.
3. Nyholm S. Artificial intelligence and human enhancement: Can AI technologies make us more (artificially) intelligent? *Camb Q Healthc Ethics*. 2024;33(1):76–88. DOI: 10.1017/S0963180123000464.
4. Duan YQ, Edwards JS, Dwivedi YK. Artificial intelligence for decision making in the era of Big Data – evolution, challenges and research agenda. *International Journal of Information Management*. 2019;48:63–71. DOI: 10.1016/j.ijinfomgt.2019.01.021.
5. Autor DH. Why are there still so many jobs? The history and future of workplace automation. *Journal of Economic Perspectives*. 2015;29(3):3–30. DOI: 10.1257/jep.29.3.3.
6. Brown T, Mann B, Ryder N, Subbiah M, Kaplan JD, Dhariwal Prafulla, Neelakantan A, Shyam P, Sastry G, Askell A, Agarwal S, Herbert-Voss A, Krueger G, Henighan T, Child R, Ramesh A, Ziegler D, Wu J, Winter C, Hesse Chris, Chen Mark, Sigler E, Litwin M, Gray S, Chess B, Clark J, Berner Ch, McCandlish S, Radford A, Sutskever Ilya, Amodei D. Language models are few-shot learners. In: Larochelle H, Ranzato M, Hadsell R, Balcan MF, Lin H, editors. *Advances in Neural Information Processing Systems*. Vol. 33. 2020. P. 1877–1901.
7. Wang L, Törngren M, Onori M. Current status and advancement of cyber-physical systems in manufacturing. *Journal of Manufacturing Systems*. 2015;37:517–527. DOI: 10.1016/j.jmsy.2015.04.008.
8. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nature Medicine*. 2019;25(1):44–56. DOI: 10.1038/s41591-018-0300-7.
9. Karpov OE, Hramov AE. *Information Technology, Computing Systems and Artificial Intelligence in Medicine*. M.: DPK Press; 2022. 480 p. (in Russian).
10. Fedorov AA, Kurkin SA, Hramova MV, Hramov AE. Neurotechnology and artificial intelligence as key factors in the customization of the lifelong learning route. *Informatics and Education*. 2023;38(3):5–15. DOI: 10.32517/0234-0453-2023-38-3-5-15.
11. Holmes W, Tuomi I. State of the art and practice in AI in education. *European Journal of Education*. 2022;57(4):542–570. DOI: 10.1111/ejed.12533.
12. Carleo G, Cirac I, Cranmer K, Daudet L, Schuld M, Tishby N, Vogt-Maranto L, Zdeborová L. Machine learning and the physical sciences. *Rev Mod Phys*. 2019;91(4):045002. DOI: 10.1103/RevModPhys.91.045002.
13. Silver D, Hubert T, Schrittwieser J, Antonoglou I, Lai M, Guez A, Lanctot M, Sifre L, Kumaran D, Graepel T, Lillicrap T, Simonyan K, Hassabis D. A general reinforcement learning

- algorithm that masters chess, shogi, and Go through self-play. *Science*. 2018;362(6419):1140–1144. DOI: 10.1126/science.aar6404.
14. Bandettini PA. What's new in neuroimaging methods? *Ann NY Acad Sci*. 2009;1156:260–293. DOI: 10.1111/j.1749-6632.2009.04420.x.
 15. Hramov AE, Maksimenko VA, Pisarchik AN. Physical principles of brain-computer interfaces and their applications for rehabilitation, robotics and control of human brain states. *Phys Rep*. 2021;918:1–133. DOI: 10.1016/j.physrep.2021.03.002.
 16. Marzbani H, Marateb HR, Mansourian M. Neurofeedback: a comprehensive review on system design, methodology and clinical applications. *Basic Clin Neurosci*. 2016;7(2):143–158. DOI: 10.15412/J.BCN.03070208.
 17. Mihara M, Miyai I, Hattori N, Hatakenaka M, Yagura H, Kawano T, Okibayashi M, Danjo N, Ishikawa A, Inoue Y, Kubota K. Neurofeedback using real-time near-infrared spectroscopy enhances motor imagery related cortical activation. *PLoS ONE*. 2012;7(3):e32234. DOI: 10.1371/journal.pone.0032234.
 18. Hallett M. Transcranial magnetic stimulation: a primer. *Neuron*. 2007;55(2):187–199. DOI: 10.1016/j.neuron.2007.06.026.
 19. Rossi S, Hallett M, Rossini P, Pascual-Leone A, The Safety of TMS Consensus Group. Safety, ethical considerations, and application guidelines for the use of transcranial magnetic stimulation in clinical practice and research. *Clin Neurophysiol*. 2009;120(12):2008–2039. DOI: 10.1016/j.clinph.2009.08.016.
 20. Brunoni A, Nitsche M, Bolognini N, Bikson M, Wagner T, Merabet L, Edwards D, Valero-Cabre A, Rotenberg A, Pascual-Leone A, Ferrucci R, Priori A, Boggio P, Fregni F. Clinical research with transcranial direct current stimulation (tDCS): Challenges and future directions. *Brain Stimul*. 2012;5(3):175–195. DOI: 10.1016/j.brs.2011.03.002.
 21. Benabid A, Chabardes S, Mitrofanis J, Pollak P. Deep brain stimulation of the subthalamic nucleus for the treatment of Parkinson's disease. *Lancet Neurol*. 2009;8(1):67–81. DOI: 10.1016/S1474-4422(08)70291-6.
 22. Cometa A, Falasconi A, Biasizzo M, Carpaneto J, Horn A, Mazzoni A, Micera S. Clinical neuroscience and neurotechnology: An amazing symbiosis. *iScience*. 2022;25(10):105124. DOI: 10.1016/j.isci.2022.105124.
 23. Daly JJ, Wolpaw JR. Brain-computer interfaces in neurological rehabilitation. *Lancet Neurol*. 2008;7(11):1032–1043. DOI: 10.1016/S1474-4422(08)70223-0.
 24. Plass-Oude Bos D, Reuderink B, van de Laar B, Gürkök H, Mühl C, Poel M, Nijholt A, Heylen D. Brain-Computer Interfacing and Games. In: Tan D, Nijholt A (eds) *Brain-Computer Interfaces. Human-Computer Interaction Series*. London: Springer; 2010. P. 149–178. DOI: 10.1007/978-1-84996-272-8_10.
 25. Marshall D, Coyle D, Wilson S, Callaghan M. Games, gameplay, and BCI: The State of the Art. *IEEE Transactions on Computational Intelligence and AI in Games*. 2013;5(2):164–176. DOI: 10.1109/tciaig.2013.2263555.
 26. Campbell M, Hoane AJ, Hsu Fh. Deep Blue. *Artificial Intelligence*. 2002;134(1–2):57–83. DOI: 10.1016/S0004-3702(01)00129-1.
 27. Hsu FH, Anantharaman T, Campbell M, Nowatzyk A. A grandmaster chess machine. *Scientific American*. 1990;263(4):44–51. <https://www.jstor.org/stable/24997060>.
 28. He K, Zhang X, Ren S, Sun J. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In: *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*; 2015. p. 1026–1034. DOI: 10.1109/ICCV.2015.123.
 29. Silver D, Schrittwieser J, Simonyan K, Antonoglou I, Huang A, Guez A, Hubert T, Baker L, Lai M, Bolton A, Chen Y, Lillicrap T, Hui F, Sifre L, van den Driessche G, Graepel Th, Hassabis D. Mastering the game of go without human knowledge. *Nature*. 2017;550(7676):354–359. DOI: 10.1038/nature24270.

30. Logothetis NK. What we can do and what we cannot do with fMRI. *Nature*. 2008;453(7197): 869–878. DOI: 10.1038/nature06976.
31. Luck SJ. *An Introduction to the Event-Related Potential Technique*. Massachusetts: MIT Press; 2014. 388 p.
32. Hämäläinen M, Hari R, Ilmoniemi RJ, Knuutila J, Lounasmaa OV. Magnetoencephalography-theory, instrumentation, and applications to noninvasive studies of the working human brain. *Rev. Mod. Phys.* 1993;65(2):413–497. DOI: 10.1103/RevModPhys.65.413.
33. Ferrari M, Quaresima V. A brief review on the history of human functional near-infrared spectroscopy (fNIRS) development and fields of application. *NeuroImage*. 2012;63(2):921–935. DOI: 10.1016/j.neuroimage.2012.03.049.
34. Grau C, Ginhoux R, Riera A, Nguyen TL, Chauvat H, Berg M, Amengual JL, Pascual-Leone A, Ruffini G. Conscious brain-to-brain communication in humans using non-invasive technologies. *PLoS ONE*. 2014;9(8):e105225. DOI: 10.1371/journal.pone.0105225.
35. Nam CS, Traylor Z, Chen M, Jiang X, Feng W, Chhatbar PY. Direct communication between brains: A systematic PRISMA review of brain-to-brain interface. *Front Neurobot*. 2021;15:656943. DOI: 10.3389/fnbot.2021.656943.
36. Maksimenko VA, Hramov AE, Frolov NS, Lüttjohann A, Nedaivozov VO, Grubov VV, Runnova AE, Makarov VV, Kurths J, Pisarchik AN. Increasing human performance by sharing cognitive load using brain-to-brain interface. *Front Neurosci*. 2018;12:949. DOI: 10.3389/fnins.2018.00949.
37. Kurkin S, Gordleeva S, Savosenkov A, Grigorev N, Smirnov N, Grubov VV, Udoratina A, Maksimenko V, Kazantsev V, Hramov AE. Transcranial Magnetic Stimulation of the Dorsolateral Prefrontal Cortex Increases Posterior Theta Rhythm and Reduces Latency of Motor Imagery. *Sensors*. 2023;23(10):4661. DOI: 10.3390/s23104661.
38. Musk E. An integrated brain-machine interface platform with thousands of channels. *J Med Internet Res*. 2019;21(10):e16194. DOI: 10.2196/16194.
39. Pisarchik AN, Maksimenko VA, Hramov AE. From novel technology to novel applications: Comment on “An Integrated Brain-Machine Interface Platform With Thousands of Channels” by Elon Musk and Neuralink. *J. Med. Internet. Res.* 2019;21(10):e16356. DOI: 10.2196/16356.
40. Opie NL, John SE, Rind GS, Ronayne SM, Wong YT, Gerboni G, Yoo PE, Lovell T, Scordas T, Wilson SL, Dornom A, Vale T, O’Brien TJ, Grayden DB, May CN, Oxley ThJ. Focal stimulation of the sheep motor cortex with a chronically implanted minimally invasive electrode array mounted on an endovascular stent. *Nat Biomed Eng*. 2018;2(12):907–914. DOI: 10.1038/s41551-018-0321-z.
41. Filipova IA. Neurotechnologies: development, practical application and regulation. *Bulletin of St Petersburg State University Law*. 2021;12(3):502–521. DOI: 10.21638/spbu14.2021.302.
42. Lecomte P. Umwelt as the foundation of an ethics of smart environments. *Humanit Soc Sci Commun*. 2023;10:925. DOI: 10.1057/s41599-023-02356-9.
43. Ienca M, Andorno R. Towards new human rights in the age of neuroscience and neurotechnology. *Life Sci Soc Policy*. 2017;13(5):1–27. DOI: 10.1186/s40504-017-0050-1.
44. Kellmeyer P. Big brain data: On the responsible use of brain data from clinical and consumer-directed neurotechnological devices. *Neuroethics*. 2018;14:83–98. DOI: 10.1007/s12152-018-9371-x.
45. Mittelstadt BD, Allo P, Taddeo M, Wachter S, Floridi L. The ethics of algorithms: Mapping the debate. *Big Data and Society*. 2016;3(2):2053951716679679. DOI: 10.1177/2053951716679679.
46. Floridi L, Cowls J, Beltrametti M, Chatila R, Chazerand P, Dignum V, Luetge C, Madelin R, Pagallo U, Rossi F, Schafer B, Valcke P, Vayena E. AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds Mach*. 2018;28(4):689–707. DOI: 10.1007/s11023-018-9482-5.
47. Calo R. *Artificial Intelligence Policy: A Primer and Roadmap*. *UCD L Rev*. 2017;51:399.
48. Battista D. Political communication in the age of artificial intelligence: an overview of deepfakes and their implications. *Society Register*. 2024;8(2):7–24. DOI: 10.14746/sr.2024.8.2.01.

49. Gambín AF, Yazidi A, Vasilakos A, Haugerud H, Djenouri Y. Deepfakes: current and future trends. *Artif Intell Rev.* 2024;57(3):64. DOI: 10.1007/s10462-023-10679-x.
50. Chesney R, Citron D. Deepfakes and the new disinformation war: The coming age of post-truth geopolitics. *Foreign Affairs.* 2019;98(1):147–155.
51. Maras MH, Alexandrou A. Determining authenticity of video evidence in the age of artificial intelligence and in the wake of Deepfake videos. *The International Journal of Evidence and Proof.* 2019;23(3):255–262. DOI: 10.1177/1365712718807226.
52. Citron DK, Chesney R. Deep fakes: A looming challenge for privacy, democracy, and national security. *California Law Review.* 2019;107(6):1753–1819.
53. Memarian B, Doleck T. ChatGPT in education: Methods, potentials and limitations. *Computers in Human Behavior: Artificial Humans.* 2023;1(6):100022. DOI: 10.1016/j.chbah.2023.100022.
54. Adeshola I, Adepoju AP. The opportunities and challenges of ChatGPT in education. *Interactive Learning Environments.* 2024;32(10):6159–6172. DOI: 10.1080/10494820.2023.2253858.
55. Berger S, Rossi F. AI and neurotechnology: learning from AI ethics to address an expanded ethics landscape. *Communications of the ACM.* 2023;66(3):58–68. DOI: 10.1145/3529088.
56. Huang C, Zhang Z, Mao B, Yao X. An overview of artificial intelligence ethics. *IEEE Transactions on Artificial Intelligence.* 2022;4(4):799–819. DOI: 10.1109/TAI.2022.3194503.
57. Bostrom N. Ethical issues in advanced artificial intelligence. In: *Machine Ethics and Robot Ethics.* London: Routledge; 2020. P. 69–75. DOI: 10.4324/9781003074991-7.
58. Mittelstadt B. Principles alone cannot guarantee ethical AI. *Nat Mach Intell.* 2019;1(11):501–507. DOI: doi.org/10.1038/s42256-019-0114-4.
59. Ruiz S, Valera L, Ramos P, Sitaram R. Neurorights in the Constitution: from neurotechnology to ethics and politics. *Philosophical Transactions B.* 2024;379(1915):20230098. DOI: 10.1098/rstb.2023.0098.
60. Robinson JT, Rommelfanger KS, Anikeeva PO, Etienne A, French J, Gelinis J, Grover P, Picard R. Building a culture of responsible neurotech: Neuroethics as socio-technical challenges. *Neuron.* 2022;110(13):2057–2062. DOI: 10.1016/j.neuron.2022.05.005.
61. Illes J. *Neuroethics: Anticipating the Future.* NY: Oxford University Press; 2017.
62. Sample M., Racine E. Pragmatism for a digital society: The (in)significance of artificial intelligence and neural technology. In: Friedrich O., Wolkenstein A., Bublitz C., Jox R. J., Racine E., editors. *Clinical Neurotechnology meets Artificial Intelligence. Advances in Neuroethics.* Cham: Springer; 2021. P. 81–100. DOI: 10.1007/978-3-030-64590-8_7.
63. Savage N. How AI and neuroscience drive each other forwards. *Nature.* 2019;571(7766):S15–S17. DOI: 10.1038/d41586-019-02212-4.
64. Raichle ME. A brief history of human brain mapping. *Trends Neurosci.* 2009;32(2):118–126. DOI: 10.1016/j.tins.2008.11.001.
65. Friston KJ. Modalities, modes, and models in functional neuroimaging. *Science.* 2009;326(5951):399–403. DOI: 10.1126/science.1174521.
66. Lefaucheur JP, Aleman A, Baeken C, Benninger DH, Brunelin J, Di Lazzaro V, Filipović SR, Grefkes C, Hasan A, Hummel FC, Jääskeläinen SK, Langguth B, Leocani L, Londero A, Nardone R, NguyenJ-P, Nyffeler Th, Oliveira-Maia AJ, Oliviero A, Padberg F, Palm U, Paulus W, Poulet E, Quartarone A, Rachid F, Rektorová I, Rossi S, Sahlsten H, Schecklmann M, Szekely D, Ziemann U. Evidence-based guidelines on the therapeutic use of repetitive transcranial magnetic stimulation (rTMS): An update (2014-2018). *Clin Neurophysiol.* 2020;131(5):474–528. DOI: 10.1016/j.clinph.2019.11.002.
67. Benabid AL, Chabardes S, Mitrofanis J, Pollak P. Deep brain stimulation of the subthalamic nucleus for the treatment of Parkinson’s disease. *Lancet Neurol.* 2009;8(1):67–81. DOI: 10.1016/S1474-4422(08)70291-6.
68. Fisher R, Salanova V, Witt T, Worth R, Henry T, Gross R, Oommen K, Osorio I, Nazzaro J, Labar D, Kaplitt M, Sperling M, Sandok E, Neal J, Handforth A, Stern J, DeSalles A, Chung S,

- Shetter A, Bergen D, Bakay R, Henderson J, French J, Baltuch G, Rosenfeld W, Youkilis A, Marks W, Garcia P, Barbaro N, Fountain N, Bazil C, Goodman R, McKhann G, Krishnamurthy KB, Papavassiliou S, Epstein Ch, Pollard J, Tonder L, Grebin J, Coffey R, Graves N, the SANTE Study Group. Electrical stimulation of the anterior nucleus of thalamus for treatment of refractory epilepsy. *Epilepsia*. 2010;51(5):899–908. DOI: 10.1111/j.1528-1167.2010.02536.x.
69. Bonomo R, Elia AE, Bonomo G, Romito LM, Mariotti C, Devigili G, Cilia R, Giossi R, Eleopra R. Deep brain stimulation in Huntington’s disease: a literature review. *Neurological Sciences*. 2021;42(11):4447–4457. DOI: 10.1007/s10072-021-05527-1.
 70. Ramachandran VS, Rogers-Ramachandran D. Phantom limbs and neural plasticity. *Arch Neurol*. 2000;57(3):317–320. DOI: 10.1001/archneur.57.3.317.
 71. Khorev V, Kurkin S, Badarin A, Antipov V, Pitsik E, Andreev A, Grubov V, Drapkina O, Kiselev A, Hramov A. Review on the use of brain computer interface rehabilitation methods for treating mental and neurological conditions. *J Integr Neurosci*. 2024;23(7):125. DOI: 10.31083/j.jin2307125.
 72. Kuo MF, Nitsche MA. Effects of transcranial electrical stimulation on cognition. *Clin EEG Neurosci*. 2012;43(3):192–199. DOI: 10.1177/1550059412444975.
 73. Anguera JA, Boccanfuso J, Rintoul JL, Al-Hashimi O, Faraji F, Janowich J, Kong E, Larraburo Y, Rolle C, Johnston E, Gazzaley A. Video game training enhances cognitive control in older adults. *Nature*. 2013;501(7465):97–101. DOI: 10.1038/nature12486.
 74. Grubov VV, Khramova MV, Goman S, Badarin AA, Kurkin SA, Andrikov DA, Pitsik E, Antipov V, Petushok E, Brusinskii N, Bukina T, Fedorov AA, Hramov AE. Open-loop neuroadaptive system for enhancing student’s cognitive abilities in learning. *IEEE Access*. 2024;12:49034–49049. DOI: 10.1109/access.2024.3383847.
 75. Bukina TV, Khramova MV, Kurkin SA, Andrikov DA, Goman SS, Dedkov AE, Hramov AE. Neuroeducational software recommendational service as a tool for personalizing the educational process. *Informatics and Education*. 2024;39(5):50–62. DOI: 10.32517/0234-0453-2024-39-5-50-62.
 76. Donati AR, Shokur S, Morya E, Campos DS, Moioli RC, Gitti CM, Augusto PB, Tripodi S, Pires CG, Pereira GA, Brasil FL, Gallo S, Lin AA, Takigami AK, Aratanha MA, Joshi S, Bleuler H, Cheng G, Rudolph A, Nicolelis MAL. Long-term training with a brain-machine interface-based gait protocol induces partial neurological recovery in paraplegic patients. *Sci Rep*. 2016;6:30383. DOI: 10.1038/srep30383.
 77. Poppe C, Elger BS. Brain–computer interfaces, completely locked-in state in neurodegenerative diseases, and end-of-life decisions. *Journal of Bioethical Inquiry*. 2024;21(1):19–27. DOI: 10.1007/s11673-023-10256-5.
 78. Kübler A, Neumann N. Brain-computer interfaces—the key for the conscious brain locked into a paralyzed body. *Prog Brain Res*. 2005;150:513–525. DOI: 10.1016/S0079-6123(05)50035-9.
 79. Lebedev MA, Nicolelis MA. Brain-machine interfaces: past, present and future. *Trends Neurosci*. 2006;29(9):536–546. DOI: 10.1016/j.tins.2006.07.004.
 80. Veena N, Anitha N. A review of non-invasive BCI devices. *Int J Biomed Eng Technol*. 2020;34(3):205–233. DOI: 10.1504/IJBET.2020.111471.
 81. Grill WM, Norman SE, Bellamkonda RV. Implanted neural interfaces: biochallenges and engineered solutions. *Annu Rev Biomed Eng*. 2009;11:1–24. DOI: 10.1146/annurev-bioeng-061008-124927.
 82. Salatino JW, Ludwig KA, Kozai TDY, Purcell EK. Glial responses to implanted electrodes in the brain. *Nat Biomed Eng*. 2017;1(11):862–877. DOI: 10.1038/s41551-017-0154-1.
 83. Rossi S, Hallett M, Rossini PM, Pascual-Leone A, The Safety of TMS Consensus Group. Safety, ethical considerations, and application guidelines for the use of transcranial magnetic stimulation in clinical practice and research. *Clin Neurophysiol*. 2009;120(12):2008–2039. DOI: 10.1016/j.clinph.2009.08.016.
 84. Kringelbach ML, Jenkinson N, Owen SL, Aziz TZ. Translational principles of deep brain stimulation. *Nature Reviews Neuroscience*. 2007;8(8):623–635. DOI: 10.1038/nrn2196.

85. Bergey GK, Morrell MJ, Mizrahi EM, Goldman A, King-Stephens D, Nair D, Srinivasan S, Jobst B, Gross RE, Shields DC, Barkley G, Salanova V, Olejniczak P, Cole A, Cash SS, Noe K, Wharen R, Worrell G, Murro AM, Edwards J, Duchowny M, Spencer D, Smith M, Geller E, Gwinn R, Skidmore Ch, Eisenschenk S, Berg M, Heck Ch, Van Ness P, Fountain N, Rutecki P, Massey A, O'Donovan C, Labar D, Duckrow RB, Hirsch LJ, Courtney T, Sun FT, Seale CG. Long-term treatment with responsive brain stimulation in adults with refractory partial seizures. *Neurology*. 2015;84(8):810–817. DOI: 10.1212/WNL.0000000000001280.
86. Deer TR, Pope JE, Hayek SM, Bux A, Buchser E, Eldabe S, De Andrés JA, Erdek M, Patin D, Grider JS, Doleys DM, Jacobs MS, Yaksh TL, Poree L, Wallace MS, Prager J, Rauck R, DeLeon O, Diwan S, Falowski SM, Gazelka HM, Kim P, Leong M, Levy RM, McDowell G II, McRoberts P, Naidu R, Narouze S, Perruchoud C, Rosen SM, Rosenberg WS, Saulino M, Staats P, Stearns LJ, Willis D, Krames E, Huntoon M, Mekhail N. The Polyanalgesic Consensus Conference (PACC): Recommendations on Intrathecal Drug Infusion Systems Best Practices and Guidelines. *Neuromodulation*. 2017;20(2):96–132. DOI: 10.1111/ner.12538.
87. Deisseroth K. Optogenetics: 10 years of microbial opsins in neuroscience. *Nat Neurosci*. 2015; 18(9):1213–1225. DOI: 10.1038/nn.4091.
88. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*. 2015;521(7553):436–444. DOI: 10.1038/nature14539.
89. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser L, Polosukhin I. Attention is all you need. In: Guyon I, Luxburg UV, Bengio S, Wallach H, Fergus R, Vishwanathan S, Garnett R, editors. *Advances in Neural Information Processing Systems*. vol. 30. Curran Associates, Inc.; 2017. p. 5998–6008.
90. Goodfellow IJ, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Bengio Y. Generative adversarial nets. In: *Proceedings of the 28th International Conference on Neural Information Processing Systems - Volume 2. NIPS'14*. Cambridge: MIT Press; 2014. p. 2672–2680.
91. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. In: *Proceedings of the 26th International Conference on Neural Information Processing Systems - Volume 1. NIPS'12*. Red Hook, NY, USA: Curran Associates Inc.; 2012. p. 1097–1105.
92. He K, Zhang X, Ren S, Sun J. Deep Residual Learning for Image Recognition. In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*; 2016. P. 770–778. DOI: 10.1109/CVPR.2016.90.
93. Qi CR, Su H, Mo K, Guibas LJ. PointNet: Deep learning on point sets for 3D classification and segmentation. In: *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*; 2017. p. 77–85. DOI: 10.1109/CVPR.2017.16.
94. Girshick R, Donahue J, Darrell T, Malik J. Rich feature hierarchies for accurate object detection and semantic segmentation. In: *2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*; 2014. p. 580–587. DOI: 10.1109/CVPR.2014.81.
95. Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM, Thrun S. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*. 2017;542(7639):115–118. DOI: 10.1038/nature21056.
96. LeCun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*. 1998;86(11):2278–2324. DOI: 10.1109/5.726791.
97. Chen C, Seff A, Kornhauser A, Xiao J. DeepDriving: Learning affordance for direct perception in autonomous driving. In: *2015 IEEE International Conference on Computer Vision (ICCV)*; 2015. p. 2722–2730. DOI: 10.1109/ICCV.2015.312.
98. Loquercio A, Maqueda AI, del Blanco CR, Scaramuzza D. DroNet: Learning to fly by driving. *IEEE Robotics and Automation Letters*. 2018;3(2):1088–1095. DOI: 10.1109/LRA.2018.2795643.
99. Pinto L, Gupta A. Supersizing self-supervision: Learning to grasp from 50K tries and 700 robot hours. In: *2016 IEEE International Conference on Robotics and Automation (ICRA)*; 2016. p. 3406–3413. DOI: 10.1109/ICRA.2016.7487517.

100. Kober J, Bagnell JA, Peters J. Reinforcement learning in robotics: A survey. *The International Journal of Robotics Research*. 2013;32(11):1238–1274. DOI: 10.1177/0278364913495721.
101. Koopman P, Wagner M. Autonomous vehicle safety: An interdisciplinary challenge. *IEEE Intelligent Transportation Systems Magazine*. 2017;9(1):90–96. DOI: 10.1109/MITS.2016.2583491.
102. Leveson NG. *Engineering a Safer World: Systems Thinking Applied to Safety*. Cambridge: MIT Press; 2011. DOI: 10.7551/mitpress/8179.001.0001.
103. Jurafsky D, Martin JH. *Speech and Language Processing*. Harlow: Pearson Education; 2020.
104. Goodman B, Flaxman S. European Union regulations on algorithmic decision-making and a “right to explanation”. *AI Magazine*. 2017;38(3):50–57.
105. Ribeiro MT, Singh S, Guestrin C. “Why Should I Trust You?”: Explaining the predictions of any classifier. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. KDD '16. New York, NY, USA: Association for Computing Machinery; 2016. p. 1135–1144. DOI: 10.1145/2939672.2939778.
106. Caruana R, Lou Y, Gehrke J, Koch P, Sturm M, Elhadad N. Intelligible models for healthcare: Predicting pneumonia risk and hospital 30-day readmission. In: *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. KDD '15. NY: Association for Computing Machinery; 2015. P. 1721–1730. DOI: 10.1145/2783258.2788613.
107. Bassett DS, Khambhati AN. A network engineering perspective on probing and perturbing cognition with neurofeedback. *Ann. NY Acad. Sci.* 2017;1396(1):126–143. DOI: 10.1111/nyas.13338.
108. Jiang F, Jiang Y, Zhi H, Dong Y, Li H, Ma S, Wang Y. Artificial intelligence in healthcare: past, present and future. *Stroke Vasc Neurol*. 2017;2(4):230–243. DOI: 10.1136/svn-2017-000101.
109. Yuste R, Goering S, Arcas BAY, Bi G, Carmena JM, Carter A, Fins JJ, Friesen P, Gallant J, Huggins JE, Illes J, Kellmeyer P, Klein E, Marblestone A, Mitchell C, Parens E, Pham M, Rubel A, Sadato N, Sullivan LS, Teicher M, Wasserman D, Wexler A, Whittaker M, Wolpaw J. Four ethical priorities for neurotechnologies and AI. *Nature*. 2017;551(7679):159–163. DOI: 10.1038/551159a.
110. Plis SM, Hjelm DR, Salakhutdinov R, Allen EA, Bockholt HJ, Long JD, Johnson HJ, Paulsen JS, Turner JA, Calhoun VD. Deep learning for neuroimaging: a validation study. *Front Neurosci*. 2014;8:229. DOI: 10.3389/fnins.2014.00229.
111. Pisarchik AN, Maksimenko VA, Andreev AV, Frolov NS, Makarov VV, Zhuravlev MO, Runnova AE, Hramov AE. Coherent resonance in the distributed cortical network during sensory information processing. *Sci Rep*. 2019;9(1):18325. DOI: 10.1038/s41598-019-54577-1.
112. Dedkov AE, Andrikov DA, Hramov AE. A review of ways to measure brain cognitive load and machine learning methods for their identification from EEG data. *Medical Doctor and Information Technologies*. 2024(3):20–31. DOI: 10.25881/18110193_2024_3_20.
113. Shatte ABR, Hutchinson DM, Teague SJ. Machine learning in mental health: A scoping review of methods and applications. *Psychol Med*. 2019;49(9):1426–1448. DOI:10.1017/S0033291719000151.
114. Grubov VV, Nazarikov SI, Kurkin SA, Utyashev NP, Andrikov DA, Karpov OE, Hramov AE. Two-stage approach with combination of outlier detection method and deep learning enhances automatic epileptic seizure detection. *IEEE Access*. 2024;12:122168–122182. DOI: 10.1109/access.2024.3453039.
115. Pitsik EN, Maximenko VA, Kurkin SA, Sergeev AP, Stoyanov D, Paunova R, Kandilarova S, Simeonova D, Hramov AE. The topology of fMRI-based networks defines the performance of a graph neural network for the classification of patients with major depressive disorder. *Chaos, Solitons and Fractals*. 2023;167:113041. DOI: 10.1016/j.chaos.2022.113041.
116. Zhang S, Zhao H, Wang W, Wang Z, Luo X, Hramov A, Kurths J. Edge-centric effective connection network based on multi-modal MRI for the diagnosis of Alzheimer’s disease. *Neurocomputing*. 2023;552:126512. DOI: 10.1016/j.neucom.2023.126512.
117. Park KH, Sun S, Lim YH, Park HR, Lee JM, Park K, Jeon B, Park H-P, Kim HCh, Paek SH. Clinical outcome prediction from analysis of microelectrode recordings using deep learning in

- subthalamic deep brain stimulation for Parkinson's disease. PLoS ONE. 2021;16(1):e0244133. DOI: 10.1371/journal.pone.0244133.
118. Hubel DH, Wiesel TN. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. J Physiol. 1962;160(1):106–154. DOI: 10.1113/jphysiol.1962.sp006837.
 119. Behrouz A, Zhong P, Mirrokni V. Titans: Learning to Memorize at Test Time. arXiv:250100663 arXiv Preprint. 2024. DOI: 10.48550/arXiv.2501.00663.
 120. Stanton SJ, Sinnott-Armstrong W, Huettel SA. Neuromarketing: Ethical implications of its use and potential misuse. J Bus Ethics. 2017;144:799–811. DOI: 10.1007/s10551-016-3059-0.
 121. Luna-Nevarez C. Neuromarketing, ethics, and regulation: An exploratory analysis of consumer opinions and sentiment on blogs and social media. J Consum Policy. 2021;44(4):559–583. DOI: 10.1007/s10603-021-09496-y.
 122. Maron DF. Science career ads are disproportionately seen by men. Scientific American [Electronic resource] Available from: <https://www.scientificamerican.com/article/science-career-ads-are-disproportionately-seen-by-men>. 2018.
 123. Goering S, Yuste R. On the necessity of ethical guidelines for novel neurotechnologies. Cell. 2016;167(4):882–885. DOI: 10.1016/j.cell.2016.10.029.
 124. Leefmann J, Levallois C, Hildt E. Neuroethics 1995–2012. A bibliometric analysis of the guiding themes of an emerging research field. Front Hum Neurosci. 2016;10:336. DOI: 10.3389/fnhum.2016.00336.
 125. Mehrabi N, Morstatter F, Saxena N, Lerman K, Galstyan A. A survey on bias and fairness in machine learning. ACM Computing Surveys. 2021;54(6):1–35. DOI: 10.1145/3457607.
 126. Barocas S, Selbst AD. Big Data's Disparate Impact. California Law Review. 2016;104(3):671–732. DOI: 10.2139/ssrn.2477899.
 127. O'Neil C. Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. New York: Crown Publishing Group; 2016. 259 p.
 128. Caliskan A, Bryson JJ, Narayanan A. Semantics derived automatically from language corpora contain human-like biases. Science. 2017;356(6334):183–186. DOI: 10.1126/science.aal4230.
 129. Raghavan M, Barocas S, Kleinberg J, Levy K. Mitigating bias in algorithmic hiring: evaluating claims and practices. In: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency. FAT* '20. New York, NY, USA: Association for Computing Machinery; 2020. P. 469–481. DOI: 10.1145/3351095.3372828.
 130. Bartlett R, Morse A, Stanton R, Wallace N. Consumer-lending discrimination in the FinTech era. Journal of Financial Economics. 2022;143:30–56. DOI: 10.1016/j.jfineco.2021.05.047.
 131. Angwin J, Larson J, Mattu S, Kirchner L. Machine bias: There's software used across the country to predict future criminals. And it's biased against blacks. ProPublica [Electronic resource]. URL: <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>. 2016.
 132. Van Deursen AJAM, Van Dijk JPM. The digital divide shifts to differences in usage. New Media & Society. 2014;16(3):507–526. DOI: 10.1177/1461444813487959.
 133. Yadav BR. The Ethics of Understanding: Exploring Moral Implications of Explainable AI. International Journal of Science and Research (IJSR). 2024;13(6):1–7. DOI:10.21275/SR24529122811.
 134. Kundu S. AI in medicine must be explainable. Nat Med. 2021;27(8):1328. DOI: 10.1038/s41591-021-01461-z.
 135. McDermid JA, Jia Y, Porter Z, Habli I. Artificial intelligence explainability: the technical and ethical dimensions. Philos Trans A Math Phys Eng Sci. 2021;379(2207):20200363. DOI: 10.1098/rsta.2020.0363.
 136. Angelov PP, Soares EA, Jiang R, Arnold NI, Atkinson PM. Explainable artificial intelligence: an analytical review. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery. 2021;11(5):e1424. DOI: 10.1002/widm.1424.
 137. Dwivedi R, Dave D, Naik H, Singhal S, Omer R, Patel P, Qian B, Wen Z, Shah T, Morgan G, Ranjan R. Explainable AI (XAI): Core ideas, techniques, and solutions. ACM Computing Surveys. 2023;55(9):835. DOI: 10.1145/3561048.

138. Minh D, Wang HX, Li YF, Nguyen TN. Explainable artificial intelligence: a comprehensive review. *Artif Intell Rev.* 2022;55:3503–3568. DOI: 10.1007/s10462-021-10088-y.
139. Ponce-Bobadilla AV, Schmitt V, Maier CS, Mensing S, Stodtmann S. Practical guide to SHAP analysis: Explaining supervised machine learning model predictions in drug development. *Clin Transl Sci.* 2024;17(11):e70056. DOI: 10.1111/cts.70056.
140. Islam R, Andreev AV, Shusharina NN, Hramov AE. Explainable machine learning methods for classification of brain states during visual perception. *Mathematics.* 2022;10(15):2819. DOI: 10.3390/math10152819.
141. Bhati D, Neha F, Amiruzzaman M. A survey on explainable artificial intelligence (xai) techniques for visualizing deep learning models in medical imaging. *Journal of Imaging.* 2024;10(10):239. DOI: 10.3390/jimaging10100239.
142. Lai T. Interpretable medical imagery diagnosis with self-attentive transformers: a review of explainable AI for health care. *BioMedInformatics.* 2024;4(1):113–126. DOI: 10.3390/biomedinformatics4010008.
143. Catherine S. B. A. Unlocking the Black Box: Exploring the Future of Explainable AI in Real-World Applications. In: *Multidisciplinary Research in Arts, Science & Commerce.* Vol. 25. The Hill Publication; 2025. P. 37–38.
144. Hassija V, Chamola V, Mahapatra A, Singal A, Goel D, Huang K, Scardapane S, Spinelli I, Mahmud M, Hussain A. Interpreting black-box models: a review on explainable artificial intelligence. *Cogn Comput.* 2024;16:45–74. DOI: 10.1007/s12559-023-10179-8.
145. Karpov OE, Andrikov DA, Maksimenko VA, Hramov AE. Explainable artificial intelligence for medicine. *Medical Doctor and Information Technologies.* 2022;(2):4–11. DOI: 10.25881/18110193_2022_2_4.
146. Černevičienė J, Kabašinskas A. Explainable artificial intelligence (XAI) in finance: a systematic literature review. *Artif Intell Rev.* 2024;57:216. DOI: 10.1007/s10462-024-10854-8.
147. Craig T. *Privacy and Big Data: The Players, Regulators, and Stakeholders.* Sebastopol: O'Reilly Media, Inc.; 2011. 91 p.
148. Curzon J, Kosa TA, Akalu R, El-Khatib K. Privacy and artificial intelligence. *IEEE Transactions on Artificial Intelligence.* 2021;2(2):96–108. DOI: 10.1109/TAI.2021.3088084.
149. Chesterman S. Artificial intelligence and the problem of autonomy. *Notre Dame Journal on Emerging Technologies.* 2020;1:210–250.
150. Novelli C, Taddeo M, Floridi L. Accountability in artificial intelligence: what it is and how it works. *AI and Soc.* 2024;39:1871–1882. DOI: 10.1007/s00146-023-01635-y.
151. Oullier O, editor. *Le cerveau et la loi: analyse de l'émergence du neurodroit.* Paris: Centre d'analyse stratégique; 2012. Available from: http://archives.strategie.gouv.fr/cas/system/files/cas-dqs_dt-neurodroit_11septembreduit_0.pdf.
152. Eagleman DM. Pourquoi les sciences du cerveau peuvent éclairer le droit. In: Oullier O, editor. *Le cerveau et la loi: analyse de l'émergence du neurodroit.* Paris: Centre d'analyse stratégique; 2012. P. 33–52.
153. Ernst E, Merola R, Samaan D. Economics of artificial intelligence: Implications for the future of work. *IZA Journal of Labor Policy.* 2019;9(1):4. DOI: 10.2478/izajolp-2019-0004.
154. McGaughey E. Will robots automate your job away? Full employment, basic income and economic democracy. *Industrial Law Journal.* 2022;51(3):511–559. DOI: 10.1093/indlaw/dwab010.
155. Russell S. *Human Compatible: Artificial Intelligence and the Problem of Control.* New York: Viking; 2019.
156. Torres P. Superintelligence and the future of governance: On prioritizing the control problem at the end of history. In: *Artificial Intelligence Safety and Security.* Chapman and Hall/CRC; 2018. p. 357–374. DOI: 10.1201/9781351251389-24.
157. Totschnig W. The problem of superintelligence: Political, not technological. *AI and Soc.* 2019;34:907–920. DOI: 10.1007/s00146-017-0753-0.
158. Yandex: Risk Assessment Approach : electronic document. Available from: <https://yandex.ru/>.

159. da Silva Castanheira J, Orozco Perez HD, Misic B, Baillet S. Brief segments of neurophysiological activity enable individual differentiation. *Nat Commun.* 2021;12:5713. DOI: 10.1038/s41467-021-25895-8.
160. Shvab K. *Technologies of the Fourth Industrial Revolution.* M.: Eksmo; 2018. 317 p.
161. Ienca M. On artificial intelligence and manipulation. *Topoi.* 2023;42(3):833–842. DOI: 10.1007/s11245-023-09940-3.
162. Lemsieh H, Abarar I. Artificial intelligence: A technological tool to manipulate the psychology and behavior of consumers: Theoretical research. *International Journal of Accounting, Finance, Auditing, Management and Economics.* 2024;5(6):432–449. DOI: 10.5281/zenodo.12186136.
163. Pisarchik AN, Musatov VYu, Runnova AE, Pchelintseva CV, Hramov AE. Method of person identification by eeg-response to ambiguous images. Patent No. 2653239 C1 Russian Federation, IPC G06K 9/66, A61B 5/0476, H04L 9/32: appl. 02.05.2017: publ. 07.05.2018. Assignee: Yuri Gagarin Saratov State Technical University (in Russian).
164. Panch T., Mattie H, Celi LA. The “inconvenient truth” about AI in healthcare. *NPJ Digit Med.* 2019;2:77. DOI: 10.1038/s41746-019-0155-4.



Шендерюк-Жидков Александр Владимирович — родился в Калининграде (1982). Окончил с отличием экономический факультет Калининградского государственного университета по специальности «Финансы и кредит» (2004). Сенатор, член Совета Федерации Федерального Собрания Российской Федерации от Калининградской области, первый зампред Комитета по бюджету и финансовым рынкам Совета Федерации, доктор h. s. БФУ им. И. Канта.

Россия, 103426 Москва, ул. Б. Дмитровка, 26
 Совет Федерации Федерального Собрания Российской Федерации
 E-mail: a.shenderyuk@yandex.ru
 ORCID: 0009-0006-3206-8124
 AuthorID (eLibrary.Ru): 1210392



Максименко Владимир Александрович — родился в Саратове (1987). Окончил с отличием факультет нелинейных процессов Саратовского государственного университета (2012). Доктор физико-математических наук (2020, СГТУ). Лауреат Премии Президента Российской Федерации в области науки и инноваций для молодых учёных (2020). Ведущий научный сотрудник Национального университета Сингапура и ведущий научный сотрудник АНО «Неймарк» (Нижний Новгород).

Россия, 603104 Нижний Новгород, ул. Нартова, д. 6, пом. П1а
 Автономная некоммерческая организация «Неймарк»
 E-mail: maximenkovl@gmail.com
 ORCID: 0000-0002-4632-6896
 AuthorID (eLibrary.Ru): 742773



Храмов Александр Евгеньевич — родился в Саратове (1974). Окончил с отличием физический факультет Саратовского государственного университета по специальности «Радиофизика и электроника» (1996). Доктор физико-математических наук (2005, СГУ), профессор (2008), член-корреспондент РАН (2025). Главный научный сотрудник ФГАУ «Цифровые промышленные технологии» (Москва). С 2025 года — директор Научно-исследовательского института прикладного искусственного интеллекта и цифровых решений Российского экономического университета им. Г. В. Плеханова

Россия, 115054 Москва, Стремянный переулок, 36
 Российский экономический университет имени Г. В. Плеханова
 Россия, 127018 Москва, ул. Сущёвский Вал, 18
 ФГАУ «Цифровые промышленные технологии»
 E-mail: hramov.ae@rea.ru
 ORCID: 0000-0003-2787-2530
 AuthorID (eLibrary.Ru): 34834