https://doi.org/10.38044/2686-9136-2025-6-12

СТАТЬИ

ИСПОЛЬЗОВАНИЕ ПЕРСОНАЛЬНЫХ ДАННЫХ ДЛЯ ОБУЧЕНИЯ МОДЕЛЕЙ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В ПРАВЕ ЕВРОПЕЙСКОГО СОЮЗА

А. А. Олифиренко[1,2]

[1]Саратовская государственная юридическая академия
410056, Россия, Саратов, ул. Вольская, 1

[2]Саратовский государственный технический университет
им. Ю. А. Гагарина
410008, Россия, Саратов, ул. Политехническая, 77к1

Аннотация

Принятие европейского Акта об искусственном интеллекте (AI Act) закрепило обязательное регулирование жизненного цикла систем искусственного интеллекта в Европейском союзе при сохранении действия Общего регламента по защите данных (GDPR). Стадия обучения ИИ-моделей оказалась в зоне пересечения двух режимов: Акт об ИИ ориентирует участников на качество и репрезентативность наборов данных, управление рисками и документирование процесса обучения, тогда как Общий регламент о защите данных фиксирует принципы правомерности, минимизации, ограничения целей и сроков хранения, а также предоставляет субъекту персональных данных комплекс гарантий и средств защиты. В практических проектах это создает риск юридически дефектного обучения, когда стремление к репрезентативности реализуется через избыточный сбор и повторное использование данных. Статья исследует допустимость и организацию обучения ИИ-моделей при совместном применении Акта об ИИ и Общего регламента о защите данных. Цель исследования состоит в обосновании правовой модели, позволяющей выстроить соразмерные технические и организационные гарантии, сохранить качество обучения и одновременно обеспечить законность обработки персональных данных и защиту основных прав. Методология включает нормативно-догматический анализ требований Акта об ИИ к системе управления рисками и управлению данными, сопоставление с принципами Общего регламента о защите данных и процедурными инструментами обеспечения законности обработки, а также систематизацию типовых управленческих артефактов, используемых при разработке и эксплуатации высокорисковых ИИ-систем. Результаты исследования представлены как согласованная модель поведения участников на стадии обучения. Обосновано прикладное разграничение «ИИ-система» и «ИИ-модель»: система квалифицируется как организационно-техническая оболочка, включающая модель, инфраструктуру, интерфейсы ввода и вывода, мониторинг и взаимодействие с человеком, тогда как модель рассматривается как алгоритмическое ядро, обученное на данных и применяемое для вывода результатов; данное разграничение используется для распределения обязанностей провайдера и лиц, внедряющих или эксплуатирующих систему. Предложен механизм согласования репрезентативности и достоверности наборов данных с принципом минимизации через документируемую инвентаризацию признаков, обоснование необходимости каждого класса данных и исключение нерелевантных атрибутов с одновременной оценкой риска косвенной дискриминации. Разработана модель

соразмерности защитных мер, связывающая выбор псевдонимизации, анонимизации, агрегирования, синтетической генерации и дифференциальной приватности с чувствительностью данных, контекстом использования и уровнем риска для основных прав, подтверждаемым результатами оценки воздействия на защиту данных и оценки воздействия на основные права. Сформулирован практический контур правового обеспечения жизненного цикла обучения: постановка цели и правового основания, ограничение повторного использования наборов данных, контроль доступа и журналирование операций, правила сроков хранения и удаления, а также процедуры пересмотра параметров обучения и мониторинга после внедрения. Предложенная модель повышает предсказуемость правоприменения и задает воспроизводимый порядок согласования требований Акта об ИИ и Общего регламента о защите данных на стадии обучения.

Ключевые слова

искусственный интеллект, обучение модели искусственного интеллекта, правовое регулирование персональных данных, минимизация данных, соразмерность мер, Общий регламент о защите данных, Регламент об искусственном интеллекте, комплаенс, жизненный цикл ИИ

Конфликт интересов	Автор сообщает об отсутствии конфликта интересов.
Финансирование	Исследование не имеет спонсорской поддержки.
Для цитирования	Олифиренко, А. А. (2025). Использование персональных данных для обучения моделей искусственного интеллекта в праве Европейского союза. *Цифровое право*. https://doi.org/10.38044/2686-9136-2025-6-12

Поступила: 03.07.2025; принята в печать: 20.08.2025

ARTICLES

USING PERSONAL DATA IN AI MODEL TRAINING UNDER EU LAW

Artem A. Olifirenko[1,2]

[1]Saratov State Law Academy
1, Volskaya St., Saratov, Russia, 410056

[2]Yuri Gagarin State Technical University of Saratov
77-1, Polytechnicheskaya St., Saratov, Russia, 410008

Abstract

The adoption of the EU Artificial Intelligence Act (AI Act) established mandatory life-cycle regulation of AI systems in the European Union while preserving the validity of the General Data Protection Regulation (GDPR). The training stage of AI models has consequently become a point of intersection between two regulatory regimes: while the AI Act emphasizes data quality and representativeness along with risk management and documentation of training processes, the GDPR sets out the applicable principles of lawfulness, data minimization, purpose, and storage limitation, as well as providing data subjects with a set of safeguards and remedies. In practical terms, this interaction creates a risk of legally defective model training due to the pursuit of representativeness through excessive data collection and repeated re-use of personal data. This article examines the permissibility

and organization of AI model training under the joint application of the AI Act and the GDPR. The research sets out to substantiate a legal model that enables proportionate technical and organizational safeguards while preserving training quality and ensuring the lawfulness of personal data processing that respects the fundamental rights of data subjects. As well as combining doctrinal legal analysis of the AI Act requirements on risk management and data governance with a comparative assessment of the GDPR principles and procedural tools for ensuring lawful processing, the methodology involves a systematization of typical governance artefacts used in the development and deployment of high-risk AI systems. The results are presented as an integrated compliance-by-design model for actors involved in the training stage. A practical distinction between an "AI system" and an "AI model" is substantiated: whereas an AI system is qualified as an organizational and technical envelope comprising the model, infrastructure, input and output interfaces, monitoring, and human interaction, an AI model is treated as the algorithmic core trained on data and used to infer outputs. This distinction can be applied to allocate obligations between the provider and entities deploying or operating the system. The proposed mechanism for reconciling dataset representativeness and accuracy with the GDPR data minimization principle through a documented feature inventory is based on a necessity rationale for each class of data and the exclusion of irrelevant attributes alongside an assessment of indirect discrimination risks. The choice of safeguards (pseudonymization, anonymization, aggregation, synthetic generation, and differential privacy) to data sensitivity, use context, and the level of risk to fundamental rights is carried out on the basis of a proportionality model. This model is supported by the outcomes of a data protection impact assessment and a fundamental rights impact assessment. Finally, a practical legal governance loop for the training life cycle is formulated to cover the determination of the purpose and legal basis, limits on dataset re-use, access control and logging, as well as retention and deletion rules, along with procedures for revisiting training parameters and monitoring after deployment. The proposed model increases legal certainty and provides a reproducible framework for aligning the AI Act and GDPR during the training stage.

Keywords

artificial intelligence, AI model training, data protection law, data minimization, proportionality of safeguards, GDPR, AI Act, compliance, AI life cycle

Conflict of interest	The author declares no conflict of interest.
Financial disclosure	The study has no sponsorship.
For citation	Olifirenko, A. A. (2025). Using personal data in AI model training under EU law. *Digital Law Journal*. https://doi.org/10.38044/2686-9136-2025-6-12

Submitted: 3 July 2025, accepted: 20 Aug 2025

Введение

Принятие Регламента 2024/1689 об искусственном интеллекте (Artificial Intelligence Act, далее — Акт об ИИ)[1] ознаменовало переход Европейского союза от декларативного этического подхода к нормативно-правовой детализации жизненного цикла систем искусственного интеллекта (далее — ИИ-системы). В отличие от прежнего, преимущественно этического, дискурса новый акт закрепляет юридически обязательные требования к проектированию, обучению и эксплуатации систем искусственного интеллекта при одновременном сохранении действия

[1] Regulation 2024/1689, of the European Parliament and of the Council of 13 June 2024 laying down Harmonised Rules on Artificial Intelligence and Amending Regulations 300/2008, 167/2013, 168/2013, 2018/858, 2018/1139 and 2019/2144 and Directives 2014/90/EU, 2016/797 and 2020/1828 (Artificial Intelligence Act), 2024 O.J. (L 2024/1689) 1.

Общего регламента о защите данных (*General Data Protection Regulation,* далее — *GDPR*)[2]. Уже в ст. 2(7) Акта об ИИ подчеркивается, что он «не затрагивает и не уменьшает» объем защиты, гарантированной иным законодательством Союза, прежде всего *GDPR*, формируя для участников рынка режим двойной нормативной подотчетности[3]. В доктрине при этом просматриваются по меньшей мере два конкурирующих подхода: часть авторов предлагает рассматривать Акт об ИИ как *lex specialis* (специальное регулирование по отношению к общему режиму *GDPR*) для высокорисковых систем (Hadwick, 2024, pp. 947–948; Kindt, 2025, p. 31) (сходная аргументация встречается и в политико-правовых аналитических материалах, сопровождающих обсуждение и принятие Акта об ИИ, где отдельно фиксируются риски пересечений и конкуренции регуляторных режимов[4]); другие, напротив, настаивают на автономной значимости принципов защиты данных и воспринимают Акт об ИИ лишь как техническое уточнение уже существующих обязанностей (Winau, 2023; De Hert & Hajduk, 2024, p. 292; van Bekkum, 2025, p. 2). Автор настоящей работы придерживается промежуточной позиции, исходящей из невозможности свести один режим к другому без потери содержания.

Нормативная конструкция взаимодействия Акта об ИИ и *GDPR* не может быть сведен к простому суммированию обязанностей. Акт об ИИ формулирует углубленные технические и процессуальные требования (качество данных, управление рисками, документация, оценка воздействия на права), тогда как *GDPR* закрепляет фундаментальные принципы обработки персональных данных и архитектуру прав субъектов. Нарушение последних способно превратить даже формально выстроенный процесс обучения модели в юридически дефектный. Характерно, что аналитические материалы[5] к Акту об ИИ прямо подчеркивают: соблюдение самих по себе требований данного документа не гарантирует соответствия требованиям *GDPR*[6], а оценка влияния искусственного интеллекта на основные права[7] (*Fundamental Rights Impact Assesment,* далее — *FRIA*) в работах А. Мантелеро и других авторов рассматривается как самостоятельный способ контроля, как отдельная процедура, выходящая за рамки сугубо оценки рисков для защиты персональных данных и конфиденциальности (Mantelero, 2024, p. 9; Kaminski & Malgieri, 2024, pp. 314–315).

Наибольшее напряжение между двумя режимами проявляется на обучающей стадии, где сходятся противоположные методологические векторы. С одной стороны, статья 10 Акта об ИИ закрепляет императив высококачественных, репрезентативных и, по возможности, свободных от ошибок обучающих наборов наряду с обязанностью провайдера внедрять процедуры

[2] Regulation 2016/679, of the European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with regard to the Processing of Personal Data and on the Free Movement of Such Data, and repealing Directive 95/46/EC (General Data Protection Regulation), 2016 O.J. (L 119) 1.

[3] Artificial Intelligence Act, 2024 O.J. (L 2024/1689) 45.

[4] Voss, A. (2022). *Better regulation: Potential legal overlaps & contradictions with other laws*. European Parliament. https://www.axel-voss-europa.de/wp-content/uploads/2022/03/AVoss_AIA_BetterRegulation.pdf

[5] European Data Protection Board. (2024) *Statement 3/2024 on data protection authorities' role in the Artificial Intelligence Act framework.* https://www.edpb.europa.eu/system/files/2024-07/edpb_statement_202403_dpasroleaiact_en.pdf

[6] Cuatrecasas. (2024). *EU AI Act: A pioneering legal framework on artificial intelligence. Practical guide.* https://www.cuatrecasas.com/resources/eu-ai-act-practical-guide-669112324626b904914005.pdf?v1.94.0.20251119

[7] Под основными правами ("fundamental rights") понимаются права и свободы, закрепленные в Хартии Европейского союза об основных правах, то есть в надотраслевом стандарте защиты личности в правопорядке ЕС, охватывающим, в частности, достоинство, свободы, равенство и солидарность, а также права граждан и гарантии правосудия; в этом значении термин используется и в контексте регулирования ИИ. См.: Charter of Fundamental Rights of the European Union, 2012 O.J. (C 326) 391.

выявления и коррекции алгоритмических смещений. С другой стороны, *GDPR* квалифицирует любую загрузку данных в модель искусственного интеллекта (далее — ИИ-модель) как обработку персональных данных и предписывает принцип минимизации: обрабатывать лишь те данные, которые «адекватны, релевантны и ограничены необходимым для целей, для которых они обрабатываются» (ст. 5(1)(c) *GDPR*)[8]. Здесь важно отметить, что требование репрезентативности фактически размывает правило минимизации и способствует расширению массивов данных, тогда как другие исследователи предлагают рассматривать эти требования как комплементарные при условии строгого риск-ориентированного подхода (*risk-based approach*) и прозрачной документации об источниках и о структуре датасетов[9] (Paullada et al., 2021, pp. 1, 4, 10)[10]. В настоящей работе минимизация и репрезентативность рассматриваются не как автоматически согласованные принципы, а как напряженная связка, требующая дополнительных гарантий: процедур оценки воздействия на защиту данных (*Data Protection Impact Assessment, далее — DPIA*), на основные права (*FRIA*), ведения внутреннего журнала минимизации (обоснование включения каждого класса атрибутов и сроков хранения) и ограничений на вторичное использование обучающих, валидационных и тестовых наборов персональных данных и их производных версий (включая объединенные датасеты и повторно размеченные выборки) вне первоначально заявленных целей обучения.

Возникающая коллизия имеет по меньшей мере три аспекта:

1) вопрос о допустимом объеме и структуре выборки, необходимой для нивелирования предвзятости без утраты соразмерности обработки;
2) правовая судьба повторной и переносной переработки, когда исходные пользовательские данные применены для первоначального обучения, а затем привлекаются для иных целей, включая дополнительное обучение[11] и перенос обучения[12];
3) соразмерность технических и организационных мер категории риска модели с учетом дифференциации согласно ст. 6, 7 Акта об ИИ по уровням критичности и оценки тяжести воздействия на основные права (Malgieri & Santos, 2025 pp. 3, 7; Novelli et al., 2024b, pp. 1-2, 21)[13].

Особое значение имеет статья 10(5) Акта об ИИ[14], допускающая исключительную обработку специальных категорий персональных данных при разработке и проверке высокорисковых систем лишь в той мере, в какой это строго необходимо для обеспечения выявления и коррекции возможных предвзятостей (biases)[15] в обучающих, валидационных и тестовых наборах данных

[8] General Data Protection Regulation, 2016 O.J. (L 119) 35. См. об этом: International Organization for Standardization. (2022). *ISO/IEC 27001:2022: Information security, cybersecurity and privacy protection — Information security management systems — Requirements*. https://www.iso.org/standard/27001

[9] В оригинале — *"dataset governance"*, что означает управление жизненным циклом наборов данных.

[10] Niemiec, E., Davis, P., & Hauglid, M. (2024). *Will the EU AI Act help to eliminate dataset bias in medical AI?* SSRN. https://doi.org/10.2139/ssrn.5045561

[11] В первоисточнике *"fine-tuning"* характеризует дообучение модели на новых данных, в том числе на специализированных выборках, связанных с конкретной целевой задачей.

[12] В оригинале *"transfer learning"* подразумевает использование заранее обученной модели в новом контексте.

[13] Artificial Intelligence Act, 2024 O.J. (L 2024/1689) 53–55.

[14] См.: Ibid, at 58.

[15] См.: Artificial Intelligence Act, 2024 O.J. (L 2024/1689) 57–58. То есть *"biases"*, упомянутых в данной статье Акта, которые следует понимать как систематическую предвзятость (смещение), возникающую в данных или в результатах функционирования модели вследствие перекоса представительства отдельных групп, качества и полноты признаков, ошибок разметки, выбора критериев отбора и иных факторов, влияющих на распределение данных.

в рамках процедур управления данными, предусмотренных статьей 10(2)(f)–(g) Акта об ИИ[16], при условии соблюдения установленных данной нормой гарантий. Норма, ориентированная на снижение риска дискриминации и иных неблагоприятных эффектов, одновременно обостряет дилемму правомерности обработки: статья 9 *GDPR*[17] закрепляет общий запрет на обработку специальных категорий персональных данных, тогда как статья 10(5) Акта об ИИ[18] вводит узкоцелевое исключение, сопряженное с обязанностями по выявлению и снижению смещений и дополненное требованиями: обосновать невозможность достижения цели посредством обработки иных данных, включая синтетические или анонимизированные; применять современные меры безопасности и защиты данных, включая псевдонимизацию; ограничивать повторное использование и передачу третьим лицам, а также обеспечивать контроль доступа и документирование операций; удалять соответствующие данные после достижения цели коррекции смещений либо по истечении срока хранения. В доктрине данная конструкция обсуждается как вопрос о том, представляет ли указанное исключение допустимое точечное смягчение запрета статьи 9 *GDPR*[19] или, напротив, создает риск «ползучего расширения» (*function creep*)[20], при котором ссылка на устранение смещений фактически легитимирует устойчивую обработку чувствительной информации (van Bekkum, 2025, pp. 3, 7–8).

В отсутствие унифицированных методических разъяснений Европейской комиссии и Европейского управления искусственного интеллекта степень неопределенности усиливается недавними позициями надзорных органов. Так, в заключении 28/2024 Европейского совета по защите данных подчеркивается, что принцип минимизации «в полном объеме применим к разработке и развертыванию моделей» и что крупномасштабный веб-скрейпинг, т. е. автоматизированный сбор данных с веб-сайтов с помощью программных средств, не может автоматически опираться на исключение из информационной обязанности по ст. 14(5)(b) *GDPR*[21]. Тем самым подтверждается, что источники данных для обучения требуют индивидуальной юридической оценки, а массовый подход, предполагающий загрузку больших массивов данных без разбора, не соответствует стандартам *GDPR*. В литературе указанные выводы получают неоднозначную интерпретацию. С одной стороны, они рассматриваются как подтверждение того, что требования Акта об ИИ не подменяют и не устраняют автономную применимость *GDPR*: принципы *GDPR* и связанные оценочные процедуры сохраняют самостоятельное значение параллельно инструментам Акта об ИИ, включая *FRIA* (Mantelero, 2024, p. 4; Malgieri & Santos, 2025, pp. 1–2)[22]. С другой стороны, те же положения описываются как проявление структурной незавершенности многоуровневого регулирования, в рамках которого методологии оценки рисков и контуры правоприменения еще формируются через стандартизацию, подзаконные акты и руководства,

16 Artificial Intelligence Act, 2024 O.J. (L 2024/1689) 57.

17 General Data Protection Regulation, 2016 O.J. (L 119) 38–39.

18 См.: Ibid, at 58. В акте *"state of the art"* понимается как соответствующие высшему уровню развития техники и практики.

19 General Data Protection Regulation, 2016 O.J. (L 119) 38–39.

20 См. European Data Protection Board. (2020). *Guidelines 05/2020 on consent under Regulation 2016/679*. https://www.edpb.europa.eu/sites/default/files/files/file1/edpb_guidelines_202005_consent_en.pdf. Под "function creep" в данном контексте следует понимать риск последующего использования персональных данных для дополнительных задач, выходящих за пределы первоначально заявленной цели обработки.

21 European Data Protection Board. (2024). *Opinion 28/2024 on certain data protection aspects related to the processing of personal data in the context of AI models*. https://www.edpb.europa.eu/system/files/2024-12/edpb_opinion_202428_ai-models_en.pdf; General Data Protection Regulation, 2016 O.J. (L 119) 42.

22 См. также: Niemiec, E., Davis, P., & Hauglid, M. (2024). *Will the EU AI Act help to eliminate dataset bias in medical AI?* SSRN. https://doi.org/10.2139/ssrn.5045561

что объективно предполагает сосуществование и конкуренцию нескольких моделей комплаенса и enforcement-практик (Novelli et al., 2024b, p. 1; Söderlund & Larsson, 2024, p. 2; van Bekkum & Zuiderveen Borgesius, 2023, p. 2).

Указанные обстоятельства формируют ядро проблемного поля настоящего исследования. Его научная новизна обусловлена тем, что, во-первых, трансграничная цифровая экономика нуждается в оперативных, но юридически выверенных методиках подготовки обучающих наборов и документирования решений на обучающей стадии; во-вторых, существующая литература преимущественно концентрируется либо на догматическом анализе Акта об ИИ и его типовых схемах правоприменения и надзорной практики (*enforcement patterns*), либо на традиционных вопросах защиты персональных данных, не предлагая синтезированной модели согласованного применения обоих актов в области разработки и эксплуатации ИИ-систем. Отдельные подходы к оценке воздействия на права человека (*FRIA*), к управлению наборами данных через практики управления данных (*data governance*) и документирование решений на стадиях формирования и использования датасетов, а также к подготовке проверяемой (*audit-ready*) технической документации для целей комплаенса по Акту об ИИ (Mantelero, 2024, p. 4; Paullada et al., 2021, pp. 1, 7, 10; Sovrano et al., 2025, pp. 1, 4) пока не сведены в единую правовую конструкцию обучения ИИ-моделей[23].

Цель исследования состоит в обосновании и описании целостной правовой модели допустимого обучения моделей искусственного интеллекта, обеспечивающей согласованное применение Акта об ИИ и *GDPR* при сохранении качества и законности обучения.

Для достижения цели ставятся следующие задачи:

1) обозначить правовые контуры стадии обучения ИИ-моделей, выявив нормативную специфику и конкурирующие регуляторные векторы Акта об ИИ и *GDPR,* включая анализ категорий субъектов, их обязанностей и юридического значения терминов «модель», «система», «обучающие данные», «предназначенная цель»;
2) проанализировать трансформацию принципа минимизации данных в условиях потребности в больших и репрезентативных обучающих наборах и связать ее с требованиями к качеству данных и предотвращению дискриминации;
3) сформулировать нормативную модель соразмерности защитных мер, сопоставив положения (ст. 6–7, 9 и 10) Акта об ИИ[24] (риск-ориентированный подход к управлению рисками ИИ-систем) и ст. 25 *GDPR*[25] с учетом результатов *FRIA;*
4) предложить прикладную схему корпоративного комплаенса по стадиям жизненного цикла ИИ-модели (сбор данных, подготовка, обучение, развертывание, мониторинг), объединяющую требования обоих регламентов и позволяющую реализовать их на уровне процедур и документации.

Правовые контуры обучения ИИ-модели

Структурирование правового поля, в котором функционируют практики обучения искусственного интеллекта, невозможно без предварительного разграничения уровней регламентации

[23] См. об этом также: Siddik, M. (2024). *Datasheets for healthcare AI: A framework for transparency and bias mitigation.* OSF Preprints. https://doi.org/10.31219/osf.io/69ykb

[24] Artificial Intelligence Act, 2024 O.J. (L 2024/1689) 53–54, 56–58.

[25] Регламентирующую защиту персональных данных «по замыслу и по умолчанию» (*data protection by design and by default*). См.: General Data Protection Regulation, 2016 O.J. (L 119) 48.

и уточнения понятийного аппарата. Статья 2 Акта об ИИ прямо закрепляет, что его применение «не затрагивает и не уменьшает» гарантии, установленные иным правом Союза, прежде всего *GDPR*[26]. Тем самым формируется конструкция двойного регулирования: Акт об ИИ задает специальный режим для высокорисковых ИИ-систем, а *GDPR* остается горизонтальной нормой всеобщего действия, определяющей принципы обработки персональных данных и права субъектов[27].

Принцип неснижаемости стандартов *GDPR* превращает отношения между двумя актами не в классическую иерархию «общее — специальное», а в модель согласованного применения: выполнение процедур, предусмотренных Актом об ИИ, не освобождает от следования принципам законности, минимизации и целевого ограничения обработки. Любое технически корректное исполнение требований Акта об ИИ, не опирающееся на принципы *GDPR*, может быть квалифицировано как юридически дефектное. Это уже отражается в практических руководствах по комплаенсу в сфере искусственного интеллекта[28], где отдельно подчеркивается необходимость совмещать риск-ориентированный подход к управлению ИИ-системами и оценку воздействия на основные права.

Внутри этой конструкции ключевым становится вопрос о категориях правовых обязанностей. Акт об ИИ устанавливает следующие уровни требований, адресованных прежде всего провайдерам и импортерам высокорисковых систем:

- управление рисками (ст. 9)[29];
- управление данными: качество, происхождение, точность, безопасность (ст. 10)[30];
- техническая документация (ст. 11)[31];
- ведение регистрационных журналов (*record-keeping*) (ст. 12)[32];
- прозрачность и инструкции для пользователей (*transparency and instructions*) (ст. 13–14)[33] и др.

Эти нормы фактически переносят классические механизмы безопасности продукции (*product safety*) из сферы технического регулирования в область алгоритмической инженерии. *GDPR*, напротив, налагает обязанности в зависимости не от роли в цепочке поставок, а от факта обработки персональных данных[34]. В совокупности это образует *матрицу ответственности*, где каждая стадия жизненного цикла ИИ-системы соотносится с двумя наборами норм — Акта об ИИ и *GDPR*, а также с внутренними стандартами управления рисками и качеством (Söderlund & Larsson, 2024, pp. 10–11).

Примечательно, что в ст. 3(1) Акта об ИИ впервые закреплено нормативное определение ИИ-системы: машинная система, разработанная для работы с различным уровнем автономности и способная, для явных или неявных целей, выводимых человеком, генерировать результаты в виде прогнозов, рекомендаций или решений, влияющих на среду, с которой она

[26] Artificial Intelligence Act, 2024 O.J. (L 2024/1689) 45.

[27] De Luca, S. (2025, February 26). *Algorithmic discrimination under the AI Act and the GDPR*. European Parliament. https://www.europarl.europa.eu/thinktank/en/document/EPRS_ATA(2025)769509

[28] Bird & Bird. (2025). *European Union Artificial Intelligence Act: A guide*. https://www.twobirds.com/-/media/new-website-content/pdfs/capabilities/artificial-intelligence/european-union-artificial-intelligence-act-guide.pdf

[29] Ibid. at 56–57.

[30] Ibid. at 57–58.

[31] Ibid. at 58.

[32] Ibid. at 59.

[33] Ibid. at 59–61.

[34] Об этом см., в частности, следующие положения *GDPR*: ст. 5(1)(c), 6, 9, 12–23, 35 (General Data Protection Regulation, 2016 O.J. (L 119) 35, 36–37, 38–39, 39–47, 53–54).

взаимодействует[35]. Как следует из соображения 12 преамбулы, понятие ИИ-системы в Акте об ИИ[36] конструируется функционально, вокруг способности системы осуществлять алгоритмический вывод (*inference*)[37], а не вокруг конкретного семейства технологий. В сферу регулирования включаются как подходы машинного обучения[38], обучающиеся на данных, так и логико- и знание-ориентированные подходы (символьные/экспертные методы), тогда как «более простое» традиционное программное обеспечение, основанное исключительно на правилах, заданных человеком для автоматического исполнения операций, из охвата исключается. Дополнительно, согласно заключению 28/2024 Европейского совета по защите данных, подпадают также и нейросетевые архитектуры, но и системы, основанные на заранее заданных правилах (*rule-based systems*), статистические алгоритмы и гибридные решения[39]. Функционально-ориентированный подход смещает акцент с конкретной технологии на результат и контекст использования, что принципиально для правового анализа и оценки рисков; детерминированное программное обеспечение общего назначения при этом выводится за пределы действия акта.

Такое нормативное конструирование имеет два существенных следствия. Во-первых, все процедуры обучения (начальное обучение (*initial training*), переобучение (*retraining*), дообучение модели на новых данных) прямо включаются в жизненный цикл ИИ-системы и рассматриваются как объект регулирования Акта об ИИ. Во-вторых, статья 3 Акта об ИИ вводит категории «обучающие», «проверочные» и «тестовые данные» (*training, validation u testing data*), акцентируя требования к их качеству, репрезентативности и, по возможности, свободе от ошибок[40]. Эти характеристики непосредственно связаны с обязанностью провайдера по ст. 10 Акта об ИИ выявлять и устранять смещения в наборе данных (*dataset bias*)[41], т. е. структурную предвзятость, например недопредставленность отдельных групп или использование устаревших паттернов поведения, и связывают управление данными с принципом недискриминации и защитой фундаментальных прав (van Bekkum, 2025, p. 7)[42].

На этом фоне особое значение приобретает разграничение понятий «система искусственного интеллекта» (ИИ-система) и «модель искусственного интеллекта» (ИИ-модель). ИИ-система

---

35 Artificial Intelligence Act, 2024 O.J. (L 2024/1689) 46.

36 Ibid. at 4.

37 *Inference* — вывод (инференс), то есть способность ИИ-системы «выводить», как сгенерировать результаты (прогнозы, рекомендации или решения) на основе входных данных и целей, заданных человеком. Или рассуждение, с помощью которого делаются выводы по известным предпосылкам. См.: Федеральное агентство по техническому регулированию и метрологии (Росстандарт). (2024). *ГОСТ Р 71476–2024 (ИСО/МЭК 22989:2022). Искусственный интеллект. Концепции и терминология искусственного интеллекта*. https://docs.cntd.ru/document/1310068314

38 *Machine learning* — процесс автоматического обучения и совершенствования поведения системы искусственного интеллекта на основе обработки массива обучающих данных без явного программирования. См.: Федеральное агентство по техническому регулированию и метрологии (Росстандарт). (2021). *ГОСТ Р 59895–2021. Технологии искусственного интеллекта в образовании. Общие положения и терминология.* https://rst.gov.ru:8443/file-service/file/load/1682520552039

39 European Data Protection Board, *Opinion 28/2024 on Certain Data Protection Aspects Related to the Processing of Personal Data in the Context of AI Models* (Dec. 17, 2024), https://www.edpb.europa.eu/system/files/2024-12/edpb_opinion_202428_ai-models_en.pdf

40 Ibid. at 48.

41 Ibid. at 57–58.

42 См. об этом также: Niemiec, E., Davis, P., & Hauglid, M. (2024). *Will the EU AI Act help to eliminate dataset bias in medical AI?* SSRN. https://doi.org/10.2139/ssrn.5045561

в понимании Акта об ИИ определяется функционально как машинная система, способная осуществлять алгоритмический вывод (*inference*) и генерировать результаты (прогнозы, рекомендации или решения), влияющие на среду, в которой она применяется (ст. 3(1); п. 12 преамбулы)[43]. ИИ-модель, напротив, рассматривается как алгоритмическое ядро и компонент, который может быть встроен в ИИ-систему, но сам по себе не образует ИИ-систему (п. 97 преамбулы)[44]. Юридически обязанности провайдера по Акту об ИИ структурно привязаны к ИИ-системе как объекту вывода на рынок/ввода в эксплуатацию и обеспечения соответствия требованиям регламента (для высокорисковых ИИ-систем, в частности, ст. 16 Акта об ИИ)[45], тогда как на уровне модели концентрируются ключевые технические риски для персональных данных и основных прав, оценка которых в части защиты данных традиционно проводится через призму *GDPR* и смежных актов.

В настоящей работе основное внимание уделяется именно ИИ-модели как носителю рисков, связанных с обучающими данными. Для целей анализа используется в том числе понятие «базовая модель» (*foundation model*) — разновидность модели общего назначения (*general-purpose model*), описанная в пп. 97–100 преамбулы Акта об ИИ как крупная генеративная модель с миллиардом и более параметров, способная выполнять широкий спектр задач[46]. Такая модель не образует ИИ-систему до тех пор, пока не встроена в прикладной интерфейс и не связана с конкретным сценарием принятия решений; ее правовой режим, следовательно, определяется сочетанием требований Акта об ИИ к качеству и управлению данными и принципов *GDPR,* применимых к этапу обучения.

Особого внимания заслуживает понятие смещения в наборе данных (*dataset bias*). В статьях 10 и 44 Акта об ИИ[47] оно становится связующим звеном между принципом недискриминации (ст. 21 Хартии ЕС об основных правах[48]) и принципом минимизации данных. Законодатель прямо указывает, что одного лишь количественного критерия (размера выборки) недостаточно: необходим качественный анализ предвзятости и потенциальных дискриминационных эффектов, в том числе, в исключительных случаях, использование специальных категорий данных для корректирующих целей В литературе такая постановка вопроса описывается как практическая коллизия между потребностью диагностировать предвзятость по группам и ограничениями режима специальных категорий персональных данных: верификация дискриминационных исходов нередко требует доступа к чувствительным атрибутам или их приближенным заменителям, тогда как *GDPR* в типичных сценариях ограничивает сбор и использование таких данных, а согласие часто не отвечает критерию свободного волеизъявления. Одновременно подчеркивается, что наращивание массивов сверх строгой необходимости конфликтует с принципами минимизации и целевого ограничения, повышая риск «ползучего расширения» задач и масштабов обработки (*function creep*); в качестве одного из способов удержания баланса обсуждается использование синтетических данных как инструмента снижения рисков для частной жизни при сохранении задач контроля предвзятости (*bias*) (van Bekkum, 2025, p. 7; van Bekkum & Zuiderveen Borgesius, 2023, pp. 5–6, 7; Veltmeijer & Gerritsen, 2025, pp. 3174–3175).

---

43 Niemiec, E., Davis, P., & Hauglid, M. (2024). *Will the EU AI Act help to eliminate dataset bias in medical AI?* SSRN 4, 46.

44 Ibid. at 26.

45 Ibid. at 62.

46 Ibid. at 26.

47 Ibid. at 57–58, 79.

48 Charter of Fundamental Rights of the European Union, 2012 O.J. (C 326) 400.

С учетом указанной дихотомии разработчик корпоративной ИИ-системы на обучающей стадии сталкивается как минимум с тремя ключевыми задачами:

1) установить, относится ли создаваемый продукт к высокорисковым согласно ст. 6 Акта об ИИ[49] и, следовательно, подпадает ли он под жесткий режим регулирования;
2) определить, содержит ли обучающий набор персональные данные в понимании ст. 4(1) *GDPR* и специальные категории данных в смысле ст. 9 *GDPR*[50];
3) оценить наличие разумно предсказуемого риска для основных прав, что может требовать проведения процедуры *DPIA* и, в отдельных случаях, *FRIA.*

В отличие от абстрактного анализа норм процедура оценки воздействия на защиту данных (*DPIA*) позволяет продемонстрировать регулятору и самим субъектам данных, каким образом оператор учел выявленные риски и превратил их в управляемые параметры проектного цикла.

DPIA должна включать описание операций и целей обработки; правовой анализ и выбор правового основания; идентификацию и оценку рисков; перечень технических и организационных мер защиты; оценку остаточного риска и, при необходимости, консультацию с надзорным органом (ст. 36 *GDPR*)[51].

В качестве иллюстрации рассмотрим применение процедура оценки воздействия на защиту данных (*DPIA*) при обучении модели кредитного скоринга в финансовом секторе. Банк, разрабатывающий алгоритм оценки вероятности дефолта клиента, обрабатывает комплексные массивы персональных данных: кредитные истории, транзакции, сведения из процедур *KYC (Know Your Customer),* иногда дополнительные профили из внешних источников. При этом возникает реальная опасность косвенной дискриминации по косвенным социальным признакам, опосредованно связанным с защищаемыми характеристиками (например, по месту проживания, типу занятости, уровню дохода, модели потребления) и меморизации («запоминания») чувствительной информации моделью. В рамках *DPIA* банк фиксирует цели обработки, источники данных и сценарии их повторного использования. Для принятия решения о кредите правовым основанием служит статья 6(1)(b) *GDPR* (исполнение обязательства из договора), тогда как для последующего улучшения модели — статья 6(1)(f) *GDPR* (законный интерес) в сочетании с усиленной минимизацией и ограничением доступа[52]. Отбор переменных при этом осуществляется не по принципу «чем больше, тем лучше», а по доказанной значимости для качества прогнозирования и влиянию на риск дискриминации.

Следующий блок *DPIA* посвящен идентификации и оценке рисков для прав субъектов данных: непрозрачность алгоритмической логики, возможность косвенной дискриминации уязвимых групп, риск утечки информации через поведение модели. Каждый риск получает оценку вероятности и тяжести последствий, после чего формируется набор компенсирующих мер: регулярный аудит смещений (*bias*); мониторинг дрейфа данных и деградации качества; использование специальных категорий данных исключительно для тестирования справедливости и недопущение их включения в «боевую» модель. Результатом *DPIA* становится формализация остаточного риска; если он признается высоким, оператор обязан обратиться за консультацией к надзорному органу в порядке ст. 36 *GDPR*[53]. Такая консультация не является формальной,

[49] Artificial Intelligence Act, 2024 O.J. (L 2024/1689) 53–54.
[50] General Data Protection Regulation, 2016 O.J. (L 119) 33, 38–39.
[51] Ibid. at 118–119.
[52] Ibid. at 36.
[53] Ibid. at 54–55.

а представляет собой механизм диалога о допустимых границах применения искусственного интеллекта в конкретном секторе.

Особенно показательна взаимосвязь между ст. 10 Акта об ИИ (управление данными)[54] и ст. 25 *GDPR* (защита персональных данных по замыслу и по умолчанию)[55]. Первая предписывает обеспечивать качество, репрезентативность и корректность данных, а также процедуры управления ими; вторая требует встроенной минимизации и защиты уже на этапе проектирования архитектуры обработки. Только наложение этих требований формирует юридически валидный пайплайн обучения: *GDPR* задает горизонтальные пределы допустимой обработки, а Акта об ИИ утверждает вертикальные процедуры доказательства ее безопасного и недискриминационного характера.

Дополнительное значение приобретает международный стандарт *ISO/IEC 42001:2023* — первый сертификационный стандарт для систем управления ИИ, интегрирующий риск-ориентированное мышление, предполагающее принятие решений с учетом вероятности и тяжести последствий (*risk-based thinking*), и механизмы постоянного улучшения[56]. Его практическая ценность заключается в том, что он создает процессную прослойку между нормами и инженерными практиками, подтверждая, что принцип минимизации *GDPR* и процедуры управления данными Акта об ИИ укоренены в системе менеджмента компании, а не применяются разово «под аудит».

Уточнение субъектного состава завершает правовые очертания обучения. Статья 3(3) Акта об ИИ вводит фигуру провайдера (*provider*) — лица, размещающего систему на рынке под своим именем[57], одновременно опираясь на традиционные категории, предусмотренные *GDPR*, — «контролер, оператор персональных данных» (*controller*) и «обработчик» (*processor*) (пп. 7–8 ст. 4 *GDPR*)[58]. На практике одна и та же компания может совмещать обе роли: при разработке выступать оператором по отношению к обучающему набору, а при выводе продукта на рынок становиться провайдером высокорисковой системы. В результате формируется управленческая матрица обязанностей, в которой по вертикали располагаются стадии жизненного цикла (сбор, подготовка, обучение, валидация, эксплуатация и мониторинг), а по горизонтали — кластеры регуляторных норм (обязанности по *GDPR,* требования Акта об ИИ, процедуры согласно стандарту *ISO/IEC 42001:2023*). Такая матрица служит инструментом распределения ответственности, планирования ресурсов комплаенса и включения управления рисками в цикл разработки ИИ-модели.

Принцип минимизации данных как ценностный фильтр и методологическая опора обучения ИИ-моделей

Принцип минимизации данных в ст. 5(1)(c) *GDPR* формулируется через триаду «адекватность — релевантность — ограниченность необходимым»[59]. Адекватность означает соответствие данных поставленной цели, релевантность — наличие прямой связи данных с задачей обработки,

[54] Artificial Intelligence Act, 2024 O.J. (L 2024/1689) 57–58.

[55] General Data Protection Regulation, 2016 O.J. (L 119) 48.

[56] International Organization for Standardization. (2023). *ISO/IEC 42001:2023: Information technology — artificial intelligence — management system*. https://www.iso.org/standard/81230.html

[57] Artificial Intelligence Act, 2024 O.J. (L 2024/1689) 45.

[58] General Data Protection Regulation, 2016 O.J. (L 119) 33.

[59] Ibid. at 35.

а ограниченность необходимым предполагает невозможность достижения цели без конкретного элемента данных. Минимизация в правовом смысле подразумевает то, что обрабатываться должен только тот объем информации, который действительно нужен для конкретной цели[60] (Finck & Biega, 2021, pp. 55–57).

На практике это превращает минимизацию в инструмент управления рисками: каждое решение о включении нового атрибута (например, даты рождения, геолокации, детализированной истории транзакций) должно оцениваться через призму необходимости для достижения заявленной цели обработки и соразмерности вмешательства в права субъекта персональных данных. Суд Европейского союза в деле *Meta Platforms Ireland v. Bundesverband der Verbraucherzentralen & Verbraucherverbände*[61] подчеркнул, что даже при наличии выраженного коммерческого интереса, включая финансирование деятельности за счет персонализированной рекламы, допустимость обработки определяется тестом необходимости и пропорциональности, а также наличием менее ограничивающих альтернатив; при этом в условиях широкого и фактически «почти тотального» мониторинга поведения пользователей интересы и основные права последних могут иметь приоритет над интересом оператора, что усиливает значение минимизации как требования исключать из обработки избыточные и идентифицирующие элементы, не являющиеся строго необходимыми для цели[62].

Дополнительный уровень сложности создает статья 10 Акта об ИИ, которая требует, чтобы обучающие выборки[63] были репрезентативными, полными и, по возможности, свободными от ошибок[64]. Репрезентативность предполагает, что выборка отражает свойства всей целевой популяции (например, база клиентов банка включает разные возрастные группы, регионы, социальные категории). Полнота означает наличие признаков, без которых задача модели не может быть решена (например, отсутствие данных о доходах делает кредитный скоринг заведомо деформированным). Требование «по возможности, свободными от ошибок» связано с очисткой выборки от дубликатов, некорректных, противоречивых и явно устаревших значений.

На уровне формальной логики требования минимизации и репрезентативности выглядят как напряженная пара: чем больше данных собрано, тем проще обеспечить статистическую устойчивость[65], но тем выше регуляторный риск избыточной обработки. Разрешение этого противоречия осуществляется через принцип соразмерности, закрепленный в ст. 52 Хартии ЕС об основных правах: вмешательство в права допускается лишь в той мере, в какой оно необходимо и пропорционально легитимной цели[66]. Для обучения моделей это означает, что использование крупных выборок оправданно только при доказанной необходимости для достижения статистической устойчивости и недискриминационного результата и при отсутствии менее

---

60 European Data Protection Board, *Guidelines 4/2019 on Article 25 Data Protection by Design and by Default* (Version 2.0, adopted October 20, 2020), https://www.edpb.europa.eu/sites/default/files/files/file1/edpb_guidelines_201904_dataprotection_by_design_and_by_default_v2.0_en.pdf

61 Case C-319/20, Meta Platforms Ireland Ltd. v. Bundesverband der Verbraucherzentralen & Verbraucherverbände, ECLI:EU:C:2022:322, ¶ 51–56, 83 (Apr. 28, 2022).
По решению органов государственной власти Российской Федерации, организация *Meta Platforms, Inc.* признана экстремистской, ее деятельность запрещена на территории России.

62 Case C-252/21, Meta Platforms Inc. v. Bundeskartellamt, ECLI:EU:C:2022:704, ¶ 108–109, 115–118, 121 (Sept. 20, 2022).

63 *Training datasets* — обучающие датасеты, массивы данных, на которых строится и корректируется модель.

64 Artificial Intelligence Act, 2024 O.J. (L 2024/1689) 57.

65 *Robustness* — устойчивость модели к вариациям входных данных.

66 Charter of Fundamental Rights of the European Union, 2012 O.J. (C 326) 407.

инвазивной альтернативы (например, работы с агрегированными данными или более грубой детализацией признаков).

Современные исследования машинного обучения подтверждают, что рост объема и детализации данных не является безусловным благом. Так, чрезмерное увеличение числа параметров модели (глубины нейронной сети, количества признаков) приводит к переобучению[67] (Rocks & Mehta, 2022, p. 15). Технически это выражается в снижении числа ошибок на обучающих данных при одновременном росте на тестовом наборе. Для правового анализа важно не само понятие дисперсии и смещения, а вывод: существует точка, после которой увеличение объема и детализации данных ухудшает качество решения и, следовательно, не может считаться необходимым в смысле ст. 5(1)(c) *GDPR*[68].

Один из практических способов совместить требования минимизации и репрезентативности — статистическая подвыборка[69]. Основанная на понятных критериях (устранение повторов, выравнивание дисбаланса групп, исключение очевидных выбросов), она одновременно снижает регуляторный риск и повышает обобщающую способность модели. В юридической плоскости это можно рассматривать как реализацию принципов минимизации и защиты данных по замыслу *(data protection by design)* (ст. 25 *GDPR*) средствами инженерного проектирования[70].

Существенную роль играют и методы обезличивания. Классическим подходом остается *k*-анонимность[71]. Исследования показывают, что разумно сконструированная *k*-анонимность может снижать риск повторной идентификации при минимальном влиянии на качество классификации (Slijepčević et al., 2021). Дополнительно используется *l*-диверсификация[72], что препятствует восстановлению личности по редким комбинациям атрибутов. Эти техники не устраняют полностью юридическую связь с понятием персональных данных, но позволяют уменьшить степень вмешательства и продемонстрировать надзорному органу, что контролер предпринял разумные шаги к минимизации риска.

В праве ЕС минимизация данных в сфере обучения ИИ-моделей трактуется не как разовое сокращение признаков, а как управленческий и юридический процесс, встроенный в деятельность контролера. Статья 5(1)(c) *GDPR* и принцип подотчетности в ст. 5(2) *GDPR* требуют, чтобы контролер мог объяснить и задокументировать, почему каждый элемент данных включен в обработку[73]. На практике это реализуется через поэтапную процедуру.

Первый этап — инвентаризация признаков. Контролер составляет полный перечень атрибутов, планируемых к использованию в обучении, и проводит их классификацию по уровню идентифицируемости. Условно можно выделить три уровня. Уровень *A* включает прямые идентификаторы (имя, паспортные данные, телефон, адрес электронной почты); их сохранение в исходном виде в обучающем наборе допустимо лишь в исключительных случаях. Как правило,

[67] Более известному как *"overfitting"* — ситуация, когда модель идеально подгоняется под обучающий набор, но теряет способность к обобщению в отношении новых случаев.

[68] General Data Protection Regulation, 2016 O.J. (L 119) 35–36.

[69] *Subsampling* — формирование уменьшенного набора данных за счет удаления дубликатов, нерелевантных и явно избыточных наблюдений.

[70] Ibid. at 48.

[71] *K-anonymity* — модель, при которой каждая запись в наборе неотличима как минимум от *k – 1* других. Если, например, *k* = 10, то каждый субъект статистически растворен в группе из не менее 10 лиц с одинаковой комбинацией основных атрибутов (возраст, регион, тип договора).

[72] *L-diversity* — условие, при котором в каждой *k*-анонимной группе присутствует не менее одного различного значения чувствительного признака, например диагноза или уровня дохода.

[73] Ibid. at 35–36.

в отношении них требуется либо удаление, либо псевдонимизация[74] (ст. 4(5) *GDPR*)[75]. Уровень *B* охватывает квази-идентификаторы (возраст, регион, профессию, комбинацию характеристик), которые по отдельности не раскрывают личность, но в совокупности могут вести к идентификации; их использование требует прохождения теста невозможности альтернативы, когда оператор должен документально подтвердить, что без этих данных невозможно обеспечить требуемое качество модели. Уровень *C* составляют менее чувствительные признаки, в отношении которых приоритет имеют укрупнение, агрегирование и обезличивание, если потери качества остаются в пределах статистически допустимого диапазона.

Результаты такой классификации фиксируются в журнале минимизации данных — внутреннем документе контролера, который совмещает управленческую и доказательственную функции. По сути, это специализированный реестр операций обработки в смысле ст. 30 *GDPR*[76], ориентированный не столько на описание целей и категорий данных, сколько на обоснование необходимости и пропорциональности использования каждого признака. Практика национальных регуляторов показывает, что именно подобные журналы становятся первым запрашиваемым документом при проверке проектов на базе ИИ: они демонстрируют, что принцип минимизации не остался декларацией, а был реализован в процессе проектирования[77].

Второй этап — отбор минимально представительного среза данных. Здесь можно говорить о двух правовых фильтрах. Во-первых, фильтр доминирующей информации: сохраняются только признаки, вклад которых в достижение цели подтвержден результатами тестирования и документирован в отчете *DPIA* или аналогичных материалах. Во-вторых, фильтр уникальности: редкие комбинации атрибутов должны быть агрегированы, чтобы каждый субъект оказался статистически скрыт в группе (по крайней мере из нескольких лиц), что снижает риск повторной идентификации, особенно при малых выборках.

Третий этап связан с применением технологий усиления приватности[78]. В качестве наиболее показательного примера в литературе обычно рассматривается дифференциальная приватность как формализованный подход к ограничению утечек о конкретных записях в наборе данных за счет добавления контролируемой случайности в результаты вычислений[79] (Dwork et al., 2006, pp. 270–271).

В техническом смысле требование описывается через устойчивость вывода алгоритма к единичным изменениям исходного набора: результат должен оставаться практически одинаково вероятным как при наличии в данных одной конкретной записи, так и при ее отсутствии. Иначе говоря, для любого возможного результата вероятность его получения на двух «соседних» наборах данных (отличающихся одной записью) может различаться лишь в строго ограниченных пределах: допускается ограниченное расхождение по масштабу вероятностей и только очень малая дополнительная вероятность отклонения за пределы этого ограничения. Эта конструкция направлена на то, чтобы по наблюдаемому выходу алгоритма нельзя было надежно восстановить факт наличия конкретного субъекта в наборе данных или извлечь

[74] *Pseudonymisation* — замена идентификаторов токенами или хешами.

[75] Ibid. at 33.

[76] Ibid. at 50–51.

[77] DLA Piper. (2025). *National data protection authority in Germany.* https://www.dlapiperdataprotection.com/?t=authority&c=DE

[78] *Privacy-enhancing technologies (PETs)* — технологии усиления защиты данных.

[79] *Differential privacy* — математически контролируемое добавление случайных искажений — «шума» — к данным или результатам вычислений, при котором вклад отдельного лица становится статистически неразличимым.

чувствительную информацию о нем. В европейской практике дифференциальная приватность все чаще рассматривается как мера, совместимая с принципом «защита данных по замыслу и по умолчанию» (*data protection by design and by default*) (ст. 25 *GDPR*)[80], при условии документирования выбранного механизма, обоснования отсутствия менее инвазивных альтернатив и оценки влияния добавленной случайности на качество и устойчивость модели при условии документирования выбранного механизма (описание примененной схемы, области применения и того, какие выходы алгоритма защищаются), обоснования отсутствия менее инвазивных альтернатив (сопоставление с вариантами, требующими меньшего объема персональных данных или меньшего вмешательства в обработку) и оценки влияния добавленной случайности на качество и устойчивость модели (в том числе на сохранение сопоставимой точности и долей ошибок для различных групп субъектов и на отсутствие систематически худших результатов для уязвимых или защищаемых групп, что важно для предотвращения дискриминационных исходов).

Особое место в сегодняшней дискуссии занимает федеративное обучение[81]. На интуитивном уровне может показаться, что сохранение данных «на месте» автоматически устраняет риски для персональных данных. Однако с точки зрения *GDPR* ситуация сложнее. В соответствии со ст. 4(1) *GDPR* персональными являются любые сведения, относящиеся к идентифицированному или идентифицируемому лицу[82], а Европейский совет по защите данных в заключении 28/2024 подчеркивал, что даже косвенные цифровые следы могут образовывать персональные данные, если существует разумная вероятность их использования для повторной идентификации[83]. Градиенты и обновления весов при определенных условиях могут быть инвертированы, что позволит восстановить отдельные характеристики исходных записей.

Это означает, что в федеративной архитектуре контролер не может автоматически считать градиенты безличной информацией. При наличии разумно предсказуемого риска реконструкции индивидуальных записей возникает высокая степень угрозы для прав и свобод субъектов, а значит, необходимо проводить *DPIA* по ст. 35 *GDPR*[84]. В рамках такой оценки контролер должен показать, что архитектурное решение дополнено достаточными гарантиями: сокращением экстремальных градиентов, использованием защищенной агрегации[85], криптографических механизмов вроде гомоморфного шифрования[86]. Эти технические меры приобретают юридическое значение, только когда описаны в *DPIA,* закреплены в политиках обработки и отражены в реестре операций обработки.

Юридический смысл такого подхода состоит в том, что минимизация данных превращается из абстрактного принципа в последовательный процесс проектных решений: от отбора признаков и структурирования датасета до выбора архитектуры обучения и применения технологий усиления защиты данных (*PETs*). Для обучения ИИ-моделей это означает, что допустимыми признаются только те конфигурации данных и процедур, по которым контролер способен

[80] General Data Protection Regulation, 2016 O.J. (L 119) 48.

[81] *Federated learning* — архитектура, при которой модель обучается распределенно, на устройствах или серверах участников, а в центральный контур передаются только обновления параметров или градиенты.

[82] Ibid. at 48.

[83] European Data Protection Board. (2024). *Opinion 28/2024 on certain data protection aspects related to the processing of personal data in the context of AI models*. https://www.edpb.europa.eu/system/files/2024-12/edpb_opinion_202428_ai-models_en.pdf

[84] General Data Protection Regulation, 2016 O.J. (L 119) 53–54.

[85] *Secure aggregation* — протоколы, не позволяющие анализировать обновления на уровне отдельных участников.

[86] *Homomorphic encryption* — шифрование, допускающее вычисления с зашифрованными данными.

убедительно показать: каждый элемент необходим, альтернативы с меньшим объемом данных не обеспечивают сопоставимого качества, а остаточный риск для фундаментальных прав снижен до приемлемого уровня и задокументирован.

Принцип соразмерности защитных мер в регулировании обучения ИИ-моделей

Принцип соразмерности защитных мер в праве Европейского союза предполагает, что объем и характер технических и организационных мер должны соответствовать уровню риска для прав и свобод субъектов данных и серьезности возможных последствий. В сфере защиты данных он отражен прежде всего в статье 25 *GDPR,* закрепляющей требование защиты данных по замыслу и по умолчанию (*data protection by design and by default*), и в статье 32 *GDPR,* устанавливающей обязанность обеспечить надлежащий уровень безопасности с учетом состояния техники, стоимости внедрения, характера, объема, контекста и целей обработки[87]. Иначе говоря, оператор обязан не просто применить набор мер, а обосновать, почему именно такой набор является достаточным и пропорциональным рискам.

Статья 10(2) Акта об ИИ развивает эту логику применительно к высокорисковым ИИ-системам: провайдеры обязаны не только защищать данные, но и формализовать управление обучающими наборами (*data governance*), документируя их происхождение, качество, репрезентативность и отсутствие систематических искажений[88]. Тем самым Акт об ИИ смещает фокус с индивидуального риска (логика *GDPR*) на риск алгоритмического воздействия на группы и общество в целом: неправильно обученная модель может приводить к дискриминации, ошибкам в доступе к социальным услугам, кредитованию, здравоохранению (Novelli et al., 2024a, p. 8). Поэтому соразмерность здесь имеет два критерия оценки: с позиции стандарта охраны прав субъекта персональных данных (*DPIA* по ст. 35 *GDPR*)[89] и риска нарушения основных прав (*FRIA*) в духе Хартии ЕС об основных правах[90].

Базовым примером соразмерной меры выступает псевдонимизация согласно ст. 4(5) *GDPR*[91]. Она позволяет отделить идентификаторы (ФИО, паспортные данные, контактную информацию) от остальной части обучающего набора, сохранив возможность контролируемого восстановления личности через отдельный ключ в защищенной системе управления ключами. Однако международный стандарт *ISO/IEC 27559:2022* справедливо указывает, что одной псевдонимизации недостаточно: при объединении разных наборов данных сохраняется риск фоновой деанонимизации (*background re-identification*), когда лицо идентифицируется путем сопоставления псевдонимизированных записей с внешними источниками[92]. Поэтому псевдонимизация рассматривается как необходимая, но не достаточная мера и должна дополняться минимизацией, укрупнением редких сочетаний и контролем доступа.

Для специальных категорий данных (ст. 9 *GDPR*)[93] и высокочувствительных обучающих наборов усиливается роль *дифференциальной приватности*[94]. В литературе данный подход

87 Ibid. at 48, 51–52.

88 Artificial Intelligence Act, 2024 O.J. (L 2024/1689) 57.

89 General Data Protection Regulation, 2016 O.J. (L 119) 53.

90 Charter of Fundamental Rights of the European Union, 2012 O.J. (C 326) 407.

91 General Data Protection Regulation, 2016 O.J. (L 119) 33.

92 International Organization for Standardization. (2023). *ISO/IEC 42001:2023: Information technology — artificial intelligence — management system.* https://www.iso.org/standard/81230.html

93 General Data Protection Regulation, 2016 O.J. (L 119) 38–39.

94 См. подробнее о дифференциальной приватности предыдущий раздел настоящей статьи.

описывается как предпочтительный для случаев, когда одного обезличивания недостаточно для снижения риска выводимости сведений о субъекте (то есть когда по результатам обработки сохраняется вероятность восстановить связь с конкретным лицом либо извлечь о нем чувствительную информацию) (Arasteh et al., 2024, p. 2; Liu et al., 2024, pp. 737–738). С правовой точки зрения настройка используемого механизма и объема вносимой случайности образует вопрос соразмерности: недостаточная степень усиления защиты делает меры фактически формальными (результаты остаются статистически различимыми в зависимости от участия конкретной записи), тогда как чрезмерное усиление может существенно ухудшить качество модели (увеличить ошибку, снизить устойчивость и привести к неравномерному росту ошибок для отдельных групп субъектов, что повышает риск дискриминационных исходов). Для небольших выборок и редких случаев (например, редких заболеваний) это требует проведения отдельной процедуры оценки воздействия на защиту данных (*DPIA*) с фиксацией того, почему выбран именно данный способ усиления защиты и каким образом оценено его влияние на точность, устойчивость и потенциальные дискриминационные эффекты.

Следующий слой соразмерных мер связан с классическими средствами информационной безопасности: контролем доступа, ролевыми моделями и шифрованием. Принцип наименьших привилегий (*least privilege*) предполагает, что разработчик, дата-сайентист или администратор получают доступ только к тем данным, которые им действительно необходимы для выполнения задачи (Menges et al., 2021, p. 2). Статья 32(1) *GDPR* прямо называет среди мер шифрование[95], обеспечивающее защиту от несанкционированного доступа как при компрометации инфраструктуры, так и при перехвате трафика[96]. Для высокорисковых систем добавляются аппаратные технологии, например безопасные анклавы[97]. Они особенно актуальны для медицинских, биометрических и иных чувствительных сценариев, где компрометация обучающего контура приводит к необратимому ущербу (Novelli et al., 2024a, p. 7; Novelli et al., 2024b, pp. 5–6; Basdekis et al., 2023, p. 2)[98].

Отдельный класс архитектурных мер соразмерности связан с сокращением объема передаваемых данных за счет изменения самого процесса обучения. Федеративное и раздельное обучение[99] часто позиционируется как решение проблем конфиденциальности (Haripriya et al., 2025, pp. 10–11; Pham et al., 2023, pp. 988–989). Однако с точки зрения *GDPR* и Акта об ИИ эти архитектуры не снимают обязанность по оценке рисков, а лишь смещают ее в другую плоскость. Как показывают исследования по инверсии градиентов (*gradient inversion*), при недостаточно защищенной агрегации обновлений возможно восстановление исходных признаков. Следовательно, градиенты и обновления весов должны рассматриваться как потенциальные персональные данные, а применение федеративного обучения — сопровождаться *DPIA* и, при высокорисковых сценариях, *FRIA*. Соразмерный набор мер в этом контексте включает обрезку экстремальных градиентов, протоколы защищенной агрегации, исключающей анализ обновлений по отдельным клиентам, а при необходимости — криптографические механизмы, включая гомоморфное шифрование.

---

95 *Encryption* — защита данных при хранении и передаче.

96 Ibid. at 51–52.

97 *Secure enclaves* — изолированные области памяти, внутри которых выполняются вычисления с данными и параметрами модели без доступа со стороны остальной системы.

98 Krause, D. (2024). *Addressing the challenges of auditing and testing for AI bias: A comparative analysis of regulatory frameworks.* SSRN. https://doi.org/10.2139/ssrn.5050631

99 *Split learning* — разделение модели на локальную и серверную части.

В практическом управлении рисками выбор сочетания мер не может быть произвольным. Организации, работающие с высокорисковыми ИИ-системами, выстраивают многоуровневую модель анализа угроз, опираясь на четыре базовых параметра: тип и чувствительность данных, масштаб и критичность проекта, профиль потенциального нарушителя и экономическую оправданность мер. Если обрабатываются специальные категории данных в крупных массивах, риск априори оценивается как высокий, что влечет обязанность проведения *DPIA* (ст. 35 *GDPR*)[100] и зачастую *FRIA*, а также применения нескольких уровней защиты одновременно: от псевдонимизации и строгого разграничения доступа до дифференциальной приватности, защищенной агрегации и аппаратных анклавов. Для обезличенных или низкочувствительных наборов достаточно базового шифрования и ролевой модели доступа при условии, что сохранен риск-ориентированный подход.

При иных подходах эта логика институционализируется через формальные рамки управления рисками. Документы из системы актов *NIST IR (NIST Incident Response),* включая *IR 8286C*, хотя и потеряли официальный статус, демонстрируют полезную методологию: угрозы сводятся в единый реестр, каждой присваивается уровень риска, а меры защиты привязываются одновременно к бизнес-целям и нормативным требованиям[101]. В европейском контексте аналогичную функцию начинает выполнять *ISO/IEC 42001:2023* — стандарт системы управления ИИ, встраивающий риск-ориентированное мышление (*risk-based thinking*) и непрерывное улучшение в процессы разработки и эксплуатации ИИ[102]. В результате принцип соразмерности перестает быть абстрактным и превращается в управленческую практику: риск идентифицируется, оценивается и «закрывается» документированным набором мер.

Для упорядочения применения защитных средств организации, работающие с обучением моделей, как правило, формируют два слоя документации. Первый — матрица рисков, в которой категориям риска (низкий, средний, высокий) соответствуют обязательные комбинации технических и организационных мер: от псевдонимизации и базового шифрования до дифференциальной приватности, защищенной агрегации и использования аппаратных анклавов. Второй — специализированная процедура оценки рисков обучения моделей (*training DPIA*), где устанавливаются критерии запуска оценки, пороговые значения для высокого риска, роли участников и взаимосвязь между *DPIA, FRIA* и требованиями Акта об ИИ к управлению данными.

В совокупности такие меры формируют согласованный контур защит, в котором псевдонимизация и минимизация выступают базовым уровнем, дифференциальная приватность и архитектурные решения (федеративное / раздельное обучение) — усилением для чувствительных сценариев, а криптографические и аппаратные механизмы — уровнем для критически важных систем. Принцип соразмерности в этой модели означает не экономию на защите, а обязательность выбора и документирования именно тех мер, которые адекватны характеру данных, целям обучения и тяжести возможных последствий для фундаментальных прав.

---

100 General Data Protection Regulation, 2016 O.J. (L 119) 53.

101 Quinn, S., Ivy, N., Barrett, M., Witte, G., & Gardner, R. K. (2022). *Staging cybersecurity risks for enterprise risk management and governance oversight* (NIST Interagency Report No. NIST IR 8286C-upd1). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.IR.8286C-upd1

102 International Organization for Standardization. (2023). *ISO/IEC 42001:2023: Information technology — artificial intelligence — management system*. https://www.iso.org/standard/81230.html

Комплаенс по жизненному циклу ИИ-модели

Жизненный цикл крупной ИИ-модели *(large-scale AI model),* т. е. программной системы, использующей методы машинного обучения и имеющей параметрический масштаб не ниже миллиардного порядка, включает несколько ключевых этапов, каждый из которых имеет самостоятельное правовое значение.

Прежде всего определяются цель и сфера применения модели. По сути, речь идет о формулировании задачи, ради которой создается или адаптируется ИИ-система: к примеру, анализ медицинских изображений, обработка естественного языка, кредитный скоринг, прогнозирование финансовых рисков и т. п. В юридической плоскости это соответствует установлению цели обработки персональных данных в смысле ст. 5(1)(b) *GDPR*[103]*:* диагностические, образовательные, управленческие, сервисные или иные задачи, подлежащие последующему документированию в политике обработки и оценке воздействия (с позиций *DPIA, FRIA* и Акта об ИИ).

На следующем этапе разработчик выбирает архитектурный подход: использовать уже существующую модель (дообучение имеющейся архитектуры нейронной сети или иной модели) либо проектировать и обучать модель с нуля[104]. От этого решения зависят состав данных, объем вычислений и набор юридических рисков: при дообучении неизбежно встает вопрос о повторной обработке ранее собранных персональных данных, при обучении с нуля — о правомерности изначального сбора и масштабах выборки.

Далее следует этап сбора и подготовки обучающих данных[105]. На этой стадии формируются обучающие выборки данных, которые будут использованы для настройки параметров модели. На практике это означает агрегирование объемных массивов данных из открытых источников, публичных реестров, ведомственных систем или коммерчески приобретенных баз. Подготовка включает очистку[106], аннотацию[107] и структурирование данных в формат, пригодный для обучения (Wang et al., 2021, pp. 776–778).

После этого наступает собственно этап обучения модели (*training*). В типичном сценарии он включает предобучение[108] и последующую тонкую настройку (*fine-tuning*). В ходе обучения применяются методы оптимизации[109] и механизмы, предполагающие участие человека, позволяющие достигать требуемой точности и управляемости поведения модели.

Следующий этап — валидация и оценка. Под валидацией понимается проверка модели на отдельной части данных, которая не использовалась при обучении, с целью убедиться в корректности и устойчивости работы. Этап оценки предполагает установление метрик точности и ошибок, а также иных критериев качества, таких как полнота, устойчивость к «шуму», *вариации входных данных* и др.[110] Одновременно исследуются смещения, включая потенциальную дискриминацию по полу, возрасту, региону или иным социально значимым признакам.

---

103 General Data Protection Regulation, 2016 O.J. (L 119) 35.

104 Так называемое создание модели *"from scratch"* — полное создание и обучение без базового прототипа.

105 То есть *"dataset creation and preprocessing"* — формирование и предварительная обработка датасета.

106 *Data cleaning* — удаление дубликатов, явных ошибок и аномалий.

107 *Annotation* — разметка данных (аннотирование): присвоение объектам в наборе данных меток (классов, тегов) для обучения и оценки модели, выполняемое вручную либо с использованием полуавтоматических средств.

108 *Pre-training* — начальная настройка модели на очень больших и разнообразных наборах данных для формирования общих способностей.

109 *Optimization methods* — математические алгоритмы минимизации функции ошибки модели.

110 См. об этом в предыдущих разделах статьи.

Завершающий технический этап — развертывание /внедрение модели (*deployment*) в эксплуатационную среду. С этого момента модель начинает использоваться в реальных бизнес- или управленческих процессах, и на разработчика/оператора ложится обязанность по непрерывному мониторингу и периодической переоценке[111] ее качества, рисков и соответствия нормативным требованиям (Wang et al., 2021, p. 781; De Silva & Alahakoon, 2022, pp. 3–4).

Формулируя цель обучения ИИ-модели, необходимо соблюдать принцип ее ограничения (*purpose limitation*). Согласно ст. 5(1)(b) *GDPR* персональные данные должны «собираться для конкретных, явных и законных целей и не подлежат дальнейшей обработке иным образом, несовместимым с этими целями»[112]. Это означает, что цель обучения должна быть заранее определена, описана в документации и связана с конкретными результатами модели; изменение цели впоследствии требует либо нового согласия субъектов персональных данных, либо иного правового основания.

Акт об ИИ дополняет эту конструкцию понятием «предназначенная цель»[113], т. е. учетом условий, контекста и предсказуемых сценариев применения. Для высокорисковых ИИ-систем законодатель прямо требует описывать и принимать к сведению все известные и разумно предсказуемые сценарии использования и злоупотребления. В соответствии со ст. 9 Акта об ИИ провайдер обязан включить предназначенную цель в систему управления рисками (*risk management system*), а любое ее изменение влечет повторную оценку рисков и при необходимости корректировку архитектуры и процессов обучения[114].

Нормы *GDPR* и Акта об ИИ в этой части согласуются: обе системы требуют прозрачности и точности формулирования целей, их документального закрепления и пересмотра при изменении сценариев использования данных.

Сбор данных для обучения должен опираться на одно из оснований, предусмотренных статьей 6(1) *GDPR:* согласие субъекта (*consent*), исполнение договора (*performance of a contract*), выполнение юридической обязанности (*legal obligation*), защита жизненно важных интересов (*vital interests*), выполнение задачи в общественных интересах или при осуществлении публичных полномочий (*public task*), а также законный интерес оператора (*legitimate interest*), если он не перевешивается правами и свободами субъекта[115]. Например, если обучение модели связано с задачами устойчивого развития или общественного здравоохранения, это может квалифицироваться как выполнение задачи в общественных интересах по ст. 6(1)(e) *GDPR*[116].

Если данные были ранее собраны для иной цели, применяется статья 6(4) *GDPR,* предписывающая проверку совместимости целей (*compatibility test*): анализируются связь между первоначальной и новой целями, контекст сбора, характер данных (обычные или специальные категории), возможные последствия для субъектов и предусмотренные меры защиты[117]. Пункт 67 преамбулы к Акту об ИИ подчеркивает необходимость прозрачности относительно исходной цели сбора и последующих трансформаций[118].

---

111 *Re-evaluation* — повторная оценка по установленным метрикам.

112 Ibid.

113 *Intended purpose* — запроектированная область и условия использования ИИ–системы.

114 Artificial Intelligence Act, 2024 O.J. (L 2024/1689) 56–57.

115 General Data Protection Regulation, 2016 O.J. (L 119) 36.

116 Ibid.

117 Ibid. at 37.

118 Artificial Intelligence Act, 2024 O.J. (L 2024/1689) 19–20.

Принцип минимизации данных (ст. 5(1)(c) *GDPR*) требует, чтобы собирались только те персональные данные, которые действительно необходимы для достижения заявленных целей[119]. В то же время статья 10 Акта об ИИ устанавливает требование репрезентативности обучающих данных, т. е. их сбалансированности и отражения ключевых характеристик целевой популяции (например, половозрастной структуры, этнического состава, географического распределения)[120]. На первый взгляд это приводит к противоречию: минимизация предполагает сокращение объема данных, а репрезентативность — расширение набора для отражения разнообразия. Однако в современной доктрине эти категории разводятся: минимизация понимается как отбор только релевантных признаков и исключение избыточной информации, тогда как репрезентативность — как баланс и непредвзятость внутри уже минимизированного множества данных. В практическом измерении это означает, в частности, возможность использовать анонимизированные или синтетические данные для увеличения разнообразия без нарушения принципа минимизации (van Bekkum, 2025, p. 7).

На этапе сбора данных статья 13 *GDPR* обязывает информировать субъектов персональных данных о ключевых параметрах обработки: целях использования данных, правовых основаниях, сроках хранения, получателях, а также о правах субъекта (право на доступ, исправление, удаление, ограничение обработки и др.)[121]. Это информационная основа прозрачности, без которой любая последующая работа с данными для обучения ИИ-модели будет юридически уязвимой.

Акт об ИИ дополняет эти требования концепцией практик управления данными[122]. В статье 10(2) Акта об ИИ провайдеру высокорисковой ИИ-системы вменяется в обязанность документировать происхождение данных[123], исходную цель их сбора, примененные методы очистки (*data cleaning*) и разметки (*data labelling / annotation*). В технической документации (приложение III к Акту об ИИ[124]) должны быть отражены используемые датасеты, процедуры подготовки и проверки, что позволяет контролирующим органам оценить как качество, так и правомерность подготовки обучающего массива.

Требование ст. 5(1)(c) *GDPR* о минимизации данных в контексте обучения ИИ-моделей означает обязанность оператора (*controller*) обосновать выбор каждого вида данных[125]. Оператор должен продемонстрировать, почему именно такой набор признаков необходим для достижения заявленной цели и что объем данных соотносится с критерием необходимости, а не с техническим удобством по принципу «чем больше, тем лучше».

С другой стороны, статья 10(3) Акта об ИИ вводит критерий репрезентативности и достоверности обучающих и валидационных данных (*representativeness and accuracy*): используемые наборы должны адекватно отражать целевую популяцию, исключать систематические ошибки и обеспечивать статистическую пригодность модели. Здесь проявляется различие в логике требований: репрезентативность предполагает широкий охват и разнообразие данных, а минимизация — их сокращение. Правовой подход позволяет согласовать эти требования: оператор фиксирует в документации (отчетах по результатам *DPIA;* технических отчетах по Акту об ИИ),

119 General Data Protection Regulation, 2016 O.J. (L 119) 35.

120 Artificial Intelligence Act, 2024 O.J. (L 2024/1689) 57–58.

121 General Data Protection Regulation, 2016 O.J. (L 119) 40–41.

122 *Data governance practices* — практики управления качеством, происхождением и безопасностью данных.

123 То есть *"data provenance"* — источник и способ получения данных. См.: Artificial Intelligence Act, 2024 O.J. (L 2024/1689) 57.

124 Artificial Intelligence Act, 2024, annex III, O.J. (L 2024/1689) 127–129.

125 General Data Protection Regulation, 2016 O.J. (L 119) 35–36.

что объем данных минимален с точки зрения прав субъекта, но достаточен для статистической устойчивости и недискриминации. Баланс достигается, в частности, через исключение явно нерелевантных признаков; укрупнение и агрегирование чувствительных данных; использование анонимизации и синтетической генерации для расширения выборки без увеличения объема персональных данных (Laato et al., 2022, p. 4).

Сам процесс обучения моделей в праве ЕС рассматривается как непрерывная обработка персональных данных (*continuous processing*). Это означает, что на обучение целиком распространяются базовые принципы и права согласно *GDPR*: законность обработки *(lawfulness)* (ст. 6)[126]; целевое ограничение и минимизация *(purpose limitation u data minimisation)* (ст. 5)[127], обеспечение безопасности *(security of processing)* (ст. 32)[128], а также специальные правила профилирования и автоматизированного принятия решений *(profiling and automated decision-making)* (ст. 22)[129]. Если итоговая модель используется для принятия решений, затрагивающих права субъектов (например, отказ в кредите, приеме на работу, доступе к услуге), возникает дополнительная обязанность информировать о логике и последствиях обработки (ст. 13 и 22 *GDPR*)[130] и обеспечивать возможность человеческого вмешательства (*human intervention*).

Статья 9 Акта об ИИ дополняет эту конструкцию требованием к системному управлению рисками[131]. Оно должно носить постоянный, цикличный и документированный характер и охватывать весь жизненный цикл ИИ-системы: от проектирования и обучения до эксплуатации и пострыночного мониторинга. Особый акцент делается на рисках дискриминации и смещения данных. Уже на стадии проектирования и обучения провайдер обязан выявлять такие риски и принимать меры по их снижению[132].

С этой целью статья 11 Акта об ИИ[133] и приложение III к нему[134] закрепляют обязанность вести подробную техническую документацию, где фиксируются архитектура системы, используемые датасеты, критерии отбора данных, процессы очистки и проверки, протоколы тестирования, методы предотвращения предвзятости и результаты оценок. Такая документация частично пересекается по функциям с реестрами операций обработки по ст. 30 *GDPR*[135], что позволяет выстраивать единую систему комплаенса: одна часть записей ориентирована на надзор по линии защиты данных, другая — на контроль соответствия Акту об ИИ.

Принцип защиты данных по замыслу и по умолчанию (ст. 25 *GDPR*) требует от провайдера искусственного интеллекта внедрять технические и организационные меры еще на стадии обучения: шифрование (защиту данных криптографическими методами), псевдонимизацию (замену идентификаторов токенами), дифференциальную приватность (контролируемое

126 Ibid. at 36–37.

127 Ibid. at 35–36.

128 Ibid. at 51–52.

129 Ibid. at 46.

130 Ibid. at 40–41.

131 Artificial Intelligence Act, 2024, annex III, O.J. (L 2024/1689) 56–57.

132 Slattery, P., Saeri, A. K., Grundy, E. A. C., Graham, J., Noetel, M., Uuk, R., Dao, J., Pour, S., Casper, S., & Thompson, N. (2024). *The AI Risk Repository: A Comprehensive Meta-Review, Database, and Taxonomy of Risks from Artificial Intelligence.* https://doi.org/10.48550/arXiv.2408.12622

133 Artificial Intelligence Act, 2024, O.J. (L 2024/1689) 58.

134 Ibid. at 127–129.

135 General Data Protection Regulation, 2016 O.J. (L 119) 50–51.

внесение «шума», исключающего выделение человека в наборе данных), а также процедуры «машинного разучивания» (целенаправленное исключение определенных данных из обученной модели)[136]. Эти меры должны быть соразмерны риску и увязаны с требованиями статьи 32 *GDPR*, предписывающей учитывать состояние техники, стоимость внедрения, характер, объем и контекст обработки[137].

Акт об ИИ не вступает с этой логикой в противоречие, а усиливает ее: статья 10(2)(f), (g) прямо указывает на необходимость проверки обучающих данных на отсутствие незаконной предвзятости и несоответствующих смещений[138], а также на согласование таких проверок с принципами и пунктами преамбулы *GDPR*, включая п. 67 о прозрачности и поддержании актуальности данных[139].

Особое значение имеет статья 10(5) Акта об ИИ, допускающая использование специальных категорий данных (чувствительные данные о здоровье, биометрии, убеждениях и т. п.) для устранения смещений в обучении[140]. Формально это вступает в противоречие со статьей 9 *GDPR*, устанавливающей общий запрет на обработку таких данных[141]. Однако Акт об ИИ вводит конкретные условия в смысле статьи 10(5)(a), (b), (e): отсутствие менее инвазивной альтернативы, применение усиленных мер псевдонимизации и других методов защиты, удаление данных сразу после выполнения корректирующей задачи. С правовой точки зрения данную норму следует рассматривать как узкое исключение, которое допустимо только при одновременном выполнении условий ст. 9(2), (4) *GDPR*[142] и документальном обосновании в *DPIA* и технической документации. Это означает, что привлечение чувствительных признаков допустимо исключительно в режиме минимально необходимого вмешательства и под постоянным контролем регулятора и внутренних органов комплаенса.

Наконец, доказательства соблюдения указанных норм формируются в единой системе документации. *GDPR* требует ведения реестров операций обработки (ст. 30)[143], регистрации инцидентов и уведомлений (ст. 33–34)[144], *DPIA* и сопутствующих отчетов. Акт об ИИ предписывает хранить протоколы тестов, логи качества, отчеты об оценке рисков и техническое досье[145]. В совокупности эти документы составляют корпус доказательств законности и прозрачности обучения. На практике оправдан подход, при котором техническая документация по Акту об ИИ интегрируется с документами по *GDPR* (политиками обработки, *DPIA*, реестрами операций), образуя сквозную систему отчетности и повышая доверие к ИИ-системе со стороны регуляторов, партнеров и пользователей.

При валидации (тестировании) модели *GDPR* продолжает действовать в полном объеме. Если валидационные и тестовые наборы содержат персональные данные, их обработка должна опираться на правовое основание (согласие, договор, законный интерес и т. д.) и соответствовать принципам минимизации и целевого ограничения. На этой стадии особенно важна

136 General Data Protection Regulation, 2016 O.J. (L 119) 48.
137 Ibid. at 51–52.
138 Artificial Intelligence Act, 2024, O.J. (L 2024/1689) 57.
139 General Data Protection Regulation, 2016 O.J. (L 119) 13.
140 Artificial Intelligence Act, 2024, O.J. (L 2024/1689) 58.
141 General Data Protection Regulation, 2016 O.J. (L 119) 38–39.
142 Ibid.
143 Ibid. at 50–51.
144 Ibid. at 52–53.
145 Artificial Intelligence Act, 2024, annex III, O.J. (L 2024/1689) 127–129.

оценка рисков дискриминации: в соответствии со ст. 13(2)(f) *GDPR* контролер обязан заранее информировать субъектов о возможных последствиях автоматизированного прогнозирования и рисках для их прав и свобод[146].

Статья 10(2), (3) Акта об ИИ предписывает, чтобы валидационные наборы обладали теми же качественными характеристиками, что и обучающие: релевантностью (связь с задачей), репрезентативностью и способностью обеспечивать проверку точности и устойчивости модели[147]. Это означает, что тестирование должно быть спроектировано, чтобы выявлять смещения и ошибки, а результаты должны фиксироваться в отчетах и логах для целей надзора, аудита и подтверждения соответствия как Акту об ИИ, так и *GDPR*.

На этапе внедрения ИИ-модели в эксплуатацию юридические обязанности оператора и провайдера существенно ужесточаются. Если модель используется для принятия решений, затрагивающих права и законные интересы физических лиц, применяется статья 22 *GDPR* об автоматизированном принятии решений, предусматривающая право субъекта на вмешательство (пересмотр решения человеком), оспаривание решения и получение объяснений относительно логики обработки[148]. Эти требования корреспондируют со статьей 14 Акта об ИИ, закрепляющей принцип человеческого надзора (*human oversight*)[149], а также со статьей 13 *GDPR*, устанавливающей обязанность информировать о целях обработки, логике принятия решений и возможных последствиях для субъектов данных[150].

В этом контексте особое значение приобретает концепция так называемого пакета прозрачности[151]. Под ним в практике применения Акта об ИИ понимается передача по всей цепочке внедрения и эксплуатации исчерпывающего набора материалов: руководств пользователя, технического описания, отчетов о тестировании, ограничений применимости, условий корректного использования и сценариев, при которых система может создавать повышенные риски. Пункт 27 преамбулы к Акту об ИИ подчеркивает, что все лица, участвующие в проектировании, интеграции и эксплуатации системы, должны быть осведомлены о ее характеристиках, возможностях и рисках[152]. В свою очередь *GDPR* в ст. 12–14 дополняет это требованием прозрачности в отношении обработки персональных данных и доступности информации для субъектов, включая сведения об их правах, основаниях обработки и механизмах их реализации[153].

Мониторинг после внедрения (*post-market monitoring*) закреплен в обоих нормативных актах. Статьи 33–34 *GDPR* обязывают контролера сообщать о нарушениях безопасности персональных данных в надзорный орган и при необходимости информировать субъектов данных[154], а также обеспечивать реализацию права на стирание *(right to erasure)* (ст. 17 *GDPR*)[155]. Статьи 61–62 Акта об ИИ, в свою очередь, требуют от провайдера регистрировать серьезные инциденты, поддерживать систему пострыночного мониторинга и фиксировать результаты

146 General Data Protection Regulation, 2016 O.J. (L 119) 41.
147 Artificial Intelligence Act, 2024, O.J. (L 2024/1689) 57.
148 General Data Protection Regulation, 2016 O.J. (L 119) 46.
149 Artificial Intelligence Act, 2024, O.J. (L 2024/1689) 60–61.
150 General Data Protection Regulation, 2016 O.J. (L 119) 40–41.
151 *Transparency pack* — совокупность документов и сведений о системе.
152 Artificial Intelligence Act, 2024, O.J. (L 2024/1689) 8.
153 General Data Protection Regulation, 2016 O.J. (L 119) 39–41.
154 Ibid. at 52–53.
155 Ibid. at 43–44.

в технической документации[156]. Формируется правовая симметрия: обе системы регулирования исходят из обязанности постоянного контроля, регулярной корректировки и документированного подтверждения соответствия.

Отдельный проблемный блок связан с принципом целевого ограничения (*purpose limitation*). В соответствии со ст. 5(1)(b) *GDPR* персональные данные могут использоваться только для конкретных, заранее определенных и законных целей и не подлежат дальнейшей обработке способом, несовместимым с этими целями[157]. Логика коммерческого использования ИИ-моделей, напротив, предполагает расширение сфер применения: модель, изначально обученная для одной задачи (например, внутреннего риск-анализа), впоследствии адаптируется для иных функций или встраивается в новые бизнес-процессы. Такое расширение цели[158] очевидно противоречит конструкции целевого ограничения.

Для частичного разрешения этой коллизии европейское право предусматривает специальный режим исследовательской дерогации (*research derogation*). Статья 89 *GDPR* допускает использование персональных данных в научных и исследовательских целях при условии дополнительных гарантий: псевдонимизации, ограничения доступа, строгого документирования исследовательских целей и процедур[159]. Пункт 159 преамбулы к *GDPR* прямо относит обучение и тестирование алгоритмов к сфере научного и технологического развития, при этом подчеркивая, что такой режим допустим, пока результаты не используются в отношении конкретных субъектов в контексте оказания услуг или принятия решений[160]. На этой стадии первична именно исследовательская цель, а не коммерческая эксплуатация.

Статья 2 Акта об ИИ ограничивает сферу действия регламента системами, которые размещаются на рынке или вводятся в эксплуатацию (*placing on the market / putting into service*)[161]. Соответственно, сугубо исследовательский этап разработки и внутреннего тестирования ИИ-моделей, когда система не предлагается на рынке и не используется для принятия решений в отношении внешних пользователей, напрямую не подпадает под действие Акта об ИИ и регулируется исключительно *GDPR*. На этом этапе применим режим исследовательской дерогации, несколько смягчающий требования при условии строгих дополнительных гарантий (ограничение доступа, псевдонимизация, минимизация и т.д.).

Ситуация принципиально меняется в момент вывода модели на рынок[162] или ее ввода в эксплуатацию во внешнем контуре. С этого момента для разработчика/провайдера возникает комплекс новых обязательств.

Во-первых, необходимо пересмотреть правовое основание обработки данных в соответствии со ст. 6 *GDPR*[163]. Если на исследовательском этапе обработка опиралась на исключение для научных целей *(research basis)* (в сочетании со ст. 89 *GDPR*)[164], то при коммерческом использовании требуется иное основание: согласие субъекта *(consent)* (ст. 6(1)(a)), необходимость

---

156 Artificial Intelligence Act, 2024, O.J. (L 2024/1689) 94–95. См. также соответствующие приложения к Акту: Ibid., annexes IV, V, O.J. (L 2024/1689) 130–132.

157 General Data Protection Regulation, 2016 O.J. (L 119) 35.

158 *Purpose creep* — постепенный выход за исходные цели обработки.

159 Ibid. at 84–85.

160 Ibid. at 30.

161 Ibid. at 45–46.

162 Имеется ввиду первое предоставление системы на рынке ЕС (*"placing on the market"*).

163 *Intended purpose* — запроектированная область и условия использования ИИ–системы.

164 Ibid. at 36–37.

обработки для исполнения договора с субъектом данных (или принятия мер до заключения договора по его запросу) *(performance of a contract)* (ст. 6(1)(b))[165], выполнение юридической обязанности, публичный интерес или законный интерес оператора *(legitimate interests)* (ст. 6(1)(f)) при выполнении теста баланса интересов[166].

Во-вторых, в силу ст. 27 Акта об ИИ и приложения IV к нему разработчик/провайдер обязан обновить описание целей эксплуатации, областей применения и ограничений системы и зафиксировать их в нормативной и технической документации[167]. Эти сведения становятся частью паспорта ИИ-системы, ее технической документации, которая должна включать описание предназначенной цели, целевой среды, категорий пользователей, ограничений и ключевых рисков.

В-третьих, обязательной становится повторная оценка воздействия на защиту данных (обновленная *DPIA* по ст. 35 *GDPR*)[168]. Переход от исследовательского режима к коммерческой эксплуатации изменяет характер обработки, масштаб воздействия и зачастую категории субъектов. Контролер обязан пересмотреть перечень рисков, обновить перечень мер защиты и зафиксировать измененный уровень остаточного риска.

В-четвертых, при наличии серьезных рисков для фундаментальных прав и свобод (по Хартии ЕС об основных правах[169]) возникает необходимость проведения отдельной оценки воздействия на основные права *(FRIA).* Она проводится как дополнение к *DPIA* применительно к высокорисковым ИИ-системам, подпадающим под приложение III к *AI Act (*Mantelero, 2024, p. 17; Malgieri & Santos, 2025, p. 69).

В результате формируется двухступенчатая модель регулирования обучения и использования ИИ-моделей. На стадии разработки и внутреннего тестирования действует исследовательский режим согласно *GDPR* (п. 159 преамбулы и ст. 89)[170], допускающий дерогацию при условии дополнительных гарантий и отсутствия внешнего воздействия на субъектов. С момента вывода продукта на рынок или ввода в эксплуатацию активируется полный комплекс обязательств одновременно по *GDPR* и Акту об ИИ. Фактический момент вывода системы на рынок становится юридической точкой отсчета смены режима: именно с этого времени разработчик обязан пересмотреть правовые основания обработки, обновить техническую и правовую документацию, встроить систему пострыночного мониторинга, усилить меры защиты данных и прав субъектов и при необходимости провести как *DPIA,* так и *FRIA.*

Заключение

Проведенный анализ позволяет утверждать, что этап обучения ИИ-моделей выступает не только технологически, но и юридически значимой фазой жизненного цикла ИИ-систем. Именно на этой стадии концентрируются ключевые обязанности по *GDPR* и Акту об ИИ, связанные с выбором цели и правового основания обработки, управлением качеством и репрезентативностью данных, минимизацией и оценкой рисков для фундаментальных прав. Специфика обучающего этапа проявляется в сочетании масштабной обработки данных, статистического обобщения и высоких рисков профилирования, дискриминации и снижения прозрачности.

[165] Ibid. at 36.

[166] Ibid. at 84–85.

[167] Artificial Intelligence Act, 2024, O.J. (L 2024/1689) 69–70, 130–131.

[168] General Data Protection Regulation, 2016 O.J. (L 119) 53–54.

[169] Charter of Fundamental Rights of the European Union, 2012 O.J. (C 326) 391.

[170] General Data Protection Regulation, 2016 O.J. (L 119) 30, 84–85.

Одновременно отсутствие унифицированных методических разъяснений на уровне институтов ЕС (Европейской комиссии, Европейского управления по искусственному интеллекту, надзорных органов по защите данных) обусловливает правоприменительную неопределенность и расхождение в национальных подходах.

С учетом выявленных противоречий и разнонаправленных регуляторных векторов предлагается прикладная модель поведения разработчиков и операторов, построенная по логике «*R&D* → *Market*» («от исследовательской стадии к стадии рыночного внедрения»). Суть модели состоит в том, что обучение рассматривается как самостоятельный юридически значимый процесс в рамках общего жизненного цикла ИИ-системы, для которого устанавливается порог перехода от исследовательской дерогации по ст. 89 *GDPR* к полному режиму Акта об ИИ. Такая конструкция не претендует на единственно возможное решение: в доктрине обсуждаются и более жесткие подходы, исходящие из необходимости изначально применять к обучающим процедурам весь комплекс требований Акта об ИИ, и, напротив, более мягкие позиции, фактически сводящие регулирование обучения к общим нормам о научной обработке данных. Предлагаемая модель занимает промежуточную позицию, увязывая момент вывода системы на рынок с усилением регуляторного режима и формализацией обязанностей по обоим актам.

В рамках данной модели разработчик (или связка «разработчик — провайдер») выстраивает документационную и организационную структуру, включающую определение и фиксацию цели обработки и правового основания (ст. 5, 6 *GDPR*); правовую и техническую экспертизу обучающих выборок (*data governance* в смысле ст. 10 Акта об ИИ); предпочтительное использование анонимизированных, агрегированных и синтетических данных там, где это допустимо без ущерба для статистического качества; проведение оценки воздействия на защиту данных (*DPIA*) в случаях высокого риска (ст. 35 *GDPR*); ведение внутренней отчетности и журналов операций; соблюдение требований прослеживаемости, устойчивости и подотчетности, закрепленных в Акте об ИИ, включая техническую документацию и протоколы тестирования. С практической точки зрения эта модель превращает набор разрозненных обязанностей в последовательную дорожную карту для корпоративного комплаенса, но одновременно повышает нагрузку на субъекты, лишенные зрелых процедур управления рисками, что уже сегодня является предметом критики со стороны представителей малого и среднего бизнеса.

Значение модели «*R&D* → *Market*» состоит в том, что она предлагает операционализированную схему согласованного применения двух параллельных режимов: горизонтальных требований *GDPR* к обработке персональных данных и вертикальных требований Акта об ИИ к высокорисковым ИИ-системам. При надлежащей имплементации такая схема позволяет не только формально соблюдать ст. 5, 6, 9, 25, 30, 35 *GDPR* и ст. 10, 11, 69 Акта об ИИ, но и выстраивать доказуемую линию поведения на случай регуляторного запроса или судебного спора. Вместе с тем остается открытым вопрос о стандартизации подобных моделей: пока речь идет о доктринальной конструкции, а не о закрепленном регулятором эталоне, от компаний требуется самостоятельная адаптация предложенного подхода к отраслевым и национальным особенностям.

Отдельным элементом предлагаемой модели выступает тест невозможности альтернативы при работе со специальными категориями данных, используемыми для устранения смещений в обучении (ст. 10(5) Акта об ИИ в сочетании со ст. 9 *GDPR*). Содержание теста сводится к необходимости документально обосновать, что привлечение чувствительных признаков (например, этнической принадлежности или состояния здоровья) является единственно необходимым и пропорциональным способом достижения заявленной исследовательской или эксплуатационной

цели и что сопоставимого уровня недискриминации или точности нельзя достичь путем использования синтетических, анонимизированных или менее инвазивных данных. Такой подход отражает более рестриктивное прочтение ст. 10(5) Акта об ИИ, ограничивающее сферу применения исключения, и тем самым снижает риск признания обработки несоразмерной при проверке надзорным органом или судом.

В совокупности результаты исследования позволяют квалифицировать обучение ИИ-модели как юридически значимый этап, от качества правового сопровождения которого зависит соответствие всей ИИ-системы требованиям *GDPR* и Акта об ИИ. Вместе с тем предложенная модель носит характер концептуального предложения и должна рассматриваться как основа для дальнейшей доктринальной дискуссии и последующей регуляторной конкретизации, прежде всего через разъяснения Европейской комиссии, Европейского управления по искусственному интеллекту и Европейского совета по защите данных.

Список литературы / References

1. Arasteh, S. T., Ziller, A., Kuhl, C., Makowski, M. R., Nebelung, S., Braren, R. F., Rueckert, D., Truhn, D., & Kaissis, G. (2024). Preserving fairness and diagnostic accuracy in private large-scale AI models for medical imaging. *Communications Medicine*, (4), Article 46 (2024). https://doi.org/10.1038/s43856-024-00462-6
2. Basdekis, I., Kloukinas, C., Agostinho, C., Vezakis, I., Pimenta, A., Gallo, L., & Spanoudakis, G. (2023). Pseudonymisation in the context of GDPR-compliant medical research. In 2023 *19th International Conference on the Design of Reliable Communication Networks (DRCN)* (pp. 1–6). IEEE. https://doi.org/10.1109/DRCN57075.2023.10108370
3. De Hert, P., & Hajduk, P. (2024). EU cross-regime enforcement, redundancy and interdependence: Addressing overlap of enforcement structures in the digital sphere after Meta. *Technology and Regulation*, *2024*, 291–308. https://doi.org/10.71265/fydwsg59
4. De Silva, D., & Alahakoon, D. (2022). An artificial intelligence life cycle: From conception to production. *Patterns*, *3*(6), Article 100489. https://doi.org/10.1016/j.patter.2022.100489
5. Dwork, C., McSherry, F., Nissim, K., & Smith, A. (2006). Calibrating noise to sensitivity in private data analysis. In S. Halevi & T. Rabin (Eds.), *Lecture notes in computer science, Vol. 3876. Theory of cryptography* (pp. 265–284). Springer. https://doi.org/10.1007/11681878_14
6. Finck, M., & Biega, A. (2021). Reviving purpose limitation and data minimisation in data-driven systems. *Technology and Regulation, 2021,* 44–61. https://doi.org/10.26116/techreg.2021.004
7. Hadwick, D. (2024). Slipping through the cracks, the carve-outs for AI tax enforcement systems in the EU AI Act. *European Papers: A Journal on Law and Integration*, *9*(3), 936–955. https://doi.org/10.15166/2499-8249/793
8. Haripriya, R., Khare, N., & Pandey, M. (2025). Privacy-preserving federated learning for collaborative medical data mining in multi-institutional settings. *Scientific Reports*, (15), Article 12482 (2025). https://doi.org/10.1038/s41598-025-97565-4
9. Kaminski, M. E., & Malgieri, G. (2025). Impacted stakeholder participation in AI and data governance. *Yale Journal of Law & Technology*, *27*(1), 247–326.
10. Kindt, E. J. (2025). EU biometric data regulation: Part 2: The AI Act. *IEEE Biometrics Council Newsletter*, *54*, 30–41.
11. Laato, S., Birkstedt, T., Mäntymäki, M., Minkkinen, M., & Mikkonen, T. (2022). AI governance in the system development life cycle: Insights on responsible machine learning engineering. In *Proceedings of the 1st*

International Conference on AI Engineering: Software Engineering for AI (pp. 113–123). ACM. https://doi.org/10.1145/3522664.3528598

12. Liu, W., Zhang, Y., Yang, H., & Meng, Q. (2024). A survey on differential privacy for medical data analysis. *Annals of Data Science*, *11*, 733–747. https://doi.org/10.1007/s40745-023-00475-3
13. Malgieri, G., & Santos, C. (2025). Assessing the (severity of) impacts on fundamental rights. *Computer Law & Security Review*, *56*, Article 106113. https://doi.org/10.1016/j.clsr.2025.106113
14. Mantelero, A. (2024). The Fundamental Rights Impact Assessment (FRIA) in the AI Act: Roots, legal obligations and key elements for a model template. *Computer Law & Security Review*, *54*, Article 106020. https://doi.org/10.1016/j.clsr.2024.106020
15. Menges, F., Latzo, T., Vielberth, M., Sobola, S., Pöhls, H. C., Taubmann, B., Köstler, J., Puchta, A., Freiling, F., Reiser, H. P., & Pernul, G. (2021). Towards GDPR-compliant data processing in modern SIEM systems. *Computers & Security*, *103*, Article 102165. https://doi.org/10.1016/j.cose.2020.102165
16. Novelli, C., Casolari, F., Hacker, P., Spedicato, G., & Floridi, L. (2024a). Generative AI in EU Law: Liability, privacy, intellectual property, and cybersecurity. *Computer Law & Security Review*, *55*, Article 106066. https://doi.org/10.1016/j.clsr.2024.106066
17. Novelli, C., Casolari, F., Rotolo, A., Taddeo, M., & Floridi, L. (2024b). AI risk assessment: A scenario-based, proportional methodology for the AI Act. *Digital Society*, *3*, Article 13. https://doi.org/10.1007/s44206-024-00095-1
18. Paullada, A., Raji, I. D., Bender, E. M., Denton, E., & Hanna, A. (2021). Data and its (dis)contents: A survey of dataset development and use in machine learning research. *Patterns*, *2*(11), Article 100336. https://doi.org/10.1016/j.patter.2021.100336
19. Rocks, J. W., & Mehta, P. (2022). Memorizing without overfitting: Bias, variance, and interpolation in over-parameterized models. *Physical Review Research*, *4*(1), Article 013201. https://doi.org/10.1103/PhysRevResearch.4.013201
20. Slijepčević, D., Henzl, M., Klausner, L. D., Dam, T., Kieseberg, P., & Zeppelzauer, M. (2021). *k*-Anonymity in practice: How generalisation and suppression affect machine learning classifiers. *Computers & Security*, *111*, Article 102488. https://doi.org/10.1016/j.cose.2021.102488
21. Sovrano, F., Hine, E., & Anzolut, S. (2025). Simplifying software compliance: AI technologies in drafting technical documentation for the AI Act. *Empirical Software Engineering*, *30*, Article 91. https://doi.org/10.1007/s10664-025-10645-x
22. Söderlund, K., & Larsson, S. (2024). Enforcement Design Patterns in EU Law: An Analysis of the AI Act. *Digital Society*, *3*, Article 41. https://doi.org/10.1007/s44206-024-00129-8
23. Van Bekkum, M. (2025). Using sensitive data to de-bias AI systems: Article 10(5) of the EU AI Act. *Computer Law & Security Review*, *56*, Article 106115. https://doi.org/10.1016/j.clsr.2025.106115
24. Van Bekkum, M., & Zuiderveen Borgesius, F. (2023). Using sensitive data to prevent discrimination by artificial intelligence: Does the GDPR need a new exception? *Computer Law & Security Review*, *48*, Article 105770. https://doi.org/10.1016/j.clsr.2022.105770
25. Veltmeijer, E., & Gerritsen, C. (2025). Legal and ethical implications of AI-based crowd analysis: The AI Act and beyond. *AI and Ethics*, *5*, 3173–3183. https://doi.org/10.1007/s43681-024-00644-x
26. Wang, L., Liu, Z., Liu, A., & Tao, F. (2021). Artificial intelligence in product lifecycle management. *The International Journal of Advanced Manufacturing Technology*, *114*, 771–796. https://doi.org/10.1007/s00170-021-06882-1
27. Winau, M. (2023). On the lack of substantive balancing and coordinated legal concretisation in the European Commission's proposal for a regulation on AI. *European Data Protection Law Review*, *9*(2), 123–135. https://doi.org/10.21552/edpl/2023/2/7

Сведения об авторе:

Олифиренко А. А. — магистрант, кафедра информационного права и цифровых технологий, Саратовская государственная юридическая академия; магистрант, кафедра «Информационная безопасность автоматизированных систем», Институт электронной техники и приборостроения, Саратовский государственный технический университет имени Ю. А. Гагарина, Саратов, Россия; специалист по защите данных, ответственный за безопасность ИИ-систем, ООО «Экосистема недвижимости "Метр квадратный"», Москва, Россия.
panolifer@gmail.com
ORCID: https://orcid.org/0000-0002-2186-281X

Information about the author:

Artem A. Olifirenko — Master's student, Department of Information Law and Digital Technologies, Saratov State Law Academy; Master's student, Department of Information Security of Automated Systems, Institute of Electronic Engineering and Instrumentation, Yuri Gagarin State Technical University of Saratov, Saratov, Russia; Data Protection Specialist, responsible for AI governance and security, "Ecosystem Real Estate 'Metr Kvadratny'" LLC, Moscow, Russia.
panolifer@gmail.com
ORCID: https://orcid.org/0000-0002-2186-281X