




«Искусственный
интеллект в фокусе
права»

ОЭСР (OECD)

Принципы ОЭСР в области искусственного интеллекта (2019)

- 1) **Инклюзивный рост, устойчивое развитие и благополучие;**
 - 2) **Человекоцентричность и справедливость;**
 - 3) **Прозрачность и понятность (объяснимость);**
 - 4) **Надежность, защищенность и безопасность ;**
 - 5) **Подотчетность**
- 

СТРАТЕГИЧЕСКИЙ ПОДХОД

По состоянию на май 2023 года в 51 стране действуют национальные стратегии в области искусственного интеллекта.

Япония (AI Strategy, 2022)

Стратегия ИИ-2022 определила пять стратегических целей :


- **человеческие ресурсы**
- **конкурентоспособность промышленности**
- **технологические системы**
- **международное сотрудничество**
- **использование ИИ в условиях кризисов, например, пандемии, стихийных бедствий.**

В основе три принципа:

1. **Общество, в котором главным является уважение человеческого достоинства.**
2. **Общество, в котором люди с различным происхождением и образованием способны стремиться и достигать своих целей и индивидуального счастья (диверсификация и инклюзивность).**
3. **Общество устойчивого развития.**

Великобритания (National AI Strategy, 2021)

Основные задачи Национальной стратегии Великобритании в области искусственного интеллекта:

- **инвестирование и планирование долгосрочных потребностей экосистемы искусственного интеллекта для сохранения лидерства в качестве научной и ИИ-сверхдержавы;**
 - **обеспечение перехода к экономике с повсеместным применением ИИ, в которой ИИ приносит пользу всем секторам экономики и регионам Великобритании;**
 - **обеспечение надлежащего управления технологиями ИИ в целях поощрения инноваций, инвестиций, защиты общества и фундаментальных ценностей.**
- 

Индия (National Strategy For Artificial Intelligence, 2018)

Структура

- 1) ИИ для всех;
- 2) ответственный ИИ для расширения социальных возможностей;
- 3) лучший поставщик решений для стран с формирующимся рынком и развивающихся экономик (за исключением Китая) по всему миру;
- 4) разработка и внедрение языковых технологий на базе искусственного интеллекта.

Приоритетные сферы внедрения ИИ

- Здравоохранение;
- Сельское хозяйство;
- Образование;
- «Умные» города и инфраструктура;
- «Умная» мобильность и транспорт.

В отдельных правовых рамках национальная стратегия в области ИИ является частью более широкой стратегии цифровизации. Например, «Цифровая Швейцария» (Digital Switzerland Strategy) – всеобъемлющая стратегия по всем цифровым направлениям, включая ИИ. Ее принципы применимы к ИИ и другим технологиям.

В целях эффективного осуществления государственной политики в области ИИ в различных правовых рамках используются такие модели, как:

- 1) создание нового правительственного или координационного органа по ИИ;
- 2) установление надзора за разработкой и внедрением стратегии в отношении действующего министерства и создание межведомственных комитетов по ИИ;
- 3) получение информации от надзорных и экспертно-консультативных органов или групп по ИИ и органов по этике данных.

Государства также создают механизмы мониторинга и оценки своих национальных стратегий в области ИИ.

Совершенствование правового регулирования в области ИИ предполагает:

- 1) установление этических рамок и принципов;
- 2) разработку и внедрение подходов, основанных на жестком законодательном регулировании;
- 3) разработку международных норм и стандартов в области ИИ;
- 4) правовые эксперименты.

17 государств разработали и внедрили национальные этические рамки и принципы разработки и внедрения ИИ, которые в значительной степени совпадают с принципами ИИ ОЭСР.

Отдельные государства (Япония, Корея и Индия) предоставляют разработчикам и операторам руководящие указания по внедрению принципов. Кроме того, Колумбия создала онлайн-платформу для мониторинга внедрения системы.

Подходы к определению понятия «искусственный интеллект»

1) Система искусственного интеллекта - это машинная система, которая для достижения явных или неявных целей делает вывод на основе получаемых входных данных, как генерировать выходные данные, такие как прогнозы, контент, рекомендации или решения, которые могут влиять на физическую или виртуальную среду. Различные системы искусственного интеллекта различаются по уровню автономности и адаптивности после применения (Рекомендации ОЭСР, Recommendation of the Council on OECD Legal Instruments Artificial Intelligence); Artificial Intelligence Act 2024.


2) Искусственный интеллект – это совокупность технологий, которые позволяют машинам действовать с более высоким уровнем интеллекта и имитировать человеческие способности чувствовать, осмысливать и действовать. Таким образом, компьютерное зрение и обработка звука могут активно воспринимать окружающий мир, получая и обрабатывая изображения, звук и речь. Механизмы обработки естественного языка и логического вывода могут позволить системам искусственного интеллекта анализировать и понимать собранную информацию (National Strategy For Artificial Intelligence, 2018; Указ Президента РФ «О развитии искусственного интеллекта в Российской Федерации»).

Европейский Союз

Нормы гражданского права о робототехнике, Резолюция Европарламента от 16 февраля 2017 г.

Призывает комиссию ЕС предложить общепринятые и универсальные определения терминов «киберфизические системы», «автономные системы», «умные автономные роботы», а также их подкатегорий.

При этом необходимо учитывать следующие характеристики умного робота:

- способность становиться автономным, используя сенсоры и/или обмениваясь данными со своей средой (совместимость), способность обмениваться этими данными и анализировать их;
 - способность самообучаться на основе приобретенного опыта и при взаимодействии (опциональный критерий);
 - наличие но меньшей мере минимальной физической поддержки;
 - способность адаптировать свои действия и поведение в соответствии с
 - условиями среды;
 - отсутствие жизни с биологической точки зрения;
- 

f) о наделении в перспективе роботов особым правовым статусом. Таким образом, по меньшей мере наиболее продвинутые автономные роботы могут создаваться как электронные липа и нести ответственность за причиненный ими ущерб в тех случаях, когда они принимают решения автономно или иным образом самостоятельно взаимодействуют с третьими лицами.



Artificial Intelligence Act 2024


Устанавливает классификацию систем искусственного интеллекта

- 1) Системы искусственного интеллекта с высоким риском;
- 2) Системы искусственного интеллекта с наименьшим риском (дипфейк);
- 3) Запрещенные системы искусственного интеллекта (когнитивно-поведенческое манипулирование людьми или конкретными уязвимыми группами; социальный рейтинг: классификация людей на основе поведения, социально-экономического статуса или личных характеристик; биометрическая идентификация и категоризация людей).

Требования к системам искусственного интеллекта с высоким риском

Данные и управление данными, техническая документация, запись события (лог),

Запрещенные системы искусственного интеллекта

- применение систем ИИ для разделения людей на категории по «чувствительным характеристикам»: пол, раса, гражданство, религия, политическая позиция;
 - применение систем ИИ в целях распознавания эмоций в деятельности правоохранительных органов, пограничных служб, учебных заведений, а также на рабочих местах;
 - предиктивные системы, используемые правоохранительными органами (например, подсказывающие, имеет ли тот или иной человек намерение совершить преступление);
 - неизбирательное извлечение биометрических данных из социальных сетей или с камер видеонаблюдения для создания базы данных.
- 

Системы искусственного интеллекта с высоким риском

- системы ИИ, используемые в критической инфраструктуре;
- системы биометрической идентификации (за исключением запрещенных);
- системы, используемые в образовании (оценивающие учащихся или определяющие, кому предоставить доступ к обучению);
- системы оценки работников: потенциальных (при найме на работу) или уже сотрудничающих с организацией;
- системы оценки кредитоспособности;
- «администрирование» правосудия (по сути, системы-«помощники» судьи в исследовании обстоятельств дела и их соответствия нормам закона).

В перечень высокорисковых также включены системы ИИ, которые используются в целях влияния на избирателей в политических кампаниях, и рекомендательные алгоритмы соцсетей.

Все высокорисковые системы признаются таковыми, если они представляют существенный риск причинения вреда здоровью, безопасности или основополагающим правам граждан, оказывают существенное влияние на результаты принятия решений (ст. 6).



Требования к системам искусственного интеллекта с высоким риском

- к технической документации
- к риск-менеджменту,
- к процессам по сбору, анализу и качеству данных, на которых обучается и тестируется система,
- к обеспечению фиксации и записи всех событий функционирования систем.

Перед допуском на рынок продукция с системой ИИ должна пройти процедуру регистрации.

Также деплойеры (те, кто устанавливает и эксплуатирует системы) должны регистрировать продукт перед выпуском его на рынок, в том числе если они являются публичными организациями .

прозрачности

ChatGPT

Генеративный ИИ, такой как ChatGPT, не отнесен к категории высокого риска, но должен соответствовать требованиям прозрачности и закону ЕС об авторском праве.

Обязательства провайдеров таких систем (foundation models — «базовые модели») должны гарантировать высокий уровень безопасности соблюдения фундаментальных прав, здоровья, безопасности, окружающей среды, демократии и верховенства права.

Требования

в том числе к самим базовым моделям, к тому, на каких данных они обучаются и развиваются (данные должны быть надежны и непредвзяты), к их возможностям оставаться предсказуемыми в своих выводах

Провайдеры должны обеспечить наличие всей необходимой технической документации и, как в случае с высокорисковыми системами, тоже зарегистрировать базовые модели. А если эти модели еще и генерируют контент, как в случае с Chat GPT или Midjourney, — провайдеры должны будут документировать и публиковать информацию о том, как соблюдаются авторские права при тренировке системы.

Раскрытие информации о том, что контент был создан искусственным интеллектом.

Публикация резюме защищенных авторским правом данных, используемых для обучения.



Штрафы

Государства-члены ЕС должны установить правила о штрафах и других принудительных мерах, которые могут также включать предупреждения и неденежные меры, применимые к нарушениям настоящего Закона.

Предусмотренные штрафные санкции должны быть действительными, пропорциональными и оказывают сдерживающее воздействие. Они должны учитывать интересы малых и средних предприятий, включая начинающие, и их экономическую жизнеспособность.

Несоблюдение требований, применяемых к системам искусственного интеллекта, влечет наложение административного штрафа в размере до 35 000 000 евро или, если нарушителем является предприятие, до 7% от его общего годового оборота по всему миру за предыдущий финансовый год.

Российская Федерация

Указ Президента РФ от 10.10.2019 № 490 «О развитии искусственного интеллекта в Российской Федерации» (вместе с «Национальной стратегией развития искусственного интеллекта на период до 2030 года»)

искусственный интеллект – комплекс технологических решений, позволяющий имитировать когнитивные функции человека (включая поиск решений без заранее заданного алгоритма) и получать при выполнении конкретных задач результаты, сопоставимые с результатами интеллектуальной деятельности человека или превосходящие их. Комплекс технологических решений включает в себя информационно-коммуникационную инфраструктуру, программное обеспечение (в том числе в котором используются методы машинного обучения), процессы и сервисы по обработке данных и поиску решений;

технологии искусственного интеллекта – совокупность технологий, включающая в себя компьютерное зрение, обработку естественного языка, распознавание и синтез речи, интеллектуальную поддержку принятия решений и перспективные методы искусственного интеллекта;

Распоряжение Правительства РФ от 19.08.2020 № 2129-р «Об утверждении Концепции развития регулирования отношений в сфере технологий искусственного интеллекта и робототехники до 2024 года»

Отсутствие однозначного понимания содержания терминов «искусственный интеллект», «робот», «умный робот», «робототехника», «интеллектуальный агент» приводит к терминологическим проблемам при формировании регулирования.

В то же время с учетом прикладного характера применения этих технологий в самых разных областях может потребоваться формулирование разных определений в зависимости от отрасли применения технологий искусственного интеллекта и робототехники. По этой причине в рассматриваемый в Концепции временной период предлагается по возможности избегать внедрения в законодательство Российской Федерации единого для всех отраслей нормативного определения указанных терминов. При этом крайне важно в этот период провести работу по построению и гармонизации онтологии предметной области силами экспертного сообщества и профильных технических комитетов при Федеральном агентстве по техническому регулированию и метрологии. Там, где это необходимо для целей определения предмета регулирования конкретных нормативных правовых актов, предлагается использовать определения, содержащиеся в документах по стандартизации, либо давать определения, актуальные конкретно для этой сферы регулирования.

Федеральный закон от 24.04.2020 № 123-ФЗ «О проведении эксперимента по установлению специального регулирования в целях создания необходимых условий для разработки и внедрения технологий искусственного интеллекта в субъекте Российской Федерации - городе федерального значения Москве и внесении изменений в статьи 6 и 10 Федерального закона «О персональных данных»

2) искусственный интеллект - комплекс технологических решений, позволяющий имитировать когнитивные функции человека (включая самообучение и поиск решений без заранее заданного алгоритма) и получать при выполнении конкретных задач результаты, сопоставимые, как минимум, с результатами интеллектуальной деятельности человека. Комплекс технологических решений включает в себя информационно-коммуникационную инфраструктуру (в том числе информационные системы, информационно-телекоммуникационные сети, иные технические средства обработки информации), программное обеспечение (в том числе в котором используются методы машинного обучения), процессы и сервисы по обработке данных и поиску решений;

3) технологии искусственного интеллекта - технологии, основанные на использовании искусственного интеллекта (включая компьютерное зрение, обработку естественного языка, распознавание и синтез речи, интеллектуальную поддержку принятия решений и перспективные методы искусственного интеллекта);

КОДЕКС ЭТИКИ В СФЕРЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Кодекс этики в сфере искусственного интеллекта устанавливает общие этические принципы и стандарты поведения, которыми следует руководствоваться участникам отношений в сфере искусственного интеллекта в своей деятельности, а также механизмы реализации положений настоящего Кодекса.

Кодекс распространяется на отношения, связанные с этическими аспектами создания (проектирования, конструирования, пилотирования), внедрения и использования технологий ИИ на всех этапах жизненного цикла, которые в настоящее время не урегулированы законодательством Российской Федерации и/или актами технического регулирования.

Кодекс - единая система рекомендательных принципов и правил, предназначенных для создания среды доверенного развития технологий искусственного интеллекта в России

Носит рекомендательный характер

Присоединение осуществляется на добровольной основе

Распространяется только на гражданские разработки

ПРИНЦИПЫ ЭТИКИ И ПРАВИЛА ПОВЕДЕНИЯ

1. Главный приоритет развития технологий ИИ в защите интересов и прав людей и отдельного человека
 - 1.1. Человеко-ориентированный и гуманистический подход.
 - 1.2. Уважение автономии и свободы воли человека.
 - 1.3. Соответствие закону.
 - 1.4. Недискриминация.
 - 1.5. Оценка рисков и гуманитарного воздействия.
 2. Необходимо осознавать ответственность при создании и использовании ИИ.
 - 2.1. Риск-ориентированный подход
 - 2.2. Ответственное отношение.
 - 2.3. Предосторожность.
 - 2.4. Непричинение вреда.
 - 2.5. Идентификация ИИ в общении с человеком.
 - 2.6. Безопасность работы с данными.
 - 2.7. Информационная безопасность.
 - 2.8. Добровольная сертификация и соответствие положениям Кодекса.
 3. Ответственность за последствия применения сии всегда несет человек.
- 