Ключевые мировые тенденции в регулировании ИИ

С. В. Габуев

Заместитель директора Департамента развития искусственного интеллекта и больших данных (Минцифры России)

12 февраля 2026 г.

Актуальность исследования тенденций_

- Внедрение ИИ/GenAI становится **массовым**, а право **догоняет** рынок.
- McKinsey (2024): **65%** организаций **регулярно** используют генеративный ИИ; **78%** организаций используют ИИ **хотя бы в одной** функции (2024), **рост с 55%** годом ранее
- Stanford AI Index 2025: **59** AI-регуляторики/**правил** федеральных агентств США в **2024**; упоминания ИИ в законодательных инициативах **выросли** на **21,3%** по **75** странам (2024 vs 2023), и примерно **в 9 раз к 2016**
- Интероперабельность: государства стремятся к **совместимости** подходов

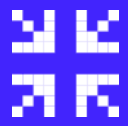
Модель регулирования ИИ: США_

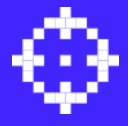
Общий подход: преимущественно «про-инновационный» – управление рисками через ведомственные политики, стандарты и закупочные требования; единого всеобъемлющего федерального «AI Act» нет.

Госуправление и закупки: курс на ускоренное внедрение ИИ в федеральных органах при одновременной институционализации риск-менеджмента (надзор OMB по закупке/использованию ИИ).

Стратегический трек: акцент на конкурентоспособность, национальную безопасность, эффективность госфункций; развитие инфраструктуры и кадров.

Модель регулирования ИИ: ЕС_





Нормативная база: Регламент (EU) 2024/1689 (AI Act) — единые правила по всему ЕС, вводятся поэтапно.

Ключевая логика: запрет отдельных практик + требования к «высокорисковым» системам + обязанности по прозрачности/управлению рисками.

Отдельный блок: требования к провайдерам GPAI/«системных» моделей (governance, оценка рисков, документация/прозрачность в установленных случаях).

Модель регулирования ИИ: КНР_

Модель регулирования: набор специализированных правил под онлайн-контент/алгоритмы и киберуправление; баланс «развитие + безопасность».

Generative AI: Interim Measures (2023) — требования к провайдерам публичных GenAI-сервисов (безопасность, законность, контент-ограничения, управление рисками).

Актуальный вектор: усиление регулирования «человеко-подобных»/эмоционально вовлекающих ИИ-сервисов (проект правил, публичное обсуждение).

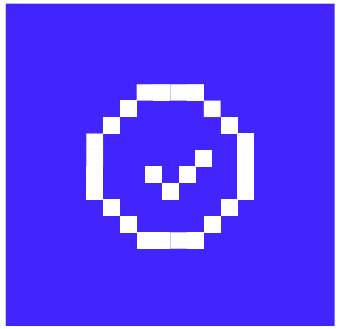
Модель регулирования ИИ: Южная Корея_

AI Basic Act: формирует общую рамку (институции, безопасность/доверие, стимулы развития), вступление в силу — январь 2026.

Содержательная ось: требования для «high-impact AI», оценка рисков, отчётность/обязанности для отдельных классов систем; сочетание регулирования и поддержки отрасли.

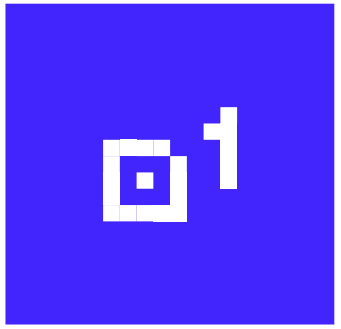
Маркировка AI-контента: движение к обязательному обозначению AI-генерации (в т.ч. в рекламе) как ответ на злоупотребления.

Дихотомия регуляторики ИИ: развитие vs безопасность_



Инновации

данные, выч. мощности, кадры, гранты, внедрение в госсектор



Риски и права

безопасность, прозрачность, ответственность, маркировка, контроль высокорисковых систем

Дискуссионные узлы для будущего регулирования_

01

«Не как в ЕС»

не «европейская» риск-классификация, а решения за отраслевыми регуляторами/ риск фрагментации и неравных стандартов

02

Про-инновационная модель

приоритет стимулирования/ недостаточность “soft law” для высоких рисков

03

Иностранные модели и суверенитет

приоритет национальных решений в госсекторе / возможное падение качества, доступности и рост стоимости

04

Данные и датасеты

«управление качеством и прослеживаемость»/ чрезмерная бюрократизация подтверждения прав/источников и с избыточными реестровыми раскрытиями

05

Ответственность

Распределение ответственности по ролям (разработчик, оператор и тд.) и компенсационных механизмов/ сложность доказывания причинно-следственной связи и “over-compliance”

Архитектура регулирования: риск-ориентированность и рамочный закон_

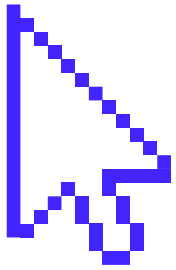
Возможность закрепления в риск-ориентированной модели, при которой объём требований зависит от сферы применения и уровня риска

Базовый закон целесообразно делать рамочным (термины, роли, принципы, полномочия), а детализацию переносить в подзаконные акты и стандарты

Дискуссионным остаётся вопрос о степени унификации: единая классификация и требования «для всех» или преимущественно отраслевые режимы при общей рамке

Стандарты испытаний и оценка качества_

- Единые подходы к испытаниям и оценке ИИ: закреплять через стандарты и методики (в т.ч. национальные) или через детальные нормы закона?
- Требования должны быть проверяемыми: методики валидации, показатели качества, процедуры аудита и испытаний
- Где проводить границу между добровольными стандартами и обязательными требованиями для высокорисковых систем?

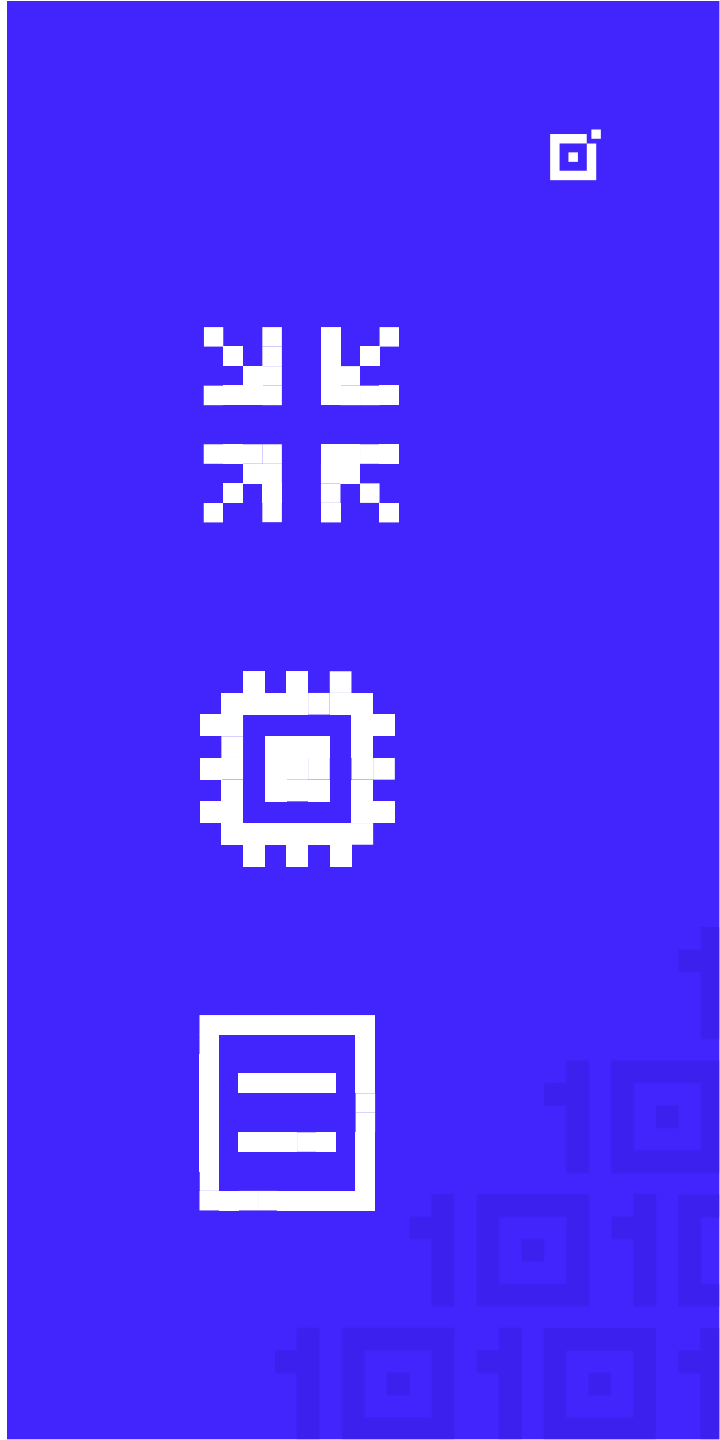


Данные и датасеты: качество, доступ, права_

- Регулирование данных фокусируется на управляемости: паспорт набора данных, метаданные, версии, прослеживаемость происхождения, базовые показатели качества

- Режимы доступа к данным: обезличивание, анонимизация, санитизация, доверенные платформы. **Нужен ли особый режим доступа к данным для ИИ?**

- Баланс прозрачности и реализуемости: чрезмерно детальные требования к подтверждению прав и «очистке» больших датасетов могут повышать издержки и снижать полноту данных



Ответственность и страхование_

Ответственность целесообразно описывать через разграничение ролей (разработчик, интегратор, оператор, заказчик) и связь ответственности с фактическим контролем над применением

Обязательность страхования - только для высокорисковых случаев, в остальных - добровольные механизмы

Иностранные модели и суверенность_

- В критических сферах (госсектор, КИИ, оборона) закрепляется приоритет отечественных решений через закупочные и контурные требования
- Предмет дискуссии: как формализовать критерии «суверенности» и режимы допуска, чтобы избежать как изоляции, так и зависимости

Бенефициары ИИ: интересы и точки регулирования_

Государство

- Эффективность госуправления и услуг
- Безопасность и технологический суверенитет
- Управляемость и соблюдение законности

Общество

- Безопасность и доверие
- Защита персональных данных и прав
- Качество, недискриминация, возможность оспаривания

Контур регулирования

- Риск-ориентированные требования
- Стандарты и испытания
- Данные и доступ (лицензии)
- Прозрачность и маркировка
- Ответственность и надзор
- Стимулы и инфраструктура

Бизнес

- Предсказуемые и исполнимые требования
- Доступ к данным и вычислительной инфраструктуре
- Защита IP, коммерческой тайны и ноу-хау

*Точки напряжения:
прозрачность vs ноу-хау
доступ к данным vs защита прав/ИБ
инновации vs безопасность

Вызовы_

Меры противодействия преступной деятельности с применением ИИ становятся все более актуальны («дипфейки»)

Необходимость маркировки контента создаваемого ИИ: соблюдению принципа достоверности информации vs удобства использования ИИ.

Регулирование прав на РИДы, созданные с использованием ИИ: интересы правообладателей vs интересы разработчиков ИИ решений.

число распространяемых дипфейков выросло:

~500 000 (2023)

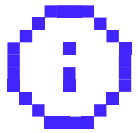
↓

~8 000 000 (2025)

86%

создателей контента (creators) уже активно используют креативный генеративный ИИ в работе (2025)

Выводы_



- Регулирование ИИ должно быть в первую очередь направлено на создание условий для его развития.
- Ключевую роль играет доступ к данным и выч. мощностям необходимым для обучения системы ИИ.
- Нужны понятные и прозрачные правила обезличивания персональных данных для их использования разработчиками ИИ. Необходимо активнее использовать для этих целей и государственные данные.
- Все это с учетом существующих рисков и угроз в сфере информационной безопасности.
- Другим важным направлением регулирования ИИ является защита наших граждан от незаконного использования ИИ.