



**95**  
ПЕРВЫЙ  
ПО ПРАВУ



ИНСТИТУТ ИНФОРМАЦИОННОЙ  
И МЕДИАБЕЗОПАСНОСТИ МГЮА

# Регулирование генеративного ИИ в целях противодействия распространению фейковой информации за рубежом

Никишин Владимир Дмитриевич - к.ю.н., директор Института информационной и медиабезопасности МГЮА, доцент кафедры информационного права и цифровых технологий Университета имени О.Е. Кутафина (МГЮА)

---

Москва, 2026

# Вводные положения

Отсутствие унифицированного подхода к дефиниции дипфейка и правому режиму генеративного ИИ и технологии глубокого синтеза создает значительные сложности в разработке эффективных механизмов противодействия фейковизации и связанным с ней рискам;

В законодательной терминологии англоязычных стран понятие «deepfake» является общеупотребимым. В других странах понятие «дипфейк», как правило, раскрывается через определение использования **генеративного искусственного интеллекта с определенной целью** (ввести в заблуждение, изменить содержание и т.д.);

В рамках настоящего доклада будут рассмотрены ключевые концепции регулирования генеративного ИИ в различных юрисдикциях - Китай, Южная Корея, Европейский Союз, Соединенные Штаты Америки

## Информация к размышлению



В декабре 2025 года **16% всех проверенных фактов** (из 1605 статей) в Европе касались **контента, созданного или измененного с помощью ИИ**. Это свидетельствует о том, что ИИ стал центральным драйвером дезинформации (*55 отчет European Digital Media Observatory*)

Только один ИИ-инструмент (Grok от xAI) сгенерировал около **3 миллионов изображений** сексуального характера **за 11 дней** (с 29 декабря по 9 января). Это примерно 190 изображений в минуту, и около 23 000 из них содержали изображения несовершеннолетних (*Daily Digital Digest*)

Тестирование шести популярных ИИ-моделей показало, что даже лучшая из них (ChatGPT) **дает полностью корректные ответы лишь в 59,7% случаев**. У худшей (Grok) этот показатель составил всего 39,6%, а доля откровенно неверных ответов достигла 21,8% (*PPC Land*)

# Общий сравнительный обзор юрисдикций



Юрисдикция	Основной подход	Главные требования / Инструменты	Направленность регулирования
Китай	Административно-превентивный	<b>Обязанности провайдеров:</b> маркировка (явная/неявная), аутентификация пользователей (по паспорту/телефону), проверка алгоритмов, хранение логов (6 мес.). Запрет на создание и распространение ложной информации.	<b>Защита общества и госбезопасности.</b> Приоритет общественных интересов над частными. Борьба с анонимностью.
Европейский союз	Антропоцентричный (Риск-ориентированный)	<b>AI Act:</b> Прозрачность для дипфейков (маркировка контента провайдерами). Запрет манипулятивных практик (эксплуатация уязвимостей). Штрафы до 7% оборота. Обязанность онлайн-платформ оценивать риски от дипфейков и маркировать синтетический контент.	<b>Защита фундаментальных прав и свобод человека.</b> Баланс между инновациями и безопасностью. Запрет на нанесение значительного вреда.
США (Федеральный уровень)	Точечное регулирование. Акцент на инновационное развитие. (Узконаправленный)	Борьба с <b>дипфейками интимного содержания</b> . Обязанность платформ удалять такой контент за 48 часов по требованию жертвы. Введение уголовной ответственности (до 3 лет лишения свободы).	<b>Защита частных лиц от специфического вреда</b> (порнография без согласия, вымогательство). Минимум барьеров для инноваций.
США (Штаты, на примере Калифорнии)	Узконаправленное регулирование	<b>Широкий спектр:</b> маркировка в метаданных для генеративных ИИ-систем, удаление/маркировка «фейкового контента» в предвыборный период, запрет на цифровые копии актеров без согласия, уголовная ответственность за ИИ-детскую порнографию.	<b>Защита демократических процессов, прав личности (актеров, граждан) и несовершеннолетних.</b> Заполнение пробелов федерального законодательства.
Южная Корея	Рамочно-специализированный	<b>Основной закон:</b> Требования прозрачности к генеративному ИИ (уведомление о синтетической природе контента). Отсутствие прямых запретов на типы ИИ. Специальный запрет на дипфейки в предвыборной агитации за 90 дней до голосования.	Баланс между <b>европейским риск-ориентированным подходом</b> (без жестких запретов) и <b>китайским фокусом на инновации</b> . Защита избирательного процесса.



# Китайская народная республика



В Китае **отсутствует законодательное регулирование технологии генеративного ИИ (глубинного синтеза)**. Однако положения о правилах использования таких технологий урегулированы на подзаконном уровне (на уровне министерств и ведомств).

- Временные меры по управлению сервисами генеративного ИИ с 15 августа 2023 года;
- Положение об управлении глубоким синтезом информационных интернет-услуг (является основным актом в вопросах регулирования генеративного ИИ) с 10 января 2023 года;
- Меры по выявлению синтетического контента, созданного ИИ с 1 сентября 2025 года.

## Положения устанавливают требования к операторам услуг глубокого синтеза

Оператор (поставщик) услуг глубокого синтеза **обязан** заметно обозначить сгенерированный или отредактированный контент **в разумном месте или области** с целью информирования о применении технологии глубинного синтеза, *если оператор представляет услуги, которые могут вызвать беспорядки или быть неправильно поняты общественностью*. Например:

- интеллектуальный разговор, интеллектуальное письмо или другие услуги, которые имитируют физическое лицо для создания или редактирования текста;
- услуги редактирования, которые генерируют голоса, такие как синтезированные человеческие голоса или имитированные голоса, или которые существенно изменяют характеристики личности;
- услуги редактирования, которые генерируют изображения или видео людей, такие как генерация лиц, замена лиц, манипуляция лицами, манипуляция позами или которые существенно изменяют черты личности;
- услуги по созданию и редактированию захватывающих и реалистичных сцен;
- иные услуги, функция которых заключается в создании или существенном изменении информационного содержания.





# Китайская народная республика



Положение о глубоком синтезе предусматривают, что поставщики услуг глубокого синтеза **должны внедрять и раскрывать правила управления**, включая внедрение системы аутентификации реальной идентификационной информации.

Аутентификация пользователей, использующих сервисы технологий глубокого синтеза осуществляется на основе:

- Временные меры по управлению сервисами генеративного ИИ;
- Положение об управлении глубоким синтезом информационных интернет-услуг (**является основным актом в вопросах регулирования генеративного ИИ**);
- Меры по выявлению синтетического контента, созданного ИИ.

**NB:** в случаях, когда персональные данные используются для обучения алгоритмов, а функции позволяют редактировать биометрическую информацию, такую как лицо или голос, субъекты данных должны быть уведомлены об этом и от них должно быть получено отдельное согласие

По состоянию на декабрь 2025 года число пользователей генеративного ИИ в Китае достигло **602 миллионов человек**.

*(EDGAR (Electronic Data Gathering, Analysis, and Retrieval))*

За 2025 год аудитория выросла **на 141.7%**, и теперь генеративным ИИ пользуются **42.8%** всех интернет-пользователей страны.

*(газета «Жэньминь Жибао»)*

**Модель DeepSeek V3.2 стоит всего \$0.28-\$0.42 за миллион токенов**, в то время как западные аналоги (GPT-5 от OpenAI) оцениваются в \$1.25-\$10. Это делает китайские технологии глубокого синтеза сверхдоступными.

*(TrustSolutions)*

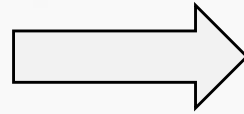


# Китайская народная республика



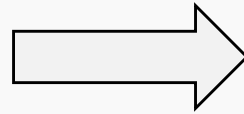
С 1 сентября 2025 вступили в силу **Меры по выявлению синтетического контента, созданного искусственным интеллектом**. Данные меры устанавливают следующие обязанности:

**Поставщики (операторы) услуг искусственного интеллекта (любые сервисы, использующие алгоритмы глубинного обучения для создания контента)**



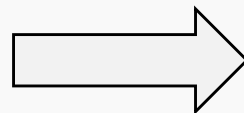
Обязанность осуществлять маркировку контента, сохранять маркировку в файлах, предупреждать пользователей об ответственности за отсутствие маркировки

**Пользователи услуг искусственного интеллекта**



Обязанность раскрывать, что контент сгенерирован ИИ, использовать инструменты маркировки от платформ (поставщиков).  
**Запрещено удалять, подделывать маркировку, создавать инструменты обхода маркировки**

**Онлайн-платформы**



Обязанность выявления ИИ-контента и обеспечение видимости маркировки

# Общий сравнительный обзор юрисдикций



Юрисдикция	Основной подход	Главные требования / Инструменты	Направленность регулирования
Китай	Административно-превентивный	<b>Обязанности провайдеров:</b> маркировка (явная/неявная), аутентификация пользователей (по паспорту/телефону), проверка алгоритмов, хранение логов (6 мес.). Запрет на создание и распространение ложной информации.	<b>Защита общества и госбезопасности.</b> Приоритет общественных интересов над частными. Борьба с анонимностью.
Европейский союз	Антропоцентричный (Риск-ориентированный)	<b>AI Act:</b> Прозрачность для дипфейков (маркировка контента провайдерами). Запрет манипулятивных практик (эксплуатация уязвимостей). Штрафы до 7% оборота. Обязанность онлайн-платформ оценивать риски от дипфейков и маркировать синтетический контент.	<b>Защита фундаментальных прав и свобод человека.</b> Баланс между инновациями и безопасностью. Запрет на нанесение значительного вреда.
США (Федеральный уровень)	Точечное регулирование. Акцент на инновационное развитие. (Узконаправленный)	Борьба с <b>дипфейками интимного содержания</b> . Обязанность платформ удалять такой контент за 48 часов по требованию жертвы. Введение уголовной ответственности (до 3 лет лишения свободы).	<b>Защита частных лиц от специфического вреда</b> (порнография без согласия, вымогательство). Минимум барьеров для инноваций.
США (Штаты, на примере Калифорнии)	Узконаправленное регулирование	<b>Широкий спектр:</b> маркировка в метаданных для генеративных ИИ-систем, удаление/маркировка «фейкового контента» в предвыборный период, запрет на цифровые копии актеров без согласия, уголовная ответственность за ИИ-детскую порнографию.	<b>Защита демократических процессов, прав личности (актеров, граждан) и несовершеннолетних.</b> Заполнение пробелов федерального законодательства.
Южная Корея	Рамочно-специализированный	<b>Основной закон:</b> Требования прозрачности к генеративному ИИ (уведомление о синтетической природе контента). Отсутствие прямых запретов на типы ИИ. Специальный запрет на дипфейки в предвыборной агитации за 90 дней до голосования.	Баланс между <b>европейским риск-ориентированным подходом</b> (без жестких запретов) и <b>китайским фокусом на инновации</b> . Защита избирательного процесса.



# Южная Корея



Основной НПА - это **Закон о развитии искусственного интеллекта и установлении доверия**, вступивший в силу в 2026 г., не устанавливает жестких запретов на использование генеративного ИИ.

Акцент сделан на требования **прозрачности и безопасности** использования данных технологий.

## Типы систем ИИ

### Высокоэффективный ИИ

Системы ИИ, представляющие значительный риск воздействия на жизнь человека, физическую безопасность или основные права

### Генеративный ИИ

Системы ИИ, разработанные для имитации структуры и характеристик входных данных для генерации выходных данных (звук, текст, изображение, видео, иной творческий контент)



# Южная Корея



Требования к уведомлению (маркировке) о том, что контент создан ИИ, **обязателен для каждой из систем ИИ, существующих в Корее.**

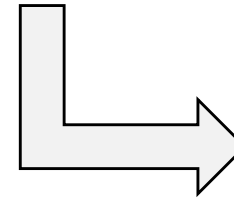
**Обязанность по маркировке возлагается на компании и учреждения, осуществляющие деятельность в области ИИ.**

Маркировка обязана быть четко различима пользователями.

Допускается отображение маркировки образом, не препятствующим показу, если контент, созданный ИИ, является частью художественного или творческого выражения.

## Компании и учреждения обязаны:

- отмечать и раскрывать факт того, что контент был сгенерирован ИИ;
- реализовывать меры по выявлению, оценке и снижению рисков использования ИИ;
- создавать системы управления рисками;
- разъяснять пользователям требования безопасности ИИ;



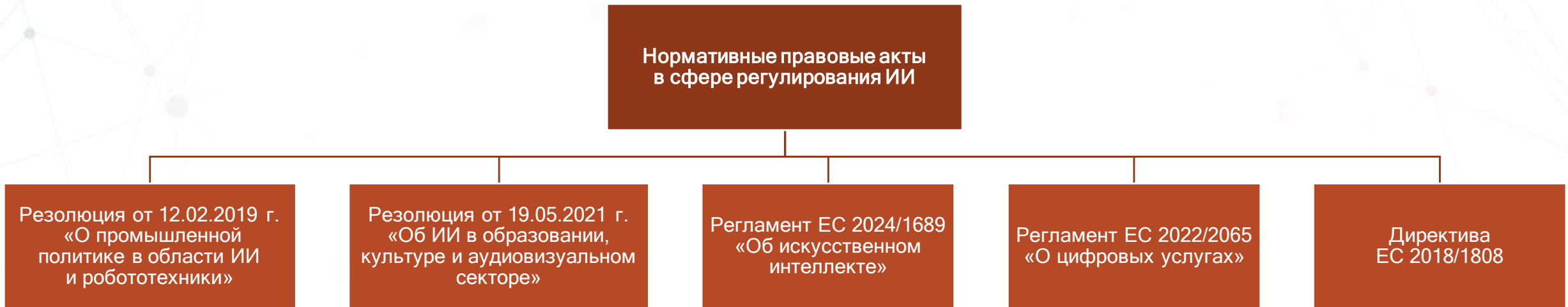
**Контроль осуществляет  
Министр науки и ИКТ.**

**НВ: отдельным актом в сфере регулирования ИИ является Закон о выборах, которым предусмотрен запрет на использование дипфейков в предвыборной компании за 90 дней до голосования**



Европейский Союз отличается от всех иных юрисдикций комплексным подходом к вопросам регулирования генеративного ИИ.

Ключевой акцент регулирования, обязательный для стран-членов ЕС, **направлен на обеспечение защиты прав и интересов граждан при создании, использовании и распространении ИИ-контента**





## Регламент ЕС 2024/1689 «Об искусственном интеллекте»

### Ключевые положения:

Генеративный ИИ для создания дипфейков **не запрещен, но подпадает под регулирование.**

### Запрещены отдельные способы использования ИИ:

- использование **в манипулятивных практиках** (используются для эксплуатирования уязвимости человека в силу возраста, социального или экономического положения, физических или психических недостатков);
- использование в высокорисковых контекстах (дистанционная биометрическая идентификация)

### REGULATION (EU) 2024/1689 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL

of 13 June 2024

laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act)

(Text with EEA relevance)

THE EUROPEAN PARLIAMENT AND THE COUNCIL OF THE EUROPEAN UNION,

Having regard to the Treaty on the Functioning of the European Union, and in particular Articles 16 and 114 thereof,

Having regard to the proposal from the European Commission,

After transmission of the draft legislative act to the national parliaments,

Having regard to the opinion of the European Economic and Social Committee <sup>(1)</sup>,

Having regard to the opinion of the European Central Bank <sup>(2)</sup>,

Having regard to the opinion of the Committee of the Regions <sup>(3)</sup>,

Acting in accordance with the ordinary legislative procedure <sup>(4)</sup>,

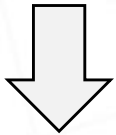
Whereas:

- (1) The purpose of this Regulation is to improve the functioning of the internal market by laying down a uniform legal framework in particular for the development, the placing on the market, the putting into service and the use of artificial intelligence systems (AI systems) in the Union, in accordance with Union values, to promote the uptake of human centric and trustworthy artificial intelligence (AI) while ensuring a high level of protection of health, safety, fundamental rights as enshrined in the Charter of Fundamental Rights of the European Union (the 'Charter'), including democracy, the rule of law and environmental protection, to protect against the harmful effects of AI systems in the Union, and to support innovation. This Regulation ensures the free movement, cross-border, of AI-based goods and services, thus preventing Member States from imposing restrictions on the development, marketing and use of AI systems, unless explicitly authorised by this Regulation.
- (2) This Regulation should be applied in accordance with the values of the Union enshrined as in the Charter, facilitating the protection of natural persons, undertakings, democracy, the rule of law and environmental protection, while boosting innovation and employment and making the Union a leader in the uptake of trustworthy AI.
- (3) AI systems can be easily deployed in a large variety of sectors of the economy and many parts of society, including across borders, and can easily circulate throughout the Union. Certain Member States have already explored the adoption of national rules to ensure that AI is trustworthy and safe and is developed and used in accordance with fundamental rights obligations. Diverging national rules may lead to the fragmentation of the internal market and may decrease legal certainty for operators that develop, import or use AI systems. A consistent and high level of protection throughout the Union should therefore be ensured in order to achieve trustworthy AI, while divergences hampering the free circulation, innovation, deployment and the uptake of AI



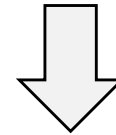
Регламент ЕС 2024/1689 «Об искусственном интеллекте» устанавливает **обязанность для провайдеров и пользователей систем ИИ, генерирующих контент.**

Провайдеры обязаны обеспечить, чтобы их система выводила контент, сгенерированный ИИ, с четким указанием на его природу. То есть, обязательная маркировка (водяные знаки, обязательное уведомление пользователя системы о сгенерированном контенте).



## В случае нарушения

Штраф в размере **до 35 миллионов евро** или **7% от глобального годового оборота компании**



## Исключение

От маркировки освобождается:

- Контент, синтетическая природа которого раскрывается надлежащим образом
- или очевидна из контекста;
- Дипфейки, используемые для «узаконенных» целей: пародия, сатира, художественное выражение.

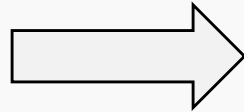


## Регламент ЕС 2022/2065 «О цифровых услугах»

Распространяется на:

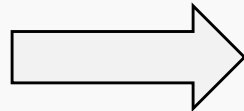
- онлайн-платформы (VLOPs);
- поисковые системы (VLOSEs);
- онлайн-маркетплейсы;
- поставщиков услуг хостинга;
- и т.д.

Ст. 26 (1)(d)



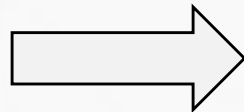
VLOPs и VLOSEs обязаны раскрывать в своих условиях политику, процедуры, меры и инструменты, применяемые для генерации контента, маркировать сгенерированный контент

Ст. 34 (1)(b)



VLOPs обязаны проводить оценку рисков, связанных с их сервисами, включая риск распространения незаконного контента, а также дипфейков, способных причинить вред здоровью граждан, гражданским правам, общественным интересам

Ст. 36



В условиях кризиса (вооруженный конфликт, терроризм, стихийное бедствие) VLOPs и VLOSEs обязаны принять меры по действию распространению ложной или вводящей в заблуждении информации, включая дипфейки



# Соединенные Штаты Америки



Ключевая цель регулирования генеративного ИИ в США - снятие правовых барьеров развития технологии для достижения технологического лидерства США в мировом рынке.

Законодательство США направлено на урегулирование вопросов применения генеративного ИИ только в определенных сферах. В первую очередь, в сфере правоохранительной деятельности.

Система регулирования генеративного ИИ представлена:

- федеральным законодательством;
- законодательством Штатов (причем более детализированное в сравнении с федеральным)

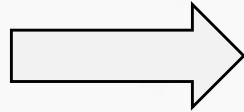


# Соединенные Штаты Америки



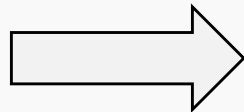
Регулирование ИИ в целях борьбы с распространением интимных изображений или дипфейков интимного содержания (без согласия лица).

**Закон TAKE It Down (19.05.2025)**



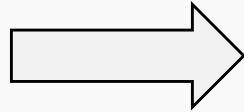
Онлайн-платформы обязаны удалить дипфейк по требованию пострадавшего лица в течении 48-ми часов. Лица, распространяющие подобные материалы, несут уголовную ответственность

**Закон AB 1831 (штат Калифорния)**



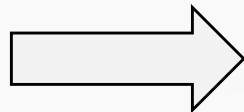
Распространяет понятие порнографии на материалы, которые были изменены в цифровом виде или сгенерированы с помощью ИИ. Направлен на борьбу с детской порнографией

**Закон SB 926 (штат Калифорния)**



Расширяет уголовное преследование за распространение не только реальных интимных фото (видео), но и дипфейков интимной тематики

**Закон SB 981 (штат Калифорния)**



Обязывает социальные сети бороться с «кражей цифровых персональных данных сексуального характера». Обязывает предоставлять пользователям понятный способ сообщить о контенте, который является поддельным интимным материалом

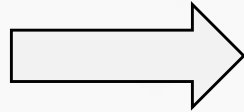


# Соединенные Штаты Америки



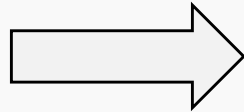
Регулирование ИИ в штате Калифорния в сфере борьбы с дипфейками.

**Закон АВ 2655 (штат Калифорния)**



Обязывает крупные онлайн-платформы (более 1 млн. пользователей) выявлять и удалять (либо маркировать) «существенно обманчивый контент» в периоды, близкие к выборам (120 дней до выборов, 60 дней после выборов). Под обманчивым контентом понимаются дипфейки, изображающие кандидата

**Закон АВ 2355 (штат Калифорния)**



Обязывает снабжать четким предупреждением всю политическую рекламу, созданную или измененную с помощью ИИ

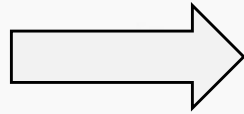


# Соединенные Штаты Америки



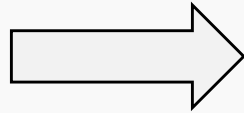
## Регулирование ИИ в штате Калифорния в сфере медиаиндустрии.

**Закон АВ 2602 (штат Калифорния)**



Вносит в Трудовой кодекс требование об обязательном получении студиями разрешения от актера в случае создания цифровой копии его голоса или образа (создание аудио-, фото-, видео-дипфейков)

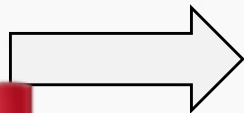
**Закон АВ 1836 (штат Калифорния)**



Вводит запрет на использование цифровых копий умерших артистов без получения явного согласия их правопреемников

# Подводя итоги

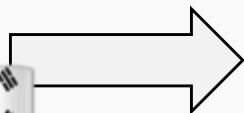
Китайская народная республика



Превентивный государственный контроль. Основные обязанности возложены на разработчиков и поставщиков услуг. Ключевой элемент регулирования - обязательная маркировка и аутентификация пользователей.

**Приоритет регулирования - национальная безопасность**

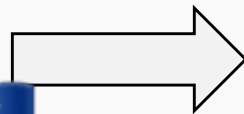
Южная Корея



Использование оценки рисков, общее регулирование ИИ-систем с наличием специализированных норм, запрещающих использование дипфейков в определенных случаях (например, выборы).

**Приоритет регулирования - защита прав и интересов граждан**

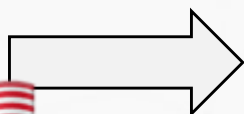
Европейский союз



Обязательства по прозрачности и маркировке систем ИИ. Возложение обязанностей на провайдеров по оценке рисков и противодействию распространению вредоносного ИИ-контента.

**Приоритет регулирования - защита прав и интересов граждан**

Соединенные Штаты Америки



Узконаправленное регулирование ИИ в определенных сферах. Основное регулирование представлено на уровнях штатов. **Приоритет регулирования - поддержка инноваций и технологического лидерства при минимальных законодательных запретах.**

# Иной зарубежный опыт

## Великобритания

Великобритании введена в действие норма, устанавливающая уголовную ответственность за создание или заказ создания интимного изображения взрослого лица без согласия (в том числе с использованием технологий генерации изображения). Эта мера введена через **Data (Use and Access) Act 2025 (раздел 138)**, которым внесены изменения в **Sexual Offences Act 2003**

## Казахстан

Принят и действует рамочный закон об искусственном интеллекте, устанавливающий обязательную маркировку (ст. 21); в правоприменительной и экспертной повестке он рассматривается как основание для обязательной маркировки синтетического контента (изображения, видео, аудио, текст) как меры против злоупотреблений и дипфейков.

**Закон Республики Казахстан от 17 ноября 2025 года № 230-VIII ЗРК «Об искусственном интеллекте»**

# Перечень предложений для регулирования



ЗАРУБЕЖНАЯ ПРАКТИКА	ПРЕДЛОЖЕНИЯ ПО РЕГУЛИРОВАНИЮ В РОССИЙСКОЙ ФЕДЕРАЦИИ
<b>КНР: разделение обязанностей провайдера, пользователя и платформы</b>	Закрепить модель: провайдер маркирует контент; пользователь не удаляет маркировку и раскрывает ИИ-происхождение при публикации; платформа обеспечивает отображение маркировки и реагирует на жалобы
<b>КНР: явная и скрытая маркировка</b>	Ввести две формы маркировки: видимую для человека и машиночитаемую в метаданных файла. Видимая маркировка обязательна для контента, способного ввести в заблуждение
<b>КНР: запрет удаления и подделки маркировки</b>	Установить ответственность за удаление, сокрытие, подмену или обход маркировки ИИ-контента, включая распространение инструментов для такого обхода
<b>США / Калифорния / Южная Корея: особый режим для выборов</b>	Ввести обязательную маркировку при использовании ИИ в агитации, хранение согласий лиц, чьи образ или голос использованы, и ускоренное удаление вводящих в заблуждение дипфейков в избирательный период
<b>КНР / Калифорния: защита лица, голоса и цифрового образа</b>	Закрепить необходимость отдельного согласия на создание и распространение дипфейка с лицом, голосом, мимикой или иными индивидуализирующими признаками человека

# СПАСИБО ЗА ВНИМАНИЕ!

## НАШИ КОНТАКТЫ



**Директор Института:**  
**Никишин Владимир Дмитриевич**

☎ Телефон: 8 (499) 244-88-88 доб. 2521

✉ Электронная почта: [vdnikishin@msal.ru](mailto:vdnikishin@msal.ru)



Институт в MAX



Институт  
во ВКонтакте



Институт в Телеграм