

Российский закон об ИИ, «национальные» модели и проверка на традиционные ценности: что уже принято, кто будет проверять и насколько система открыта

Резюме

По состоянию на **13 августа 2026 года** Россия действительно приняла федеральный закон, который вводит юридические статусы **«суверенной»** и **«национальной»** **большой фундаментальной модели ИИ** и прямо требует для получения обоих статусов пройти **«подтверждение соответствия законодательству Российской Федерации и традиционным российским духовно-нравственным ценностям»**. Это Федеральный закон от **26 июля 2026 года № 243-ФЗ**. Однако называть существующую систему уже работающей «сертификацией ИИ» пока преждевременно: основной закон вступает в силу 1 сентября 2026 года, а нормы о статусах и проверке соответствия — только **1 марта 2027 года**. Саму процедуру проверки закон не описывает ни технически, ни организационно: ее должно отдельно установить Правительство. ¹

Самое существенное на сегодня: **критерий есть, теста нет**. Закон перечисляет «традиционные ценности», но не сообщает, как переводить «патриотизм», «крепкую семью», «приоритет духовного над материальным», «историческую память» и другие нормативные категории в измеряемый pass/fail для вероятностной языковой модели. Не определены публично ни набор вопросов, ни rubric, ни количество прогонов, ни параметры inference, ни правила отказов, ни допустимая вариативность ответа, ни процедура апелляции. Правительству предстоит принять пять постановлений; подготовить их Минцифры поручено **до 1 ноября 2026 года**, включая порядок подтверждения соответствия моделям законодательству и традиционным ценностям. ²

При этом **я не нашёл официального документа, в котором было бы сказано, что проверка традиционных ценностей будет именно benchmark'ом вроде MERA**. Не нашёл и акта, назначающего MERA, Альянс в сфере ИИ, Сбер, МТС, AIRI или кого-либо ещё оператором такой проверки. Это важная граница между тем, что уже доказано, и тем, чего пока можно только опасаться. ³

Но опасение конфликта интересов не высосано из пальца. **Альянс в сфере ИИ — объединение крупнейших компаний, одновременно занимающееся нормативным регулированием, этикой, технической стандартизацией и benchmark'ами**. Его наблюдательный совет возглавляет первый зампред Сбербанка Александр Ведяхин; Национальную комиссию по этике ИИ возглавляет Андрей Незнамов, которого публичные источники одновременно идентифицируют как управляющего директора Центра регулирования ИИ Сбербанка и руководителя рабочей группы Альянса по нормативному регулированию. MERA — проект самого Альянса; репозиторий прямо перечисляет среди партнеров SberDevices/Sber AI, Yandex, Skoltech AI, MTS AI, НИУ ВШЭ, РАН и др. ⁴

Это **не доказательство подкрутки или будущего захвата процедуры сертификации**. Но структура потенциального конфликта очевидна: статус «национальной/суверенной модели» открывает государственную поддержку и может стать пропуском в сферы, где Правительство разрешит использовать только такие модели; одновременно крупнейшие потенциальные получатели этого статуса уже много лет участвуют в разработке российской AI-этики, нормативных подходов и систем оценки. ⁵

Отдельно важна формулировка закона о **«национальной» модели**. Она вовсе не обязана быть обучена с нуля на полностью российской архитектуре: закон прямо разрешает компоненты, разработанные за рубежом, **включая другие большие фундаментальные модели**, если они доступны на условиях открытой лицензии и существенные характеристики конечной модели определяет российский разработчик. То есть fine-tune/дериватив открытой Qwen, GLM или другой иностранной модели теоретически способен претендовать именно на статус **национальной**, если удовлетворит остальным условиям. Для «суверенной» категории критерий существенно строже — требуется полный контроль цикла и его полная техническая и технологическая воспроизводимость. ¹

Главный риск сейчас не в том, что опубликован явно сфальсифицированный тест. **Главный риск — в огромной дискреции, оставленной будущему подзаконному акту**. Через несколько месяцев одно постановление Правительства должно превратить политико-нормативный список ценностей в техническую процедуру, способную решить, получает ли модель привилегированный юридический статус. А публичной спецификации этой процедуры на сегодняшний день нет. ²

Что именно Россия уже приняла

Финальный закон принципиально уже не тот, что обсуждался весной

В марте 2026 года Минцифры вынесло на общественное обсуждение гораздо более широкий проект **«Об основах государственного регулирования сфер применения технологий искусственного интеллекта в Российской Федерации»**. В нем предполагались категории суверенной, национальной и «доверенной» модели, риск-ориентированное регулирование, маркировка контента и значительно более широкий охват AI-систем. Проект публиковался на федеральном портале проектов нормативных актов. ⁶

Карточка мартовского проекта: regulation.gov.ru — проект ID 166424

К июню Правительство подготовило **существенно переписанный вариант**. Он стал короче — 13 статей — и сфокусировался уже не на регулировании всего ИИ, а прежде всего на поддержке **больших фундаментальных моделей с не менее чем 1 млрд параметров**. Категория «доверенной модели» из финального текста исчезла. ⁷

25 июня Правительство внесло законопроект № 1271570-8: [карточка Государственной думы](#). 8 июля он был принят Госдумой в окончательной редакции, 17 июля одобрен Советом Федерации, 26 июля подписан президентом как № 243-ФЗ. ⁸

Официальное опубликование: publication.pravo.gov.ru

Полный текст: [Российская газета](#). ¹

«Традиционные ценности» — юридически обязательное условие статуса

Статья 4 закона устанавливает общий принцип «учета и уважения традиционных российских духовно-нравственных ценностей». Перечень практически дословно воспроизводит президентский Указ № 809: жизнь, достоинство, права и свободы, патриотизм, гражданственность, служение Отечеству, крепкая семья, созидательный труд, духовное выше материального, гуманизм, милосердие, справедливость, коллективизм, взаимопомощь, историческая память, преемственность поколений и единство народов России. ⁹

Первичный источник этого списка — [Указ Президента № 809 от 9 ноября 2022 года](#), обновленный в том числе Указом № 136 от 4 марта 2026 года. В нем «традиционные ценности» являются элементом политики национальной безопасности, а распространение «деструктивной идеологии» рассматривается как угроза; документ также предусматривает участие религиозных организаций в этой политике. ¹⁰

Но в 243-ФЗ это уже не просто декларативный принцип. Для **суверенной модели** статья 6 требует:

модель должна пройти подтверждение соответствия законодательству РФ и традиционным российским духовно-нравственным ценностям в порядке, установленном Правительством.

Ровно такое же условие установлено для **национальной модели**. ¹

При этом закон **не определяет**, что считается прохождением такого подтверждения. Нет слов о benchmark score, экзаменационной комиссии, экспертном заключении, red-teaming, количестве запросов или конкретном проценте правильных ответов. Всё это вынесено на уровень будущего постановления Правительства. ³

Что получает модель за статус

Статус не является обязательной лицензией на существование LLM в России. Закон не говорит: «ChatGPT запрещен, пока не пройдет традиционные ценности». Формально подтверждение требуется для получения **национального или суверенного статуса**. ¹

Однако последствия статуса потенциально гораздо существеннее красивого бейджа. Разработчики таких моделей получают доступ к государственным мерам поддержки, а Правительство получает право определить **случаи, когда допускается использование исключительно национальных или суверенных моделей**. Для финансового рынка такие ограничения согласуются с Банком России. ¹

То есть в законе уже находится механизм потенциального gatekeeping:

прошёл статус → доступен для льгот/поддержки + может попасть в разрешённый круг для отдельных государственных/чувствительных применений

не прошёл → продолжать существовать в коммерческом интернете можешь, но в выбранные государством сегменты можешь не попасть.

Но **всеобщей обязанности госзакупок покупать исключительно такие модели сегодня нет**. Конкретные случаи должен еще определить Кабмин; соответствующая норма начинает действовать с 1 марта 2027 года. Для уже эксплуатирующихся иностранные модели информационных систем предусмотрено переходное исключение до **1 сентября 2032 года**, если данные обрабатываются и хранятся в России. ¹

То же с субсидиями: 243-ФЗ дает разработчику модели со статусом право получать меры господдержки, но **не устанавливает универсальное правило «любая субсидия только после сертификации»**. Конкретные программы и условия еще должны быть определены. ¹

Кто реально отвечает за систему

Ключевое различие: кто **имеет полномочия по закону**, кто **готовит нормативку**, а кто всего лишь **влиятельный участник отраслевого обсуждения**.

Орган / человек	Фактическая роль	Что известно о связи с проверкой	Публичная аффилиация / биография
Правительство РФ	Главный нормативный оператор	Устанавливает процедуру проверки ценностей, процедуру присвоения статусов, эксклюзивные случаи применения российских моделей, требования к рискам	243-ФЗ ¹
Минцифры / Максут Шадаев	Профильный федеральный регулятор IT	Минцифры поручено подготовить пять правительственных постановлений до 1 ноября 2026 года	Шадаев — действующий министр; официальная биография ¹¹
Дмитрий Григоренко	Зампред Правительства, руководитель Аппарата; один из главных государственных кураторов ИИ	Сопредседатель президентской Комиссии по ИИ; правительственная AI-вертикаль работает под его кураторством	биография Правительства ¹²

Орган / человек	Фактическая роль	Что известно о связи с проверкой	Публичная аффилиация / биография
Максим Орешкин	Замруководителя Администрации Президента	Сопредседатель президентской Комиссии по развитию ИИ	Комиссия Кремля ¹³
Александр Бортников / ФСБ	Директор ФСБ; член президентской AI-комиссии	Закон предусматривает согласование доступа к государственным данным с «федеральным органом в области обеспечения безопасности», но не назначает ФСБ certifier'ом традиционных ценностей	состав президентской комиссии ¹⁴
Банк России	Отраслевой регулятор финансов	Согласует эксклюзивность национальных/суверенных моделей и риск-требования в финансовой сфере	Ст. 5 243-ФЗ ¹
Александр Бабаков	Зампред Госдумы	Руководил межфракционной рабочей группой по законодательству об ИИ; группа в марте одобрила свою концепцию и рекомендовала ее учитывать Правительству	Госдума, 18.03.2026 ¹⁵
Вадим Виноградов	Декан факультета права НИУ ВШЭ	На заседании рабочей группы Госдумы представлял структуру и концепцию разрабатываемого регулирования	Та же публикация Госдумы ¹⁵
Андрей Незнамов	Глава Национальной комиссии по этике ИИ	В 2026 году публично выступает как глава этической комиссии; ранее публично идентифицирован как управляющий директор Центра регулирования ИИ Сбербанка и глава рабочей группы Альянса по регулированию	пресс-центр РИА, 2026, Коммерсантъ ¹⁶

Орган / человек	Фактическая роль	Что известно о связи с проверкой	Публичная аффилиация / биография
Александр Ведяхин	Первый зампред Сбербанка; председатель наблюдательного совета Альянса в сфере ИИ	Не государственный certifier; публично призывал обсуждать, как «базовые ценности» транслируются нейросетями, и повышать роль комиссии по этике ИИ	Альянс , заявление на ПМЭФ ¹⁷
Валерия Воробьева	Директор Альянса в сфере ИИ	Руководитель отраслевого объединения, которое занимается нормативным регулированием, технической стандартизацией, этикой и benchmark'ами	Альянс ¹⁸

Есть еще Александр Шойтов: **13 апреля 2026 года, будучи заместителем министра цифрового развития**, он публично заявил, что, в отличие от гражданина, ИИ должен подвергаться цензуре и что на его «входе и выходе» должны находиться ограничения на определенные запросы. Это довольно прямо характеризует подход части ведомства к alignment/regulation. ¹⁹

Но называть Шойтова уверенно «нынешним ответственным за сертификацию» нельзя. Во-первых, 243-ФЗ такой функции ему не дает. Во-вторых, в мае РБК сообщал о его возможном уходе из Минцифры и о реорганизации, в рамках которой в том числе планировалось объединение департамента облачных сервисов с департаментом развития ИИ и больших данных. Окончательное увольнение Шойтова в доступных мне источниках подтверждено не было. ²⁰

А где Роскомнадзор, ФСТЭК и Минпромторг?

В финальном 243-ФЗ Роскомнадзор, ФСТЭК и Минпромторг не названы органами, подтверждающими соответствие национальной/суверенной модели традиционным ценностям. Поэтому сейчас писать «сертификат будет выдавать ФСТЭК» или «решать будет Роскомнадзор» — спекуляция. ¹

ФСТЭК и ФСБ присутствовали в **параллельном контуре информационной безопасности**. В феврале 2026 года РБК описывал прорабатываемую схему полигона, где AI-системам высокого/критического риска для КИИ могли потребоваться требования ФСТЭК и ФСБ; при этом представитель самого Минцифры сообщал РБК, что ведомство в тот момент не разрабатывало законопроект именно о таком полигоне. Это нельзя автоматически переносить на финальную процедуру статуса из 243-ФЗ. ²¹

Иными словами, сейчас существуют как минимум два разных вопроса:

security conformity для чувствительных систем;

legal + “traditional values” conformity для статуса национальной/суверенной BFM.

Подзаконные акты еще могут их связать, но федеральный закон пока этого не сделал. 22

Как к этому пришли и когда система заработает



Текущая вертикаль была институционально усилена еще до принятия закона. 26 февраля 2026 года создана **Комиссия при Президенте по вопросам развития технологий ИИ**, предназначенная для координации федеральных органов, Банка России, регионов и заинтересованных организаций; сопредседатели — Дмитрий Григоренко и Максим Орешкин. В ее состав входит, среди прочих, директор ФСБ Александр Бортников. 23

6 марта Правительство распоряжением № 435-р создало собственную **подкомиссию по развитию и внедрению ИИ**, предназначенную для оперативной межведомственной координации и взаимодействия с президентской комиссией. [Официальное сообщение Правительства](#). 24

Нормативный фундамент старше. Обновленная [Национальная стратегия развития ИИ](#), утвержденная изменениями Указа № 124 от 15 февраля 2024 года, прямо требует развивать нормативное регулирование и распространять этические нормы. В ней отдельно закреплено стимулирование внедрения этических стандартов, основанных на **Кодексе этики в сфере ИИ**. 25

Иными словами, корпоративно-общественный Кодекс этики был встроен в государственную AI-стратегию **до** нынешнего закона о национальных моделях. Это важно для понимания того, почему Альянс в сфере ИИ и его этические структуры нельзя рассматривать просто как сторонний Telegram-кружок бизнеса. Но юридические полномочия по выдаче статуса им пока не переданы. 26

По данным РАПСИ, для исполнения 243-ФЗ нужно **пять постановлений Правительства**, которые Минцифры должно подготовить до 1 ноября 2026 года. В частности, они должны определить случаи исключительного применения отечественных моделей, правила доступа к государственным данным, процедуру подтверждения соответствия закону и традиционным ценностям и порядок присвоения статусов. 27

На дату этого отчета — **13 августа 2026 года** — финальную опубликованную методику такой проверки я в открытых нормативных источниках не нашел. До установленного срока ее подготовки остается примерно два с половиной месяца. Поэтому любые рассказы о точном количестве

вопросов, «секретном benchmark», проходном балле или конкретной комиссии пока не подтверждены публичным нормативным актом. ²

Корпоративная этика, религия, Альянс и MERA

Здесь начинается наиболее интересная институциональная часть.

Национальная комиссия по этике ИИ — не государственный орган

Национальная комиссия по этике в сфере ИИ возникла как структура, сформированная участниками отрасли для реализации Кодекса этики. Уже в 2023 году «Коммерсантъ» описывал ее председателя Андрея Незнамова одновременно как **управляющего директора Центра регулирования ИИ Сбербанка** и руководителя рабочей группы «Нормативное регулирование» Альянса в сфере ИИ. В январе 2026 года Незнамов по-прежнему публично выступал как руководитель Национальной комиссии. ²⁸

В публичных сообщениях о составе комиссии фигурировали представители **Сбера, Яндекса, МТС, Иннополиса, фонда «Сколково», Сколтеха, ИТМО, МГТУ им. Баумана и других организаций**. То есть это принципиально multi-stakeholder industry/academic body, а не независимый государственный надзорный орган. ²⁹

Сам [Альянс в сфере ИИ](#) открыто указывает среди направлений своей работы **нормативное регулирование**, «разработку концепций регулирования ИИ», **этику применения и техническую стандартизацию ИИ**. Председатель его наблюдательного совета — первый зампред Сбербанка Александр Ведяхин. В публичном перечне/мероприятиях Альянса представлены крупнейшие компании российского рынка — Сбер, МТС, VK, Ростелеком, Яндекс, Альфа-Банк, Газпром нефть и другие. ³⁰

В 2026 году этический контур напрямую соединился с религиозным

30 октября 2025 года на базе **Синодальной комиссии по биоэтике РПЦ и Альянса в сфере ИИ** состоялось организационное заседание будущего Консультативного совета по этике ИИ. 26 января 2026 года соглашение о его создании было оформлено официально. ³¹

14 мая 2026 года прошло первое заседание уже созданного совета. Одной из центральных тем стала новая редакция **«Белой книги этики в сфере ИИ»**, в которую должен войти раздел о взглядах **традиционных религий** на вопросы искусственного интеллекта. ³²

Это особенно примечательно в контексте 243-ФЗ, поскольку Указ № 809 прямо говорит, что христианство, ислам, буддизм и иудаизм повлияли на формирование традиционных ценностей, а православие приписывает «особую роль». ¹⁰

Еще более прямой сигнал прозвучал на ПМЭФ в июне 2026 года: первый зампред Сбера Александр Ведяхин публично предложил поднять вопрос о том, **как «базовые ценности» транслируются через нейросети**, и одновременно говорил о необходимости повышения роли комиссии по этике

ИИ. Это не нормативный акт и не означает, что Сберу передали сертификацию. Но тематическое совпадение с принятой месяц спустя юридической конструкцией очевидно. ³³

Где здесь MERA

MERA — официальный проект экосистемы Альянса. Его нынешний GitHub описывает цель проекта как создание **«independent unified system»** измерения моделей, обещает одинаковые сценарии, метрики и adaptation strategy, «controlled and clear comparisons» и «fair and transparent leaderboards». Репозиторий благодарит SberDevices, Sber AI, Yandex, Skoltech AI, MTS AI, НИУ ВШЭ, РАН и других участников; опубликованная ACL-работа перечисляет команду авторов из индустрии и академии. [GitHub](#), [ACL paper](#). ³⁴

Сам сайт позиционирует MERA как **«независимый бенчмарк»**, причем в его батарее есть отдельные категории Ethics: `ruDetox`, `ruEthics`, `ruHateSpeech`, `ruHHH`. ³⁵

Но я не нашел ни в 243-ФЗ, ни в сообщении о пяти будущих постановлениях, ни в официальных материалах правительственной/президентской комиссии указания:

«проверку традиционных ценностей осуществлять по MERA».

Поэтому корректная формулировка сейчас такая:

MERA — очень релевантный российский прецедент того, как Альянс и крупные AI-компании создают централизованную систему оценки моделей. Но доказательств, что MERA станет юридическим экзаменом для национального статуса, пока нет. ³⁶

Насколько будущая проверка может быть открытой и воспроизводимой

Государственная методология пока представляет собой черный ящик — потому что ее еще не опубликовали

Сам федеральный закон задает семантическое требование, но не measurement protocol.

Возьмем одну ценность:

«патриотизм».

Что должна сделать модель?

Ответить «да» на вопрос «патриотизм — это хорошо»?

Отказать на просьбу критиковать государство?

Правильно изложить историю?

Как отличить законную критику органов власти от «непатриотичного» ответа?

И как вообще проверить, что миллиардопараметрическая модель **«соответствует»** ценности во всем пространстве возможных prompts?

Закон ничего этого не определяет. Аналогичная проблема возникает с «крепкой семьей», «коллективизмом», «исторической памятью» и «приоритетом духовного над материальным». ⁹

В техническом смысле возможны совершенно разные процедуры:

fixed benchmark → pass/fail

red-team prompts → экспертная комиссия

LLM-as-a-judge

ручная экспертная разметка

policy/document audit + model testing

комбинация всех методов.

С точки зрения юридической предсказуемости это совершенно разные режимы. 243-ФЗ пока оставляет выбор Кабмину. ¹

И здесь MERA действительно является полезным предупреждением

В предоставленной тобой полной транскрипции сообщества MERA есть задокументированные обсуждения того, что разработчики признают зависимость результатов от **vLLM и precision**; пользователь при попытке воспроизвести официальный Llama-3-8B-Instruct получал 0.430 против 0.414 overall и 0.641 против 0.608 на MaMuRAMu даже после подбора chat template, multi-turn, system prompt и версии кода. [filecite turn0file0](#)

Там же разработчики признают проблемы с корректной оценкой генераций в code-задачах: модель может не понять ожидаемый format/continuation contract и поэтому дать ответ, который evaluator оценивает некорректно, даже когда проблема отчасти относится к форме генерации, а не к способности программировать. [filecite turn0file0](#)

Особенно показателен эпизод, где участник предлагает удалить проблемные `ruCode/HumanEval`, а представитель проекта отвечает, что удаление действительно планировалось, но оно переставит модели на leaderboard, потребует пересчета totals и сделает прежние ссылки на результаты в СМИ неактуальными, поэтому предпочтительно сохранить старую версию и позднее выпустить новый board. [filecite turn0file0](#)

Это **не доказывает недобросовестность MERA**, тем более что в последующих версиях code-бенчмарка часть методологии менялась и исправлялась. Это показывает другую проблему: даже обычный benchmark математических и кодовых навыков быстро превращается в функцию

$$score = f(model, prompt, chat\ template, tokenizer, backend, precision, parser, generation\ settings, \dots)$$

а теперь представим то же самое для значительно менее операционализируемого понятия **«соответствие традиционным ценностям»**.

Закрытые тесты уже используются, но это не доказательство секретной сертификации

У MERA действительно есть закрытая часть оценки. В сообществе разработчики прямо объясняют, что для закрытых задач не предоставляется публичный scoring API, чтобы предотвратить утечку теста; в некоторых ветках вопросы доступны, а правильные ответы скрыты и scoring делается на серверной стороне. `filecite\turn0file0`

Само по себе это **нормальная benchmark-практика**. Если открыть все target labels, модели или их разработчики быстро оптимизируются под test set.

Проблема возникает, если такой принцип перенести из научного leaderboard в **юридически значимый допуск**, не дав одновременно:

- публичной спецификации измеряемых критериев;
- фиксированной версии protocol;
- независимого аудита;
- повторного прогона;
- статистических допусков;
- процедуры оспаривания конкретного failure;
- гарантий одинаковой среды для всех претендентов.

На 13 августа опубликованного регламента с такими гарантиями я не нашел. 2

И еще одна важная вещь: **доказательств существования уже готового секретного государственного test set на традиционные ценности я тоже не нашел**. Отсутствие опубликованной методики сегодня означает «неизвестно», а не автоматически «есть тайный список вопросов, по которому пропустят нужных».

Конфликт интересов и институциональные риски

Здесь полезно разделить реальный конфликт интересов от конспирологии.

Реальный структурный конфликт существует. Закон способен создать экономическую ценность статуса: государственная поддержка плюс возможность эксклюзивного допуска в определенные Правительством сферы. При этом российские разработчики фундаментальных моделей входят в объединение, которое официально занимается выработкой нормативных предложений, этикой и

технической стандартизацией. Председателем наблюдательного совета этого объединения является топ-менеджер Сбербанка. ⁵

Более того, сам 243-ФЗ прямо закрепляет право разработчиков суверенных и национальных моделей участвовать в **профильных отраслевых организациях, разработке этических правил и подготовке предложений по совершенствованию государственной политики в сфере ИИ**. То есть участие индустрии в rulemaking не случайный побочный эффект — оно прямо предусмотрено моделью закона. ¹

Но **не доказано** другое: что Сбер, МТС или Яндекс получают право сами принимать решение о выдаче собственного статуса; что их сотрудники составят большинство комиссии; что MERA станет обязательным экзаменом; что тест будет намеренно оптимизирован под GigaChat/Cotype/BerryLM; либо что иностранную модель будут специально заваливать. Таких документов я не нашел. ³

При этом «национальная модель» по закону совершенно легально может содержать **иностранные open-source компоненты, включая другую foundation model**. Это значит, что юридический термин «национальная» не тождественен бытовому «создана полностью с нуля в России». Именно «суверенная» категория ближе к такому требованию контроля и полной воспроизводимости. ¹

Это существенная деталь для публичной дискуссии. Будущая новость:

«Российская национальная модель получила государственный статус»

по закону **не обязательно означает**

«архитектура, pretraining и все веса разработаны в России с чистого листа».

Закон специально допускает другое. ¹

Еще один институциональный сигнал — направление этической дискуссии. Президентская Стратегия предписывает внедрять стандарты Кодекса этики; отраслевую комиссию по этому Кодексу возглавляет человек с документированной связью со Сбером и Альянсом; в 2025–2026 годах создан совместный консультативный орган Альянса и Синодальной комиссии РПЦ; а незадолго до принятия закона первый зампред Сбера публично поднимал вопрос о трансляции «базовых ценностей» через нейросети. Все отдельные звенья этой цепочки публичны. ³⁷

Но из них пока можно доказательно вывести лишь:

крупный российский AI-бизнес обладает заметным институциональным доступом к формированию этической и нормативной повестки.

Нельзя доказательно перескочить к:

«поэтому компании уже написали себе государственный экзамен и сами себя будут сертифицировать».

Для второго утверждения нужны будущие пять постановлений и, возможно, документы о составе комиссии/оператора оценки. ²⁷

Оценка рисков

Риск	Оценка на 13.08.2026	Почему
Непрозрачность operational criteria	Высокий	Федеральный закон требует соответствия ценностям, но методики пока нет; всё вынесено в постановление Правительства. ³
Политическая субъективность критерия	Высокий по конструкции	Значительная часть проверяемых понятий — нормативные категории: патриотизм, крепкая семья, коллективизм, историческая память и т. д.; Указ №809 связывает их с национальной безопасностью и «деструктивной идеологией». ¹⁰
Конфликт интересов индустрии	Средне-высокий потенциальный	Альянс объединяет потенциальных получателей статуса и одновременно занимается регулированием, этикой, стандартизацией и benchmark'ами; Сбер имеет заметную руководящую роль. Но формальным certifier'ом Альянс пока не назначен. ³⁸
Техническая валидность pass/fail	Неизвестно; потенциально высокий риск	Методика отсутствует. MERA показывает, насколько обычная оценка LLM чувствительна к inference stack и parser'y; для ценностных judgments проблема потенциально сложнее. ³⁴ filecite turn0file0
Секретный test set	Не доказан	У MERA есть закрытые targets, но опубликованного государственного теста — открытого или закрытого — пока нет.
Произвольный допуск на рынок	Средний сейчас, потенциально высокий после подзаконных актов	Общего запрета иностранных LLM нет, но Кабмин сможет определить случаи, где разрешены только national/sovereign models. ¹
Обязательность для всех госзакупок	Нет на сегодня	Закон не устанавливает blanket procurement rule; конкретные случаи должен определить Кабмин. ¹
Связка статуса с господдержкой	Да	Разработчик модели со статусом получает право доступа к соответствующим мерам поддержки. ¹

Риск	Оценка на 13.08.2026	Почему
Внешняя независимая апелляция	Не определена	243-ФЗ не описывает апелляционный механизм именно для результата value-compliance assessment. ¹
Роскомнадзор/ ФСТЭК как уже назначенный экзаменатор	Нет доказательств	В финальном законе они не назначены оператором проверки ценностей; security-контур ФСТЭК/ФСБ следует рассматривать отдельно. ³⁹

Моя итоговая оценка поэтому довольно специфична.

Оснований сейчас писать «Россия уже создала закрытый benchmark, в котором Сбер будет сертифицировать сам себя» — нет.

Зато уже есть основания написать гораздо менее сенсационную, но юридически намного более неприятную вещь:

Россия приняла закон, по которому экономически значимый статус большой модели будет зависеть в том числе от подтверждения ее соответствия политико-нормативному набору «традиционных ценностей», **при этом способ превратить эти ценности в воспроизводимый технический критерий федеральный закон вообще не определяет.**

И одновременно методологию предстоит разрабатывать внутри экосистемы, где государство уже годами тесно взаимодействует с крупнейшими российскими AI-компаниями, а созданный ими Альянс занимается сразу **регулированием, этикой, стандартизацией и benchmark'ами.** ⁵

Именно поэтому решающий документ здесь — даже не 243-ФЗ. Он уже понятен.

Решающий документ — будущее постановление Правительства о порядке подтверждения соответствия.

По нему можно будет проверить пять вещей, которые сейчас остаются неизвестными:

кто именно экзаменует; опубликованы ли критерии; кто формирует тесты; воспроизводима ли процедура; и может ли разработчик оспорить результат.

Если в постановлении окажутся публичная версия протокола, независимое управление, frozen evaluation environment, аудит, повторы, статистические допуски и апелляция — часть рисков существенно снизится.

Если получится условное:

«перечень вопросов утверждается комиссией, не публикуется; результат определяется экспертным заключением; состав экспертной организации определяется оператором»

— тогда опасение о регуляторном benchmark'e как управляемом пропуске на рынок станет уже не гипотезой, а свойством самой нормативной конструкции.

Ключевые первичные и технические источники

Федеральный закон № 243-ФЗ от 26.07.2026:

[Официальный интернет-портал правовой информации](#)

[Полный текст — Российская газета](#) 1

Законопроект № 1271570-8:

[Система обеспечения законодательной деятельности Госдумы.](#) 40

Мартовский проект Минцифры «Об основах государственного регулирования...»:

[Федеральный портал проектов нормативных актов](#) — проект ID 166424; публикация и содержание мартовской версии подтверждаются материалами РБК. 6

Указ Президента № 809 о традиционных ценностях:

[Кремль.](#) 10

Обновленная Национальная стратегия развития ИИ, Указ № 124:

[Кремль.](#) 25

Президентская Комиссия по вопросам развития ИИ:

[Положение](#)

[Состав.](#) 41

Правительственная подкомиссия по ИИ, распоряжение № 435-р:

[Правительство РФ.](#) 24

Рабочая группа Госдумы под руководством Александра Бабакова:

[Госдума, 18 марта 2026.](#) 15

Срок и содержание пяти будущих постановлений:

[РАПСИ, 7 июля 2026.](#) 27

Альянс в сфере искусственного интеллекта:

[Официальный сайт.](#) 18

Консультативный совет по этике ИИ и традиционные религии:

[Первое заседание, Патриархия.ru](#)

[Создание совета.](#) 42

Национальная комиссия по этике ИИ / Андрей Незнамов:

[Пресс-центр РИА](#)

[«Коммерсантъ» — его связь со Сбером и рабочей группой Альянса.](#) 16

MERA:

[GitHub](#)

[официальный сайт](#)

[ACL 2024 paper](#). ⁴³

Публичная транскрипция сообщества MERA, предоставленная в этой беседе: содержит обсуждения воспроизводимости, code-parsing, закрытого scoring и решения не удалять сразу проблемные code-задачи из-за изменения leaderboard и прежних ссылок в СМИ.

filecite turn0file

¹ ³ ⁵ ⁹ ³⁶ ³⁹ <https://rg.ru/documents/2026/07/31/fz243-dok.html>
<https://rg.ru/documents/2026/07/31/fz243-dok.html>

² ²⁷ https://rapsinews.ru/legislation_news/20260707/312015015.html
https://rapsinews.ru/legislation_news/20260707/312015015.html

⁴ ¹⁷ ¹⁸ ³⁰ ³⁸ <https://a-ai.ru/>
<https://a-ai.ru/>

⁶ https://amp.rbc.ru/rbcnews/technology_and_media/18/03/2026/69baed539a7947daf36b8989
https://amp.rbc.ru/rbcnews/technology_and_media/18/03/2026/69baed539a7947daf36b8989

⁷ <https://o2consult.com/publications/novaya-redaktsiya-zakona-ob-ii-chto-izmenilos/>
<https://o2consult.com/publications/novaya-redaktsiya-zakona-ob-ii-chto-izmenilos/>

⁸ ⁴⁰ https://infocenter.cntd.ru/news/read/soviet-gd-napravil-v-rassylku-zakonoproekt-o-razvitii-texnologii-iskusstvennogo-intellekta-v-rossii?rubric_id=novosti-cifrovizacii
https://infocenter.cntd.ru/news/read/soviet-gd-napravil-v-rassylku-zakonoproekt-o-razvitii-texnologii-iskusstvennogo-intellekta-v-rossii?rubric_id=novosti-cifrovizacii

¹⁰ <https://www.kremlin.ru/acts/bank/48502/print>
<https://www.kremlin.ru/acts/bank/48502/print>

¹¹ <https://government.ru/gov/persons/627/>
<https://government.ru/gov/persons/627/>

¹² <https://government.ru/gov/persons/621/>
<https://government.ru/gov/persons/621/>

¹³ ²³ ⁴¹ <https://kremlin.ru/structure/regulation/1050>
<https://kremlin.ru/structure/regulation/1050>

¹⁴ <https://kremlin.ru/structure/members/1050>
<https://kremlin.ru/structure/members/1050>

¹⁵ <https://duma.gov.ru/news/63224/>
<https://duma.gov.ru/news/63224/>

¹⁶ <https://pressria.ru/20260122/958648284.html>
<https://pressria.ru/20260122/958648284.html>

- 19 <https://www.comnews.ru/content/244742/2026-04-13/2026-w16/1009/zamglavy-mincifry-shoytov-zayavil-o-neobkhodimosti-cenzury-ii>
<https://www.comnews.ru/content/244742/2026-04-13/2026-w16/1009/zamglavy-mincifry-shoytov-zayavil-o-neobkhodimosti-cenzury-ii>
- 20 https://nsk.rbc.ru/technology_and_media/06/05/2026/69f35cbe9a794723788e0840
https://nsk.rbc.ru/technology_and_media/06/05/2026/69f35cbe9a794723788e0840
- 21 22 https://www.rbc.ru/technology_and_media/10/02/2026/69886dd69a794722bccb17b0
https://www.rbc.ru/technology_and_media/10/02/2026/69886dd69a794722bccb17b0
- 24 <https://government.ru/news/58054/>
<https://government.ru/news/58054/>
- 25 26 37 <https://www.kremlin.ru/acts/bank/50326/print>
<https://www.kremlin.ru/acts/bank/50326/print>
- 28 <https://www.kommersant.ru/doc/5810784>
<https://www.kommersant.ru/doc/5810784>
- 29 <https://t.me/s/karaulny/457531>
<https://t.me/s/karaulny/457531>
- 31 <https://patriarchia.ru/article/118056>
<https://patriarchia.ru/article/118056>
- 32 42 <https://www.patriarchia.ru/article/121140>
<https://www.patriarchia.ru/article/121140>
- 33 <https://www.fontanka.ru/2026/06/04/76459591/>
<https://www.fontanka.ru/2026/06/04/76459591/>
- 34 43 <https://github.com/MERA-Evaluation/MERA>
<https://github.com/MERA-Evaluation/MERA>
- 35 <https://mera.a-ai.ru/ru/multi/about>
<https://mera.a-ai.ru/ru/multi/about>