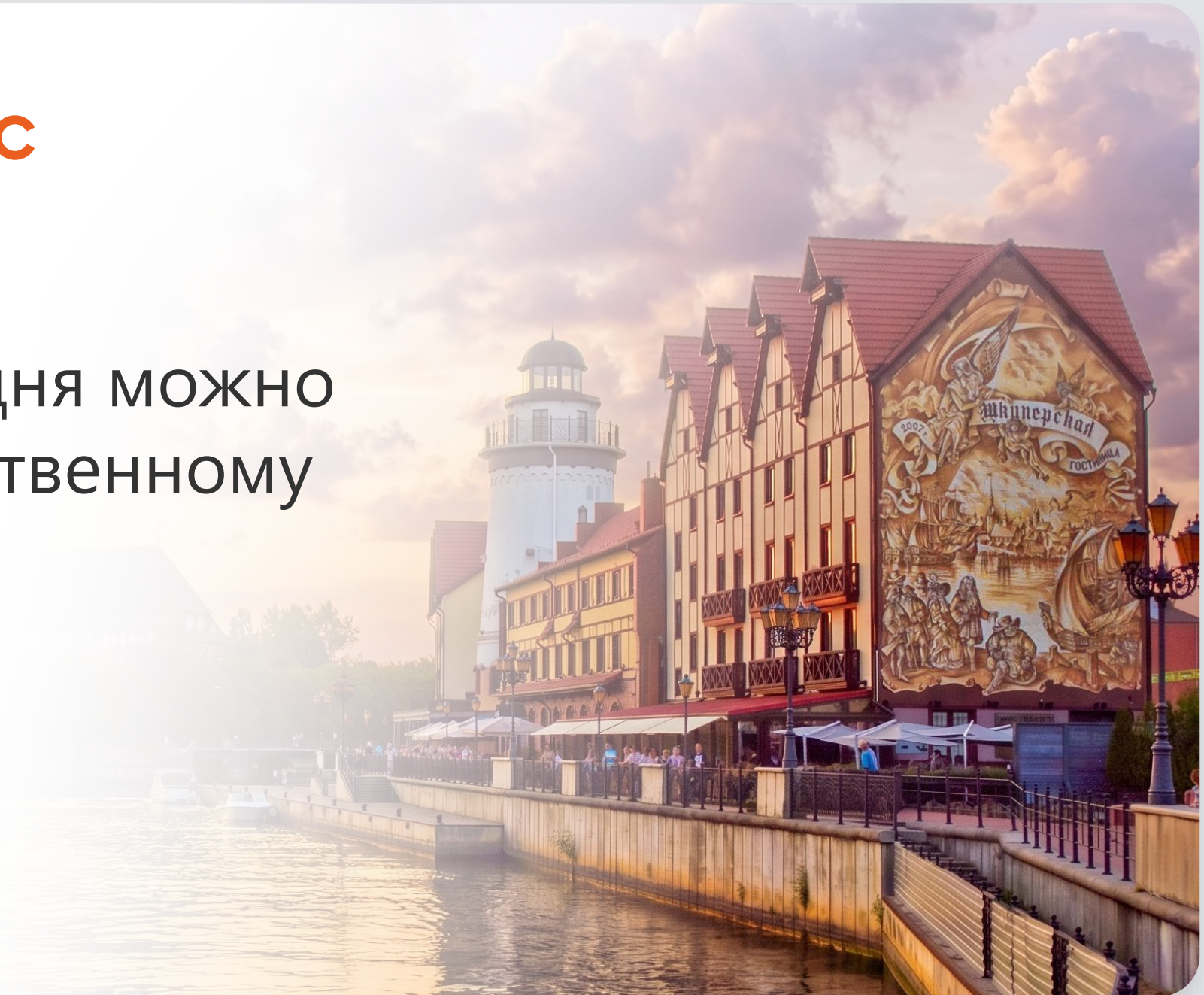




Насколько сегодня можно доверять искусственному интеллекту?

СЕМЕНОВ
Олег Станиславович

Заместитель генерального директора





~~«Быть или не быть?»~~

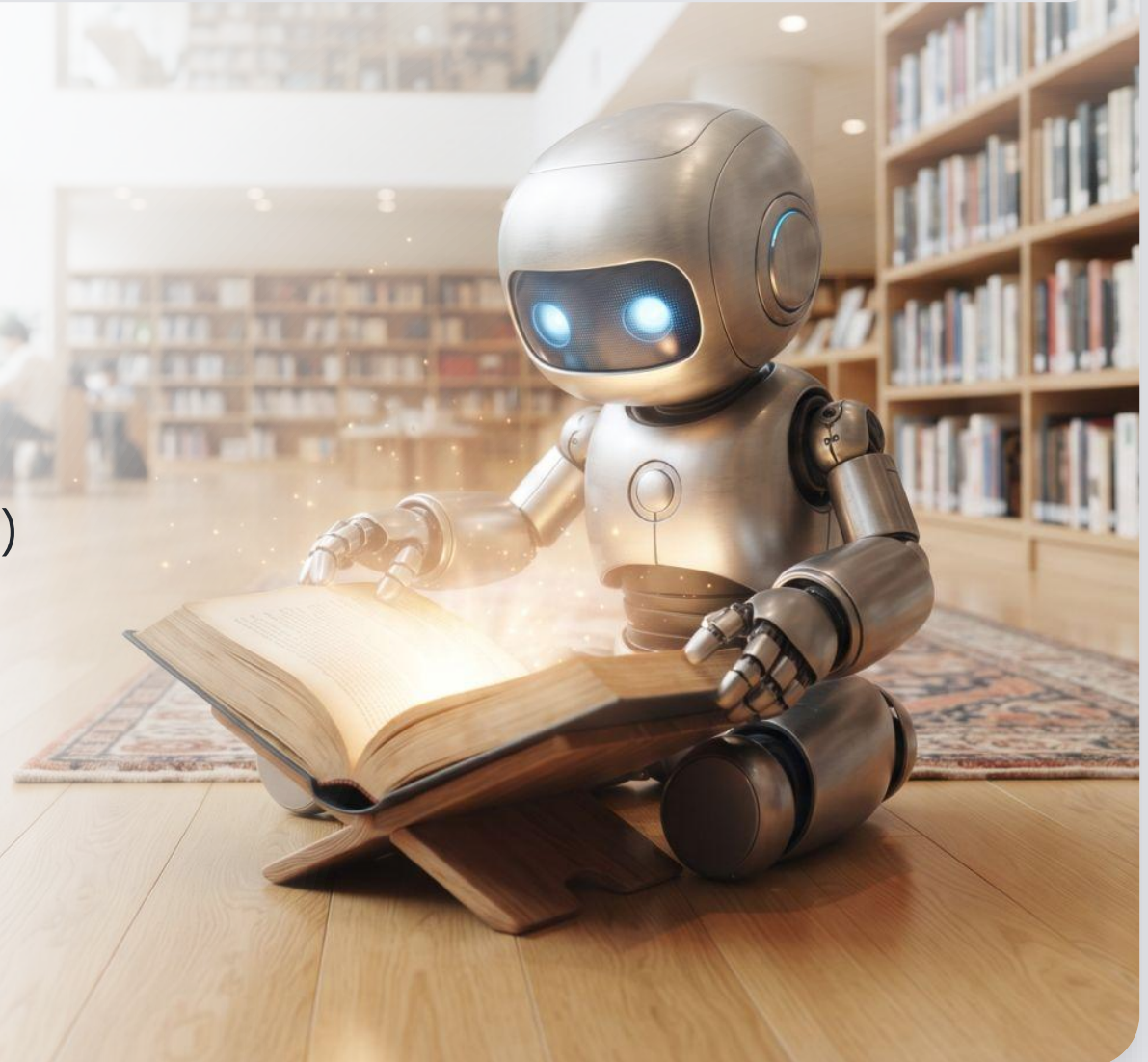


Где применить и насколько доверять?

Разберемся с понятиями

Доверие

– одно из условий взаимодействия, когда в ситуации неопределенности человек, организация (в нашем случае ИИ) будет действовать **предсказуемо** и **квалифицированно**

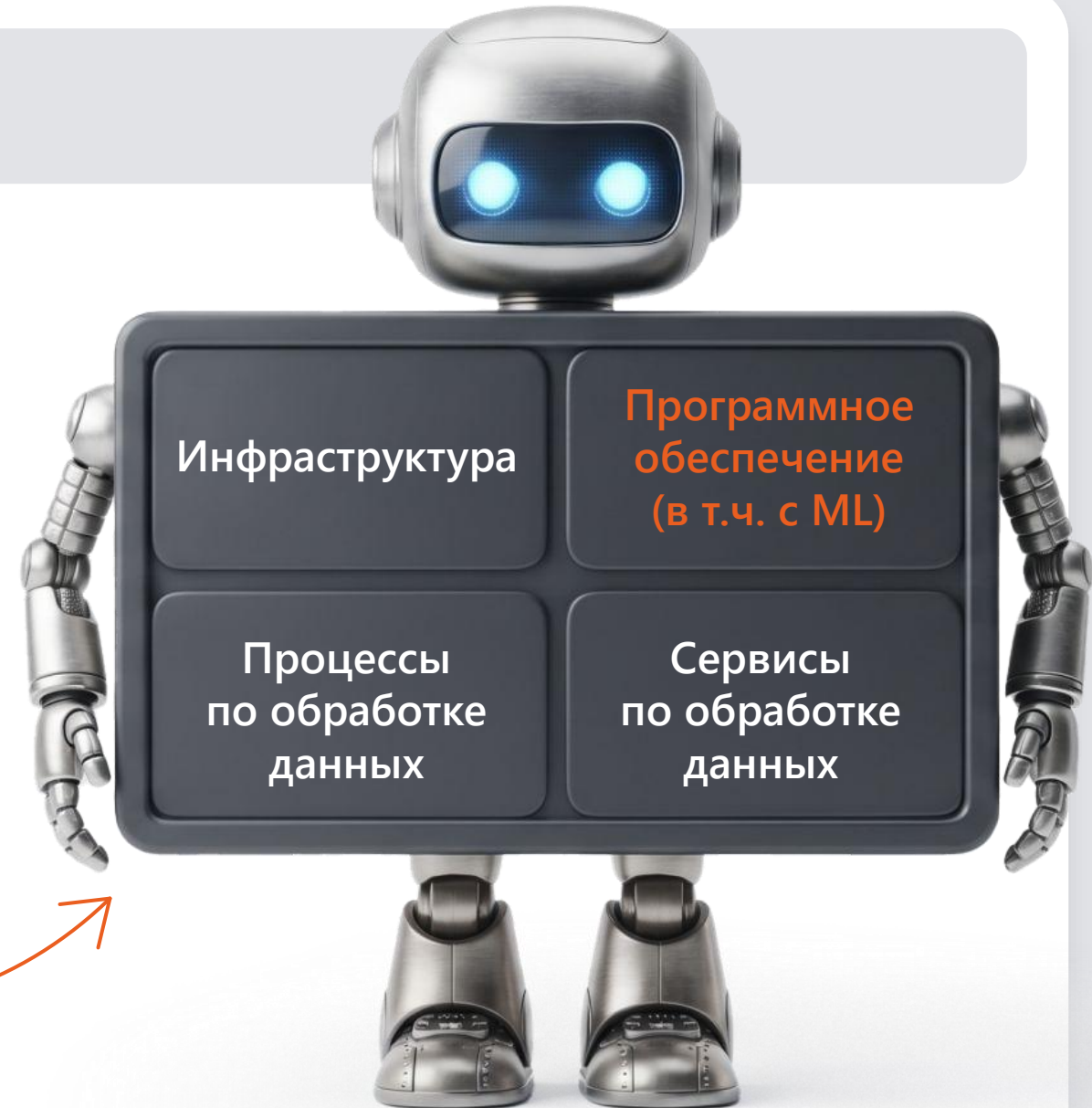


Кратко из Большой российской энциклопедии

Разберемся с понятиями

Искусственный интеллект

– комплекс технологических решений, позволяющий имитировать когнитивные функции человека и получать результаты, сопоставимые с результатами интеллектуальной деятельности человека или превосходящие их, включающий в себя...



Разберемся с понятиями

Большая фундаментальная модель

Программа для ЭВМ:

- предназначена для выполнения интеллектуальных задач
- использует алгоритмы и обучается на данных
- выводит закономерности, предоставляет информацию, обеспечивает принятие решений или прогнозирует результаты
- является основой для создания и доработки различных видов ПО
- содержит не менее 1 миллиарда параметров

Сокр. № 243-ФЗ от 26 июля 2026 г.





Риски и доверие

Основные проблемы большой фундаментальной модели



Галлюцинации



Проблема «черного ящика»



Деградация

Это не ошибка, это особенность работы



Опрос граждан РФ

Исследование НИУ ВШЭ «Информационные технологии в восприятии россиян»

Использование моделей ИИ

27,3%

2024 г.

39,2%

2025 г.

44%

2026 г.

Риски и опасения (оценка 2026 г.)

54,6%

Утрата
критического
мышления

49%

Бездумное
следование
советам ИИ

44,3%

ИИ – источник
дезинформации



«КС Ассистент»

2025 год
г. Санкт-Петербург



ИИ чат-бот
в ПК «КС Торги»



«КС Ассистент»

2026 год

Мультиагентная система

Является основой
для создания и доработки
различных видов ПО

Из определения большой фундаментальной
модели в № 243-ФЗ от 26 июля 2026 г.



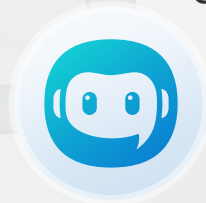
Что делает «КС Ассистент» для снижения рисков



RAG + специализированные предметные базы знаний + использование конкретных данных и документов из ПО



Использование своей поисковой модели, дообученной на предметных областях, связанных с бюджетным процессом



Формирование ответов со ссылками на первоисточники



Постоянный мониторинг показателей качества и обновление баз знаний



Использование дополнительных компактных моделей для специализированных задач

Практическое применение ИИ



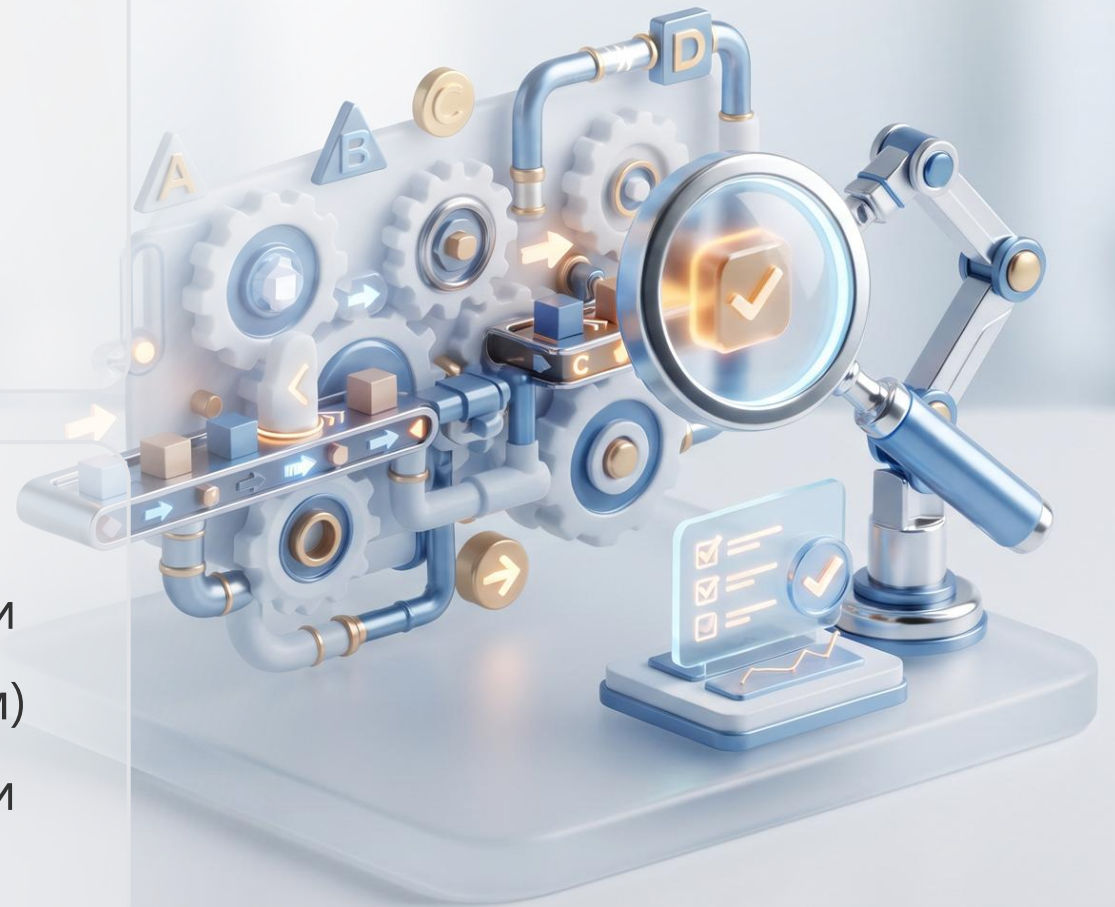
Проводим аудит бизнес-процессов, определяем:

- рутинные, трудоемкие, повторяющиеся и предполагающие работу с текстовой информацией и изображениями



При этом:

- не решаются классическими алгоритмами
- достигается эффективность (время, объем)
- не содержат чувствительной информации при работе с облачными сервисами





Что делает государство

Что делает государство

 **Федеральный закон от 26.07.2026 № 243-ФЗ** (Вступил в силу 1 сентября 2026 года)
«О поддержке развития технологий искусственного интеллекта в Российской Федерации»

 **Суверенные и национальные БФМ**
С 1 марта 2027 года

 **Порядок применения БФМ**
С 1 марта 2027 года

 **Уточняющие нормативные акты**
Ожидаются
в ноябре 2026 года

 **«КС Ассистент» готов к любому варианту**

БФМ – большие фундаментальные модели

Что делает государство



Приказ ФСТЭК России № 117 от 11 апреля 2025 г. № 117 (ред. 08.05.2026 г.)

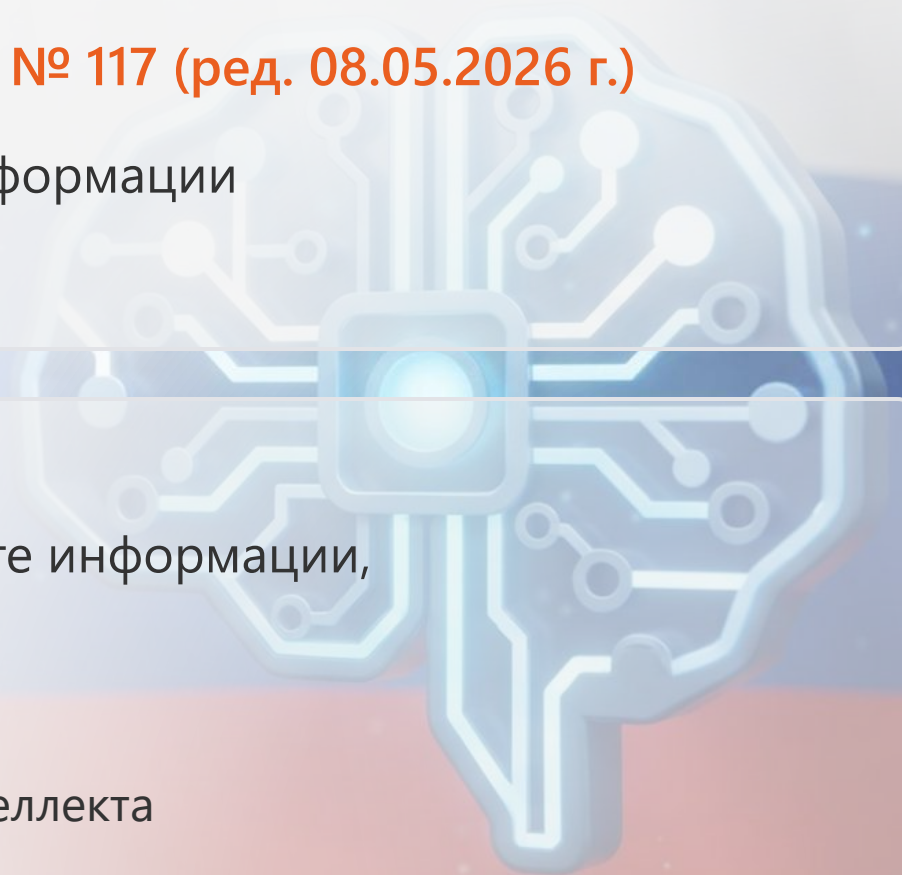
Добавлено мероприятие: обеспечение защиты информации при использовании искусственного интеллекта



Методический документ от 12 апреля 2026 г.

«Состав, содержание мероприятий и мер по защите информации, содержащейся в информационных системах»

Раздел 3.18 Обеспечение защиты информации при использовании искусственного интеллекта



Окно возможностей

До 1 сентября 2032 года

Требование применять исключительно суверенные и (или) национальные БФМ не распространяется на информационные системы при соблюдении следующих условий:



Создание и эксплуатация

БФМ созданы
и (или) эксплуатируются

до 1 марта 2027 года



Локализация данных

Обработка и хранение данных
осуществляются

на территории России

Насколько же сегодня можно доверять ИИ?

Формула доверия:

 Доверие

↓
 Предсказуемость
и квалификация

↓
 Необходимые
инженерные решения

↓
 «КС Ассистент»



Вперед, к новым
вершинам!

... с доверенными
технологиями
от «Кейсистемс»

