КОД ИБ КАЗАНЬ · 12 НОЯБРЯ 2026

IT ЭКСПЕРТ

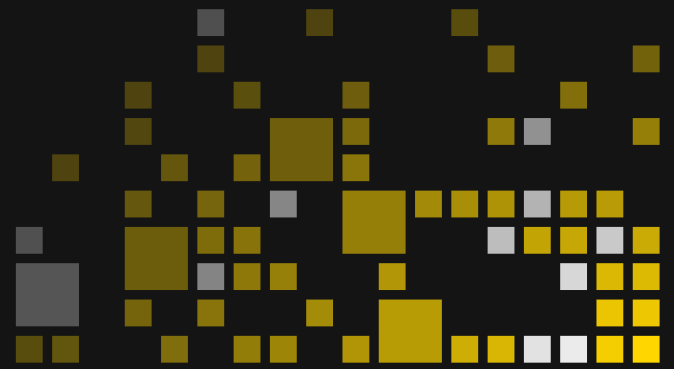
Внедрение ИИ: от договора до экспертизы

Независимая компьютерно-техническая экспертиза при спорах об ИИ-системах

Ильхом Байбулатов

Руководитель отдела IT экспертизы, судебный эксперт по производству компьютерно-технической экспертизы

Компания «IT Эксперт»



420066, г. Казань, пр. Ибрагимова, д. 56, офис 301

8 (843) 297-91-71 | info@itpravo.expert · www.itpravo.expert

Почему ИИ-проекты заканчиваются спорами

Обещания продавца должны совпадать с измеримым результатом



32%

Несоответствие ТЗ — крупнейшая доля в структуре причин.

ВЫВОД **Разрыв между обещаниями продавца и реальными возможностями системы — главный источник конфликтов.**

Типичные сценарии споров: договор и факты

Споры возникают на стыке технических и договорных неопределённостей

Сценарий	Претензия заказчика	Позиция исполнителя
Не принимают работу	Система не достигает заявленной точности	В ТЗ не было метрик; данные заказчика некачественные
Утечка данных	ИИ раскрыл персональные данные	Заказчик сам настроил доступ; это ожидаемое поведение
Дискриминация	Модель отказывает определённым группам	Данные отражают историческую реальность
Срыв сроков	Поставщик не внедрил в срок	Заказчик не предоставил данные вовремя
Скрытые расходы	Дополнительные лицензии и ресурсы	Масштабирование не было включено в договор

ВЫВОД **Ключевые вопросы спора: что стороны закрепили в договоре и что подтверждают технические доказательства.**

Независимый эксперт: нейтральный анализ фактов

НЕЗАВИСИМОСТЬ

Эксперт не связан с заказчиком или исполнителем.

ИЗМЕРИМОСТЬ

Выводы основаны на характеристиках и документированных процедурах.

ДОКАЗАТЕЛЬНОСТЬ

Заключение пригодно для арбитража, суда или переговоров.

МЕЖДИСЦИПЛИНАРНОСТЬ

Компетенции в ML, ИБ, праве и регуляторике.

ВЫВОД **Эксперт не выбирает сторону: он устанавливает, что произошло и чем это подтверждается.**

Правовые и регуляторные ориентиры

Документ	Что регулирует	Применение в экспертизе
187-ФЗ о безопасности КИИ	Безопасность критической информационной инфраструктуры	Применимо, если система связана с объектом КИИ
Приказы ФСТЭК	Защита информации, КИИ, безопасная разработка	Оценка мер защиты и аттестации объектов
ГОСТ Р ИСО/МЭК 23053	Терминология и рамки ИИ/МО	Формулировка экспертных выводов
ГОСТ Р ИСО/МЭК 23894	Управление рисками ИИ	Анализ рисков при внедрении
152-ФЗ / GDPR	Обработка персональных данных	Проверка законности обучающих данных
Рекомендации Банка России	ИИ в финансовом секторе	Экспертиза кредитного скоринга и моделей риска

ВЫВОД **Нормативные требования применяются с учётом отрасли, назначения системы и статуса объекта.**

Что исследует независимая экспертиза

Экспертиза охватывает не только модель, но весь контур внедрения

01 ТЗ И ДОГОВОР

Метрики, приёмка, права и ответственность

02 ДАННЫЕ

Качество, легальность, разметка

03 МОДЕЛЬ

Архитектура, метрики, объяснимость

04 ИНФРАСТРУКТУРА

Безопасность, развёртывание, доступ

05 ПРОЦЕССЫ

MLOps, мониторинг, human-in-the-loop

06 ЗАКЛЮЧЕНИЕ

Выводы и степень доказанности

ВЫВОД **Слабое звено может находиться на любом этапе — от договора до эксплуатации.**

Проблема «чёрного ящика»

Закрытость системы не освобождает от ответственности, но меняет методологию проверки

ПРОБЛЕМЫ

- Закрытые API и веса моделей
- Проприетарные алгоритмы поставщика
- Нет доступа к обучающим данным
- Нет прямой трассировки решения

КОСВЕННЫЕ МЕТОДЫ

- Тестовые выборки и контрольные кейсы
- Adversarial-атаки
- Анализ отклонений и ошибок
- Сравнение с baseline-моделью

ВЫВОД **Даже закрытую модель можно исследовать по наблюдаемым результатам и контрольным тестам.**

Инструментарий эксперта

Категория	Инструменты	Назначение
Анализ данных	Python, pandas, Jupyter, SQL	Исследование датасетов, поиск аномалий и пропусков
Объяснимость модели	SHAP, LIME, Yellowbrick	Оценка вклада признаков и причин решений
Adversarial-тестирование	Foolbox, ART, CleverHans	Проверка устойчивости модели к атакам
Мониторинг обучения	TensorBoard, Weights & Biases	Аудит логов обучения и метрик
Форензика артефактов	ELK, SIEM, hash-верификация	Контроль целостности данных и логов
Визуализация	Matplotlib, Seaborn, Tableau	Построение отчётов для суда и переговоров

ВЫВОД **Набор инструментов зависит от предмета спора: данные, модель, инфраструктура или процессы.**

Данные не врут? Врут.

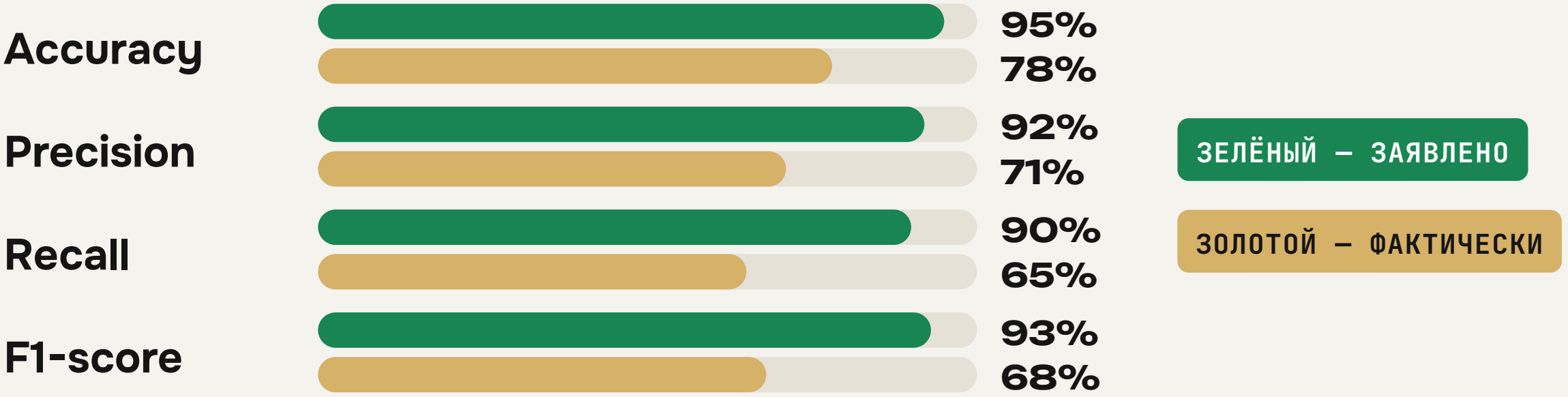
Структура дефектов в примере из исходного материала



ВЫВОД **Смещение, ошибки разметки и нерепрезентативность данных могут исказить решения модели.**

Метрики модели: ожидание vs реальность

Без согласованных метрик результат невозможно объективно принять



ВЫВОД **Разрыв метрик — классический повод для отказа от приёмки и требования возврата средств.**

Риски внедрения ИИ

Риски могут быть техническими, юридическими, финансовыми и репутационными

Риск	Вероятность	Воздействие	Как выявляет экспертиза
Репутационный ущерб	Высокая	Высокое	Публичные инциденты и ошибки модели
Регуляторные санкции	Средняя	Высокое	Проверка соответствия требованиям
Утечка данных	Средняя	Высокое	Аудит доступа и обучающих датасетов
Финансовые потери	Высокая	Среднее	Оценка ROI и фактической эффективности
Дискриминация	Средняя	Высокое	Тестирование смещений и fairness-метрики
Остановка процессов	Низкая	Критическое	Проверка отказоустойчивости и fallback

ВЫВОД **Экспертиза помогает перевести абстрактные риски в конкретные факты.**

РИТЕЙЛ • КЛИЕНТСКАЯ ПОДДЕРЖКА

LLM без контроля: микрофон в прямом эфире

СИТУАЦИЯ

Компания внедрила LLM-чатбот. Некорректная рекомендация привела к материальному ущербу. Заказчик отказался принимать систему и потребовал возврата средств.

ВЫВОДЫ ЭКСПЕРТИЗЫ

- Не реализован контроль достоверности ответов
- Нет human-in-the-loop на критичных запросах
- Нет guardrails и ограничителей тематики
- Логи не фиксируют версию модели и параметры генерации

РЕЗУЛЬТАТ

Мировое соглашение с частичным возвратом.

LLM требует особого уровня контроля и логирования.

ВЫВОД **Критичные ответы должны быть проверяемыми, а версии и параметры генерации — фиксироваться.**

Когда данные превращаются в дискриминацию

СИТУАЦИЯ

Банк внедрил ИИ-модель оценки кредитоспособности. Клиенты определённой демографической группы систематически получали отказ. Поступили жалобы и иски.

ВЫВОДЫ ЭКСПЕРТИЗЫ

- Историческое смещение в обучающей выборке
- Прокси-признаки, связанные с защищёнными категориями
- Нет fairness-метрик и тестирования
- Нет обоснования выбора признаков

РЕЗУЛЬТАТ

Модель отстранена, выплачены компенсации, проведён аудит ИИ-систем.

Fairness — юридически значимый фактор.

ВЫВОД **Проверять нужно не только точность модели, но и последствия её решений.**

Кейс 3. Срыв приёмки ИИ-проекта

СИТУАЦИЯ

После 8 месяцев внедрения точность прогнозов оказалась ниже 50%. Заказчик отказался оплачивать проект.

ВЫВОДЫ ЭКСПЕРТИЗЫ

- В ТЗ нет чётких метрик и acceptance criteria
- Обучение на синтетических, а не заводских данных
- Нет мониторинга дрейфа и переобучения
- Не переданы артефакты обучения и документация

РЕЗУЛЬТАТ

Суд частично удовлетворил иск. Договор пересмотрен, внедрение стало поэтапным.

Если критерии не согласованы заранее, спор о качестве становится спором о трактовках.

ВЫВОД **Критерии приёмки и репрезентативные данные согласуются до старта проекта.**

Как защитить проект и подготовиться к спору

Главное — заранее определить правила и сохранить доказательства

1. Закрепите в договоре метрики качества, критерии приёмки, ответственность и порядок изменения требований.
2. Определите права на модель, обучающие данные, результаты, доработки и сторонние компоненты.
3. Согласуйте доступ и передачу артефактов: датасетов, версий модели, логов и документации.
4. До запуска проверьте ТЗ, качество и правомерность данных, безопасность и сценарии отказа.
5. При инциденте зафиксируйте состояние системы, версию модели, настройки и результаты воспроизведения.
6. Соберите договоры, переписку, акты и логи; привлеките независимого эксперта до изменения системы.

ИИ-СПОР — ЭТО БИТВА ДОКАЗАТЕЛЬСТВ.
ПОБЕЖДАЕТ ТОТ, КТО СОХРАНИЛ ФАКТЫ.

Будем рады сотрудничеству!



Байбулатов Ильхом Рафикович

Руководитель отдела экспертизы компании IT Эксперт

info@itpravo.expert

www.itpravo.expert

420066, г. Казань, пр. Ибрагимова, д. 56, офис 301

8 (843) 297-91-71 | info@itpravo.expert