

Сколько стоит **не** внедрять on-premise ИИ: расчёт упущенной выгоды

Как CFO посчитать окупаемость и стоимость бездействия проекта корпоративного ИИ: методология ROI и cost of inaction, трёхлетняя модель, три отраслевых сценария и рабочий лист расчёта.

[5 категорий потерь](#)[3-летняя модель](#)[Банк · Производство · Госсектор](#)[Сквозной пример + worksheet](#)

Все денежные значения в примерах — **модельные**: они показывают, как устроен расчёт, а не сколько вы сэкономите. Замените их на данные своей компании через рабочий лист на стр. 4. Единственная подтверждённая цифра — эффект AzoneDoc — помечена отдельно.

Две стороны одного бюджета

Стоимость внедрения корпоративного ИИ известна и ограничена. Стоимость отказа от него — нет, и именно она чаще определяет экономику решения.

На комитете в центре внимания обычно одна строка — бюджет внедрения. Она конечна: пилот укладывается в несколько миллионов, промышленный запуск — в десятки. Эту сумму легко увидеть и легко испугаться. Вторая половина уравнения остаётся за кадром: пока решение откладывается, компания платит за процесс в его текущем виде. Для крупной организации эти потери — десятки и сотни миллионов в год, но они размазаны по подразделениям и не попадают ни в один отчёт.

Этап внедрения	Типовой порядок (модельно)	Характер
Пилот, 3–4 мес.	4–7 млн ₽	разовый
Промышленный запуск	10–30 млн ₽	разовый
Сопровождение, в год	15–25% от CAPEX	регулярный

Диапазоны модельные, согласованы с разбором стоимости пилота. Уточняются под конкретный проект.

Главный вопрос комитета звучит не «можем ли мы позволить себе пилот», а «сколько ещё мы готовы платить за то, чтобы его не делать».

Методология: ROI и cost of inaction

Две метрики: возврат на инвестиции отвечает «оправдается ли вложение», стоимость бездействия — «сколько стоит ничего не менять».

$$ROI = (\text{накопленный эффект} - \text{суммарные затраты}) \div \text{суммарные затраты} \times 100\%$$

Стоимость бездействия складывается из пяти слагаемых: времени на рутину, цены ошибок, простоев, цены медленных решений и конкурентных потерь. Каждое считается своим способом — разбор на стр. 3.

Четыре принципа честного расчёта

1 Коридор, а не точка

До пилота эффект — это диапазон. Считайте по нижней границе.

2 Доля применимости

На старте автоматизируется обычно 25–40% задач, а не 60–70%, как в презентациях.

3 Риск ≠ экономия часов

Снижение вероятности ошибки и сэкономленные часы — разные величины. Не складывайте.

4 Маркировка допущений

У каждого числа — источник: факт, оценка владельца или допущение.

Чего не делать. Не подставляйте суммы штрафов и санкций «по памяти» — только действующие нормы и практика вашей отрасли, согласованные с юристами.

Пять категорий упущенной выгоды

Бездействие абстрактно ровно до тех пор, пока его не разложить на счёт — от статей, что считаются в лоб, до тех, что оцениваются вилкой.

1 Время на рутину

Поиск, сверка версий, типовые ответы. Прямо через ставку и часы: **30 мин/день × N × ставка × 200 дней × доля**. Для 5 000 чел. — сотни млн ₽/год.

2 Ошибки в compliance

Покупаете не часы, а спокойствие. Через стоимость инцидента × частоту × вероятность пропуска. Штрафы — только из проверенных источников.

3 Простои в КИИ

Категория с самыми большими суммами: **стоимость часа × длительность × вероятность**. Один предотвращённый простой окупает проект.

4 Скорость решений

Медленный доступ к данным — отложенная выручка. Время поиска легко замерить до и после: удобная метрика пилота.

Категория 5 + единственная подтверждённая цифра

Конкурентная позиция. Скорость КП, разбора тендеров, ввода сотрудников. Если конкурент отвечает за день, а вы — за неделю, часть сделок уходит по скорости.

AzoneDoc (подтверждено кейсом): поиск документа в архиве — с **2–4 часов до 1–3 минут**. Остальные числа в документе модельные.

Трёхлетняя ROI-модель и сценарии

Все значения в млн ₽ и **условные** — это структура расчёта, а не прогноз. «Риски» включают предотвращённые потери. Сценарии различаются весом этой колонки.

Базовая модель — офисная компания, акцент на рутине и поиске

Год	CAPEX	OPEX	Время	Риски	Чистый	Накопл.
0 (пилот)	4–7	0	0	0	–4...–7	–4...–7
1	0	2–4	15–25	5–10	+14...+27	+7...+22
2	0	2–4	20–30	8–15	+22...+39	+29...+61
3	2–4	2–4	25–35	10–18	+27...+45	+56...+106

Сценарий А — банк: высокая цена ошибок в КУС и договорах

Год	CAPEX	OPEX	Время	Риски	Чистый	Накопл.
0	5–7	0	0	0	–5...–7	–5...–7
1	0	3–4	18–28	20–35	+34...+59	+29...+52
2	0	3–4	24–34	30–50	+50...+80	+79...+132
3	2–3	3–4	28–38	35–55	+56...+86	+135...+218

Сценарий Б — производство / энергетика (КИИ): доминируют простои

Год	CAPEX	OPEX	Время	Риски	Чистый	Накопл.
0	6–9	0	0	0	–6...–9	–6...–9
1	0	3–5	10–18	40–80	+47...+93	+41...+84
2	0	3–5	14–22	60–120	+71...+137	+112...+221
3	3–4	3–5	18–26	70–140	+81...+157	+193...+378

Госсектор — промежуточный профиль: ровный эффект на документообороте, выше CAPEX из-за закрытого контура и сертификации. Окупаемость во всех сценариях — в течение первого года.

Сквозной пример: как собрать расчёт

Гипотетический производственный холдинг, 4 000 сотрудников, из них около 1 200 — офисный персонал. Все числа иллюстративные; цель — показать логику.

Шаг 1. Рутинная. Допустим, рутинным поиском заняты ~40% из 1 200 офисных. 30 мин/день, ставка 2 000 ₽/час:
 $1200 \times 0,4 \times 0,5 \times 200 \times 2000 \text{ ₽} \approx 96 \text{ млн ₽/год}$.

Шаг 2. Простой. Пусть час простоя линии — 800 тыс. ₽, два простоя по 6 ч/год, предотвратимо ~1/3:
 $800 \times 6 \times 2 \times 0,33 \approx 3,2 \text{ млн ₽/год}$ (нижняя граница).

Шаг 3. Ошибки в договорах. Входят как управляемый риск отдельной строкой, с пометкой «уточнить у юристов» — без выдуманной суммы.

Шаг 4. Итог. Считаемая часть — **≈ 99 млн ₽/год** по нижней границе, без неоцифрованных рисков.

99 млн ₽ — это **потенциал**, а не гарантированная экономия: он превышает стоимость пилота (4–9 млн ₽) более чем в десять раз. Для комитета это аргумент — масштаб уже идущих потерь. Сколько из этого извлечь, покажет пилот.

On-premise против облака в деньгах

Параметр	Облако (SaaS / API)	On-premise
Стартовые затраты	низкие, без CAPEX	высокие: железо + развёртывание
Стоимость вдолгую	растёт с объёмом	предсказуемая, амортизируется
Данные за периметром	да — риск для ПДн, КИИ	нет — всё в контуре
152-ФЗ / ФСТЭК	отдельная проработка	реализуемо в закрытом контуре
Когда дешевле	малые объёмы, нечувств. данные	большие объёмы, регулируемые данные

Точка безразличия. Условно: при стабильном потоке от нескольких десятков тысяч запросов в месяц накопленная плата за облако догоняет CAPEX on-premise за полтора-два года — дальше своя инфраструктура дешевле. Но для ПДн и КИИ вопрос чаще не «дороже ли», а «имеем ли право выпустить данные за периметр».

Чек-лист финансовой готовности

 Выбраны 1–2 процесса с понятной болью и данными

 Определён владелец процесса с полномочиями

 Посчитана текущая стоимость: часы, ставки, ошибки, простой

 Согласованы требования ИБ и режим закрытого контура

 Определены метрики успеха/неуспеха пилота

 Заложен бюджет на пилот и на сопровождение

Посчитаем cost of inaction для вашей компании

AZONE-AI проведёт экспресс-аудит процессов, соберёт модель ROI на ваших данных и предложит архитектуру пилота в закрытом контуре. Лицензии ФСТЭК, ФСБ и Министерства обороны РФ.

Запросить расчёт ROI

azoneai.ru/contact · услуга «Развёртывание on-premise LLM»