



avito

москва 2025

Екатерина
Давыдова

Программный менеджер GenAI Авито

GenAI без иллюзий

системный подход к запуску LLM-проектов

Наши спикеры



Екатерина Давыдова

Программный менеджер
GenAI Авито



GenAI

усмиряем хаос

- Выход ChatGPT спровоцировал волну интереса к GenAI
- Хайп не угасает, но системные внедрения единичны
- Вызовы выросли еще сильнее



Структура доклада

- 01 Генерация идей
- 02 Оценка и отбор инициатив
- 03 Проверка гипотез
- 04 Пилот
- 05 Масштабирование

Генерация идей: два лагеря

1

Давайте завтра
внедрим LLM
во все наши продукты



2

Генеративный искусственный
интеллект - это что-то
для генерации ключей?



Воркшоп. Что такое GenAI?

GenAI хорош в генерации контента: текст, изображения, музыка

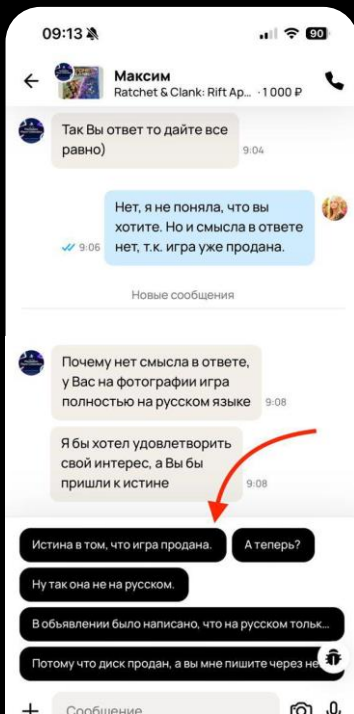
LLM - это большие языковые модели, замечательно справляются с генерацией текстов, структурированием и саммаризацией контента

Для задач с понятной детерминированной логикой, которую можно четко описать, применяйте алгоритмический подход, здесь GenAI НЕ нужен

Для задач классификации, сортировки, определения закономерностей используйте классический ИИ

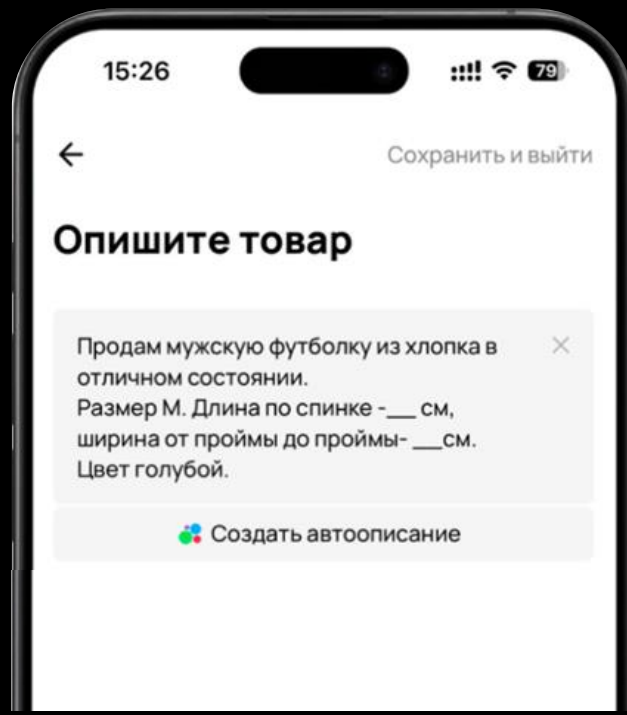
Воркшоп. Конкретные кейсы и research

+0,6% покупателей
ежедневно



+1,2% рост продаж

1.7% рост заказов
с доставкой



Воркшоп. Брейнсторм и генерация гипотез



- 1 Выбираем только те идеи, где без GenAI не обойтись
- 2 Вдумчивая проработка и оценка
- 3 На выходе - конкретная гипотеза, которая решает боль

Отбор идей

Смотрим в первую очередь
на бизнес-потенциал

Достаточно
ли отскорить по RICE?



Маркетингуемость



“Готов ли наш продукт “выстрелить” так, чтобы о нем говорили — и рынок, и пользователи?”

Технологичность

“Продукт показывает не только, что мы сделали, но и кем мы стали. Насколько мы выросли, создавая его?”

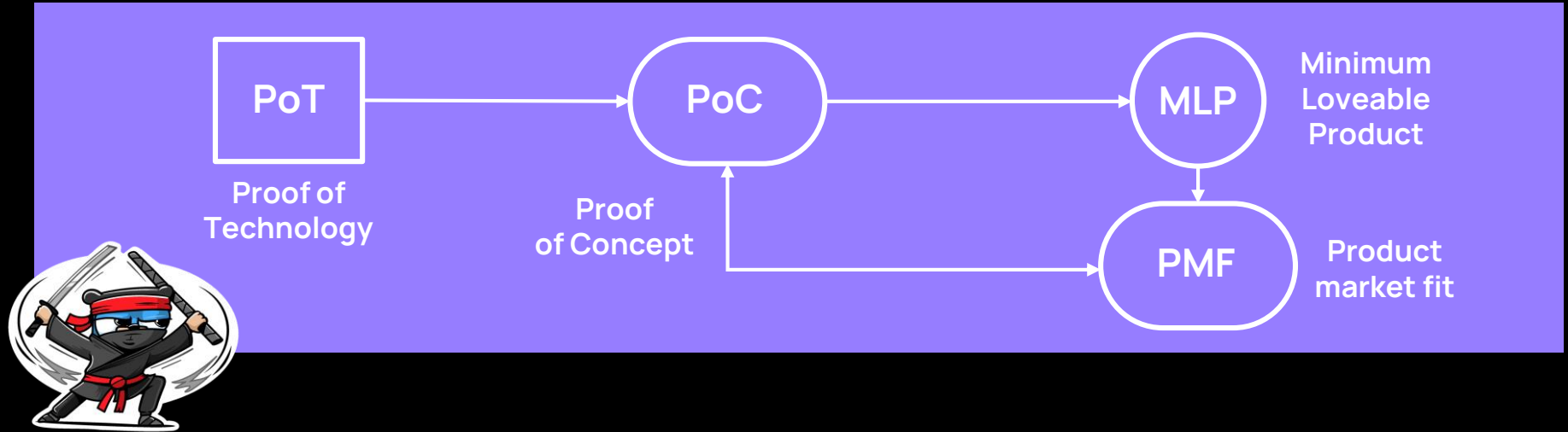


Юридические риски

- 01 Закон о персональных данных (локализация в РФ) - отсутствие трансграничной передачи
- 02 Авторские права - проверка обучающей выборки
- 03 Утечка интеллектуальной собственности - соглашение/договор с вендором
- 04 Запрещённый контент и этика - разработка guardrail



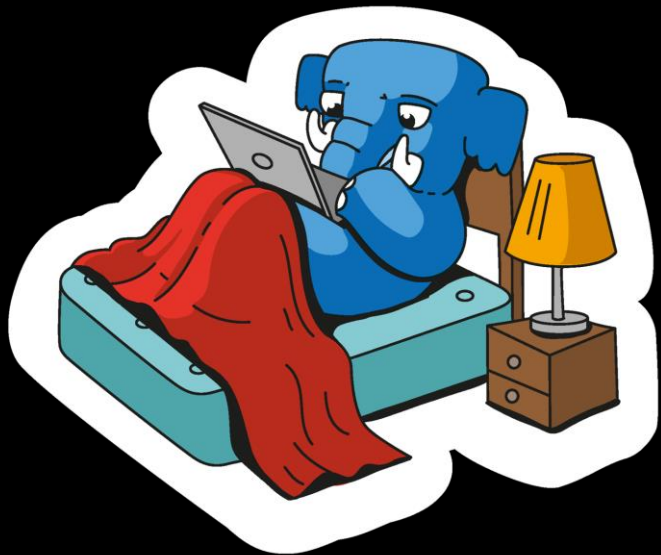
Проверка гипотез



- Три ступени к MLP

- Подходы для раннего поиска **value**

Product market fit



- F&F запуски

наша супер-сила -
это наши сотрудники

- 'Human-in-the-loop' запуски

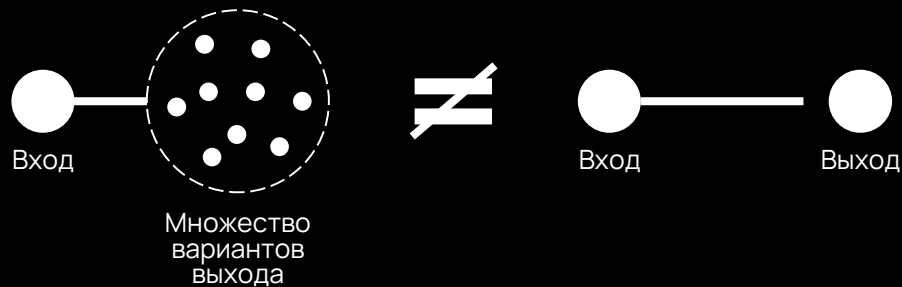
определяем прокси-метрики
для оценки бизнес-потенциала

Переходим к деливери?



LLM $\neq$ классический софт

- Модель вероятностная, не детерминированная
- Нужен отдельный пайплайн качества и детальное логирование



Оценка качества

01

Проверить LLM
по чек-листу
нельзя

02

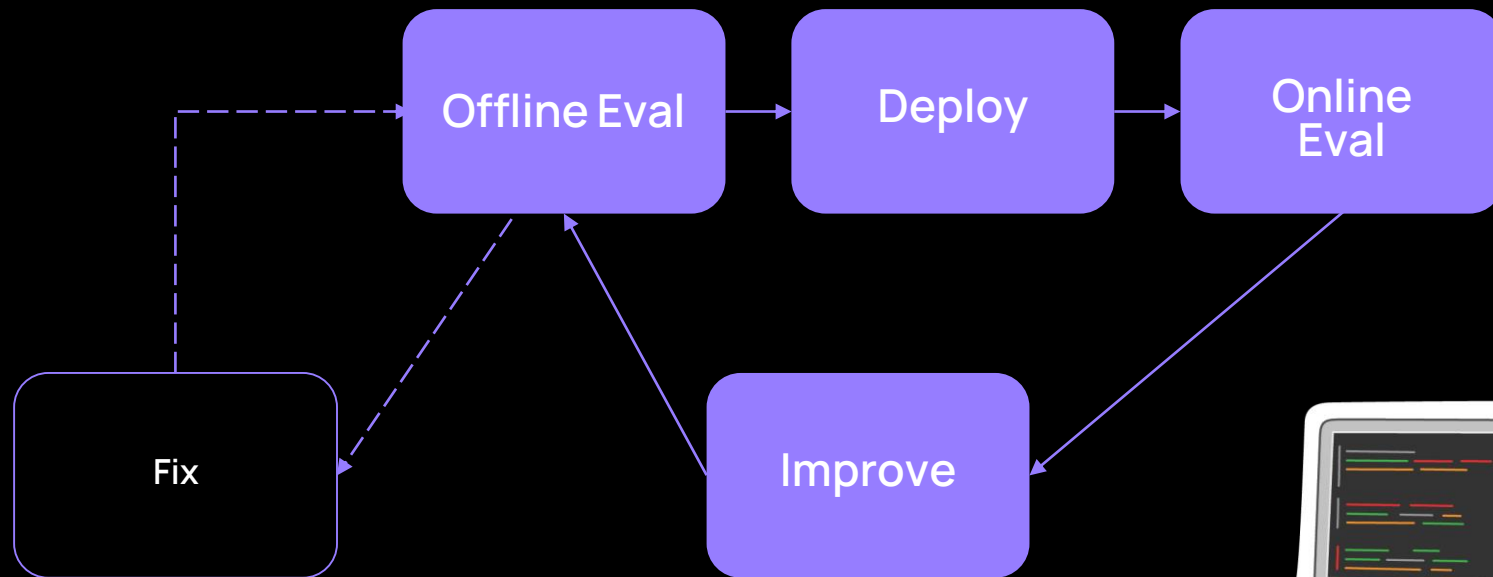
Evaluation-
фреймворк

- LLM-as-a-judge
- Human review

03

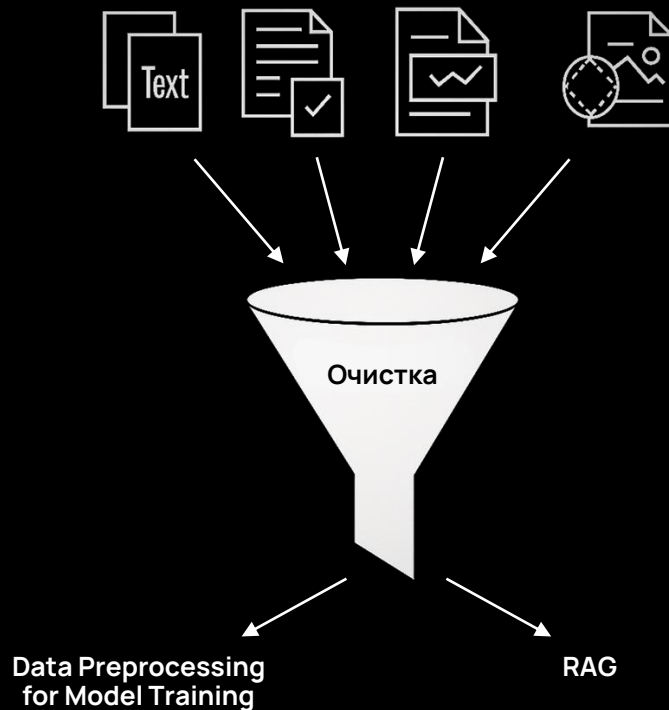
Нужен отдельный
пайплайн оценки
качества

Итеративность



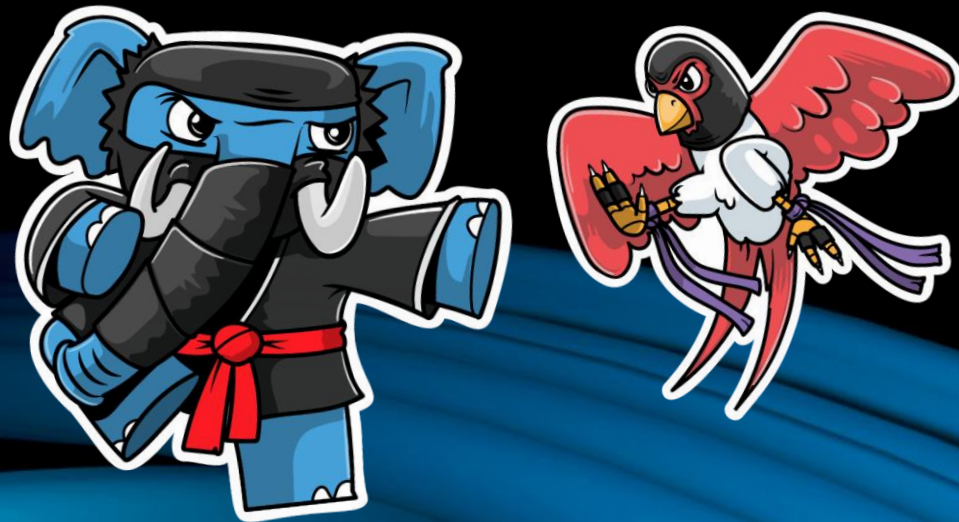
Подготовка данных

Обязательно
предусмотреть
этап и ресурсы



Guardrail

Про защиту



Команда и компетенции

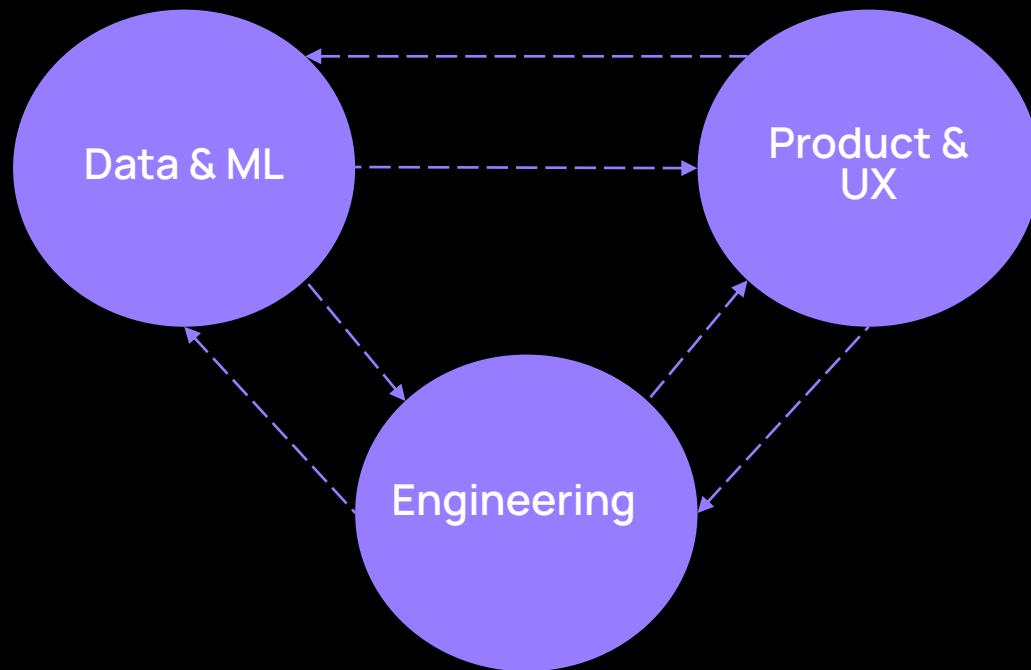
[1] Новые роли и компетенции в команде

- промптинг
- сбор и подготовка данных
- разметка данных, разный уровень экспертности
- взлом, поиск уязвимостей

[2] Повышение GenAI-грамотности



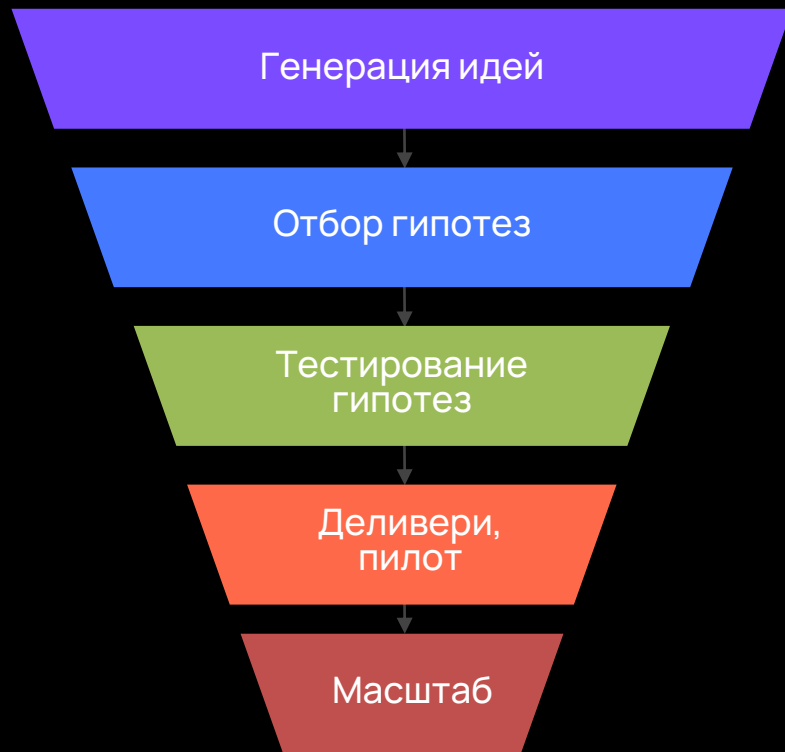
Команда: баланс трех миров



Scaling



**Масштабируем
ТОЛЬКО ТО,
что доказало
ценность**



Мониторинг

- [1] Мониторинг качества и аномалий
- [2] Latency
- [3] Контроль стоимости инференса



Платформизация



01 Создание LLM-платформы

02 Создание конкретных кубиков и фреймворков, необходимых для ряда проектов
Eval, Guardrails, Multi-agent system

Путь от хайпа к зрелой GenAI-культуре

"GenAI - это не магия, а новая инженерия. Здесь побеждает тот, кто сумел выстроить системный подход и обуздал хаос".





Q&A