Agile в AI проектах:

оценка и приоритизация эпиков

LeanDS, QUEST framework

Artemy Malkov, PhD

DATA MONSTERS

Yuliya Rubtsova, PhD

DATA MONSTERS

NVIDIA ELITE PARTNER

16 ЛЕТ

НА РЫНКЕ AI

Артемий Малков, PhD

Юлия Рубцова, PhD

Nestle Adobe intel P&G AMD Raiffeisen BANK TOYOTA

SIEMENS GE CISCO ABInBev HUAWEI Deutsche Bank

СБЕР LUKOIL X5 Тинькофф Банк МегаФон ВТБ Снежная Королева открытие

Типичный ИИ проект

Поди туда, не знаю куда,

принеси то, не знаю что

Типичный ИИ проект

Непонятно!

Непонятно, как оценить сроки выполнения задач

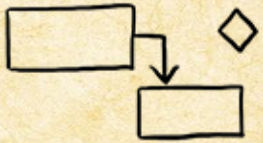
Непонятны критерии успеха

Непонятно когда все будет готово

Боссы недовольны

Квест – новая категория проектов, которым вас никто не учил

PROCESS



PROJECT

QUEST

DELIVERY

Стандартизация

Повторяемость

Неопределенность

Уникальность

DISCOVERY

Waterfall

Agile

Lean DS

Непредсказуемый прогресс

Delivery

Планомерный прогресс

Наблюдаемый инкремент

Discovery

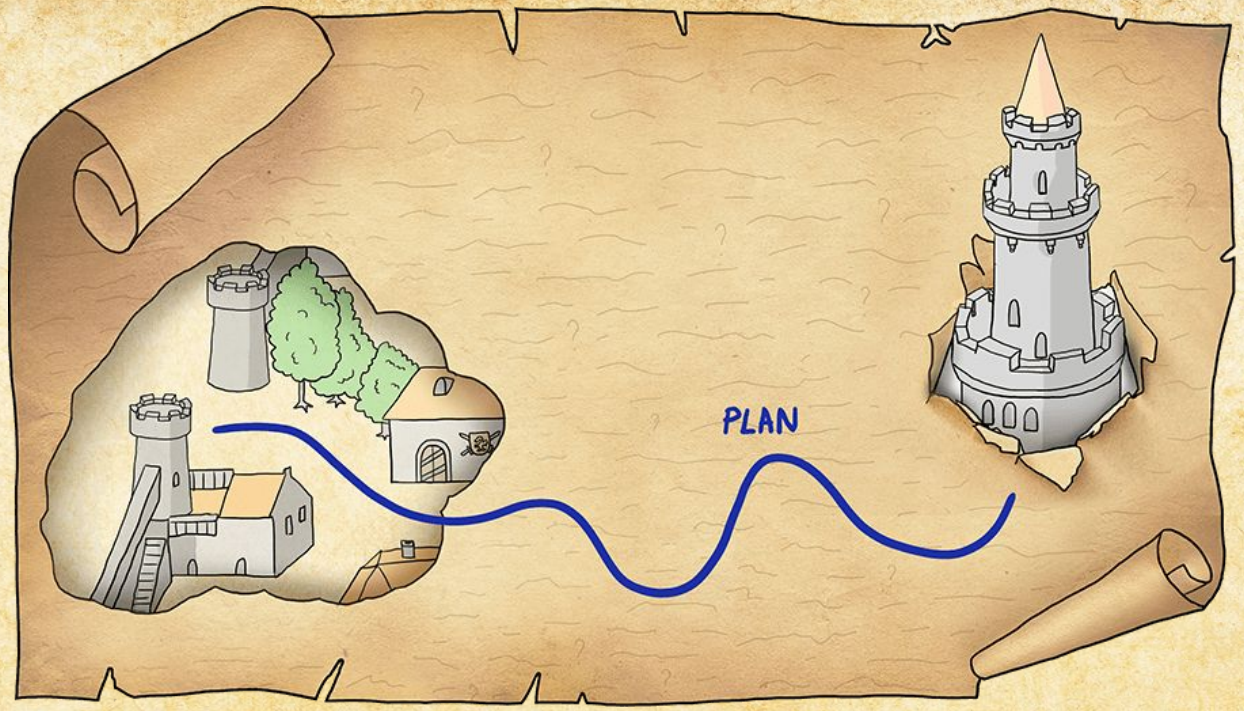
Перебор большого числа вариантов, прежде чем что-то найдется,

Долгий сбор данных по редким событиям,

Неведомое

Как планировать, когда непонятно?

Можно ли планировать непредсказуемый прогресс

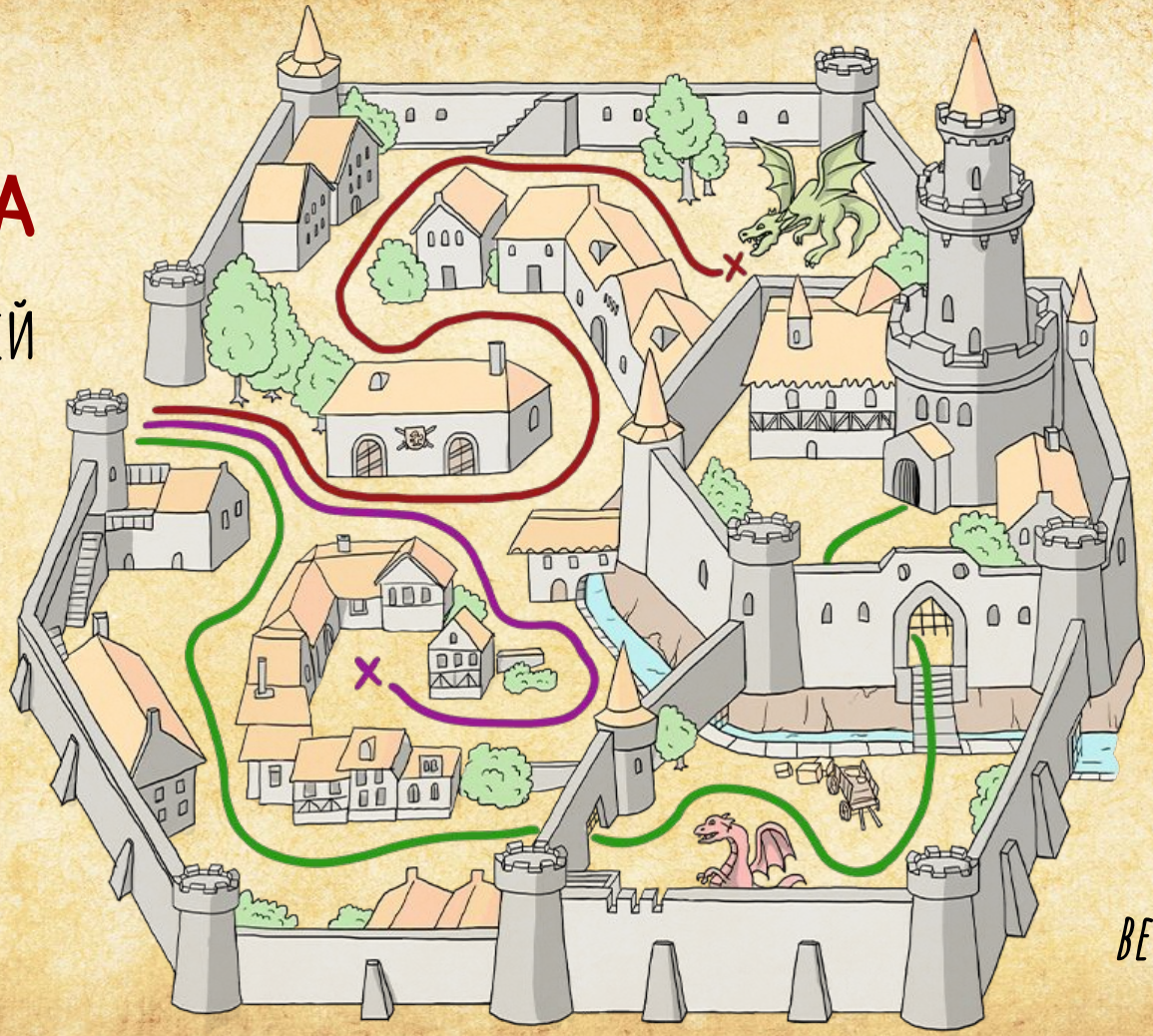


Накинуть
20 – 40%
на риски?

Как **планировать**, когда ничего непонятно?

ПРИНЦИП ЛАБИРИНТА

Множество путей и попыток достижения целей и решения задач



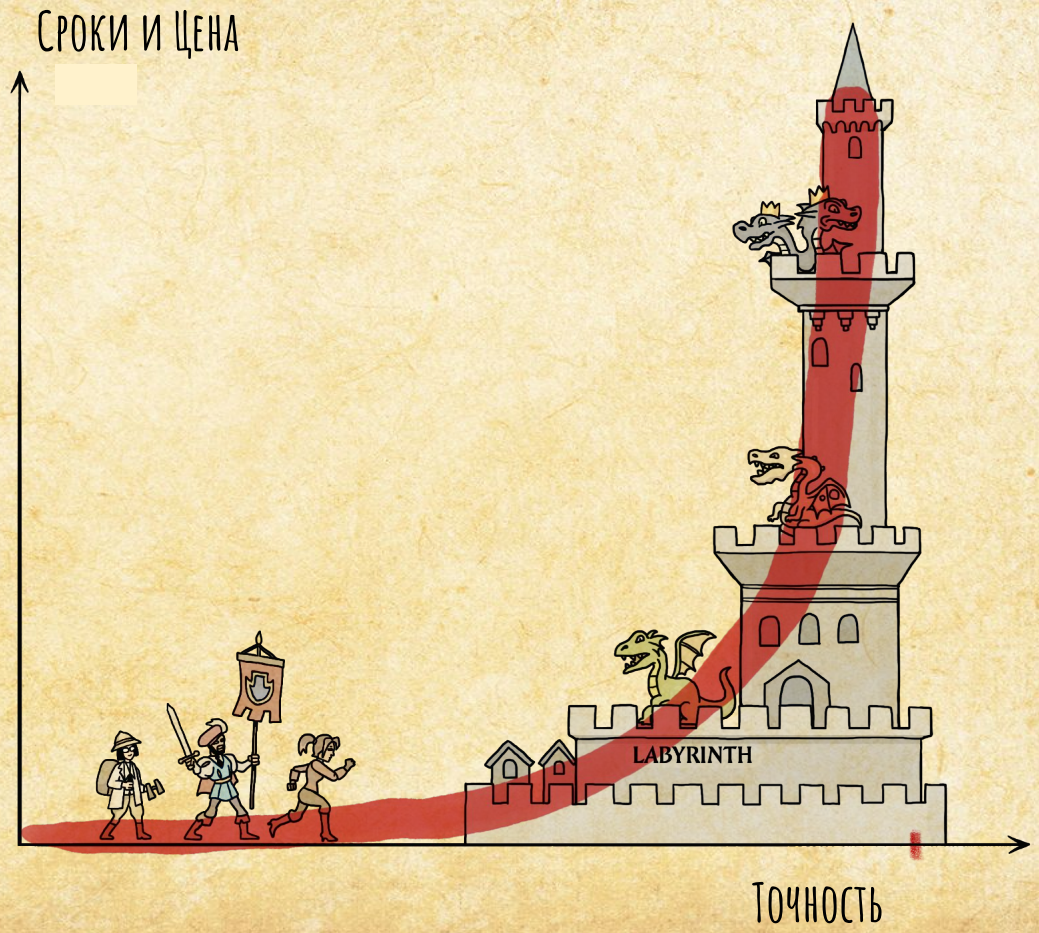
А может быть **200 – 400%** на риски?!

Метод работы венчурного фонда

Strong Problems

В отличие от традиционного software development,
в Data Science большинство задач не имеют единственного “правильного” решения.

Простую модельку с **60–70%** точности можно собрать за пару часов. Но по мере приближения к **100%** на ту же задачу может уйти несколько лет экспериментов.



7 DRAGONS

главные причины
сложностей
и задержек
в ИИ проектах

Decision-makers

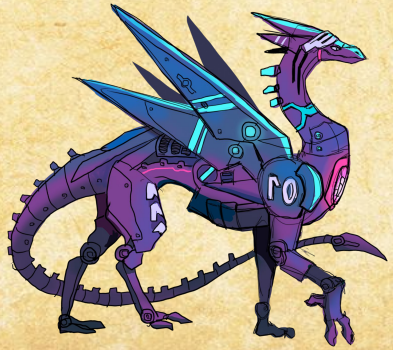
Real time

Access to data

Gigantic tail

Operations

Novelty



Sparsity and scarcity

Ну и как же планировать?

Клиенту нужен пирог. Как организовать работу?

Планирование

1. Разрезать на части
2. Приоритизировать

Profit !

Adoption

Integration

Development

Research

Как разрезать работу?

Декомпозиция

Waterfall планирование

Horizontal slicing

4 Adoption

3 Integration

Development 2

Research 1

Agile планирование

Vertical slicing

4th User Story / Epic

1st User Story / Epic

3rd User Story / Epic

2nd User Story / Epic

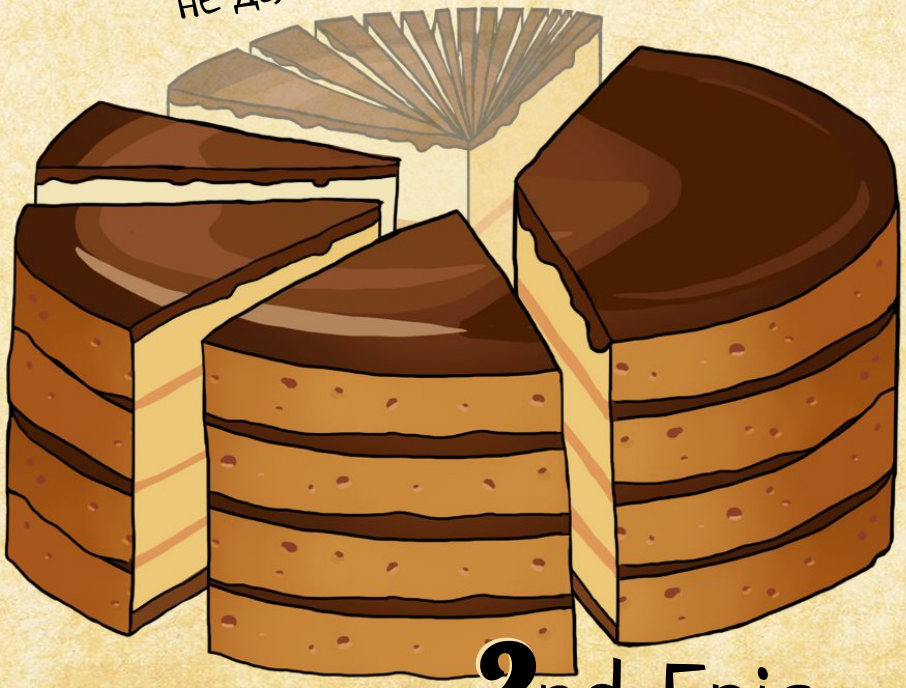
Lean DS планирование

Pareto slicing

5 6 7 8 9 10 11 12 13 14 15 16 17

не делаем вобще!

Можем себе позволить из-за избыточности путей:
принцип лабиринта



1st Epic

maximum impact,
minimal effort

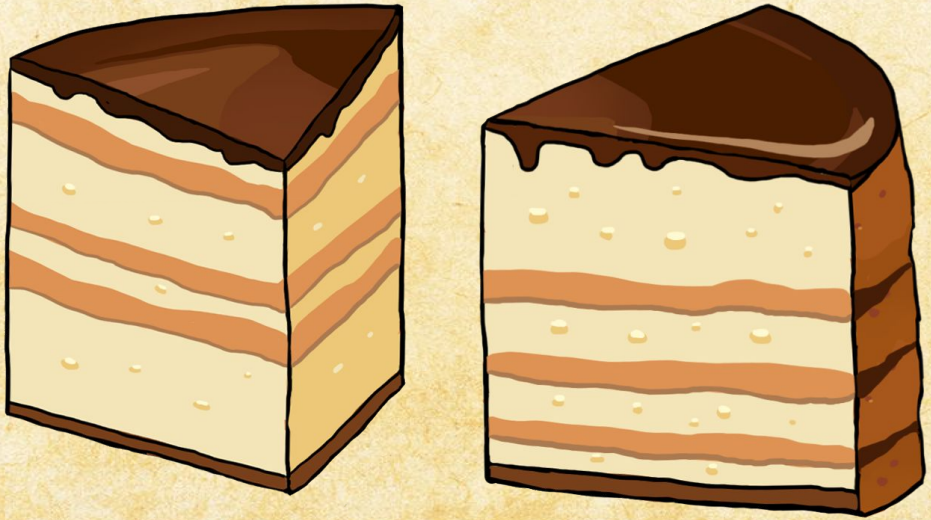
2nd Epic

3rd Epic

4th Epic

Not all stories are created equal

В каждой истории **Effort** на **R**esearch, **D**evelopment, **I**ntegration и **A**doption неравномерны.



$$Effort = Effort_R + Effort_D + Effort_I + Effort_A$$

Как выбрать кусочек?

Приоритизация

Приоритизация

Традиционный Software Product

Feature, User Story, Epic

Добавление новой Feature в работающем продукте.
Методы приоритизации: ICE / RICE
Reach, Impact, Confidence, Effort

Data Product

Data Pipeline, Data Module, Data Epic

Часто - создание нового процесса
Zero to hero — from 0 to 95% за N дней.
Высокая НЕОПРЕДЕЛЕННОСТЬ.
Непонятно как считать Confidence

$$ICE = \frac{Impact \times Confidence}{Effort}$$

Impact = [0.25, 0.5, 1, 2, 3]

Confidence = [100%, 80%, 50%, 10%]

Effort = $\sum$ Effort i

i in [R, D, I, A]

Effort суммируем для всех фаз:

Research, Development, Integration, Adoption

$$\text{R}\,\text{ICE} = \frac{\text{Impact} \times \text{Confidence}}{\text{Effort}}\ \text{Reach}$$

Impact = [0.25, 0.5, 1, 2, 3]

Confidence = [100%, 80%, 50%, 10%]

Effort = $\sum$ Effort i

i in [R, D, I, A]

Effort суммируем для всех фаз:

Research, Development, Integration, Adoption

$$DICE = \frac{Impact \times Confidence}{(\sum Effort_i \times \pi^{Dragons_i})}$$

$Effort_i$ считаем в идеальных часах, из расчета “нам повезёт”, но потом домножаем на **π** за каждого дракона (явный риск)

RESEARCH

Новизна задачи

Качество данных

DEVELOPMENT

Вариативность, длинные хвосты

Расчет в реальном времени

INTEGRATION

Доступ к данным, железу и системам

ADOPTION

Боссы, стейклохдеры, согласования

Operations, зрелость бизнес-процесса

ПРИМЕР

Посчитаем $\sum \text{Effort}_i \times \pi^{\text{Dragons}_i}$

8 недель → 39 недель

RESEARCH	$1 \text{ неделя} \times \pi^{2}$		Никогда не делали + Плохие данные	10 недель
				+
DEVELOPMENT	$3 \text{ недели} \times \pi^{0}$			3 недели
				+
INTEGRATION	$2 \text{ недели} \times \pi^{1}$		Пока нет доступа для интеграции	6 недель
				+
ADOPTION	$2 \text{ недели} \times \pi^{2}$		Нужно согласовывать + Изменить бизнес-процесс	20 недель

Получится примерно так

Эти пока не делаем – эпики с большим числом драконов

Pareto slicing

Тут еще пара эпиков с большим импактом, но есть драконы.

Берем их в работу, но если “**проснется дракон**”, то будем решать, продолжать или заморозить и искать обходной путь.

Беремся за этот эпик

максимальный impact,
минимальный effort
мало рисков

Тоже годный эпик – делаем!

Adoption: автоматизация работы людей

Примеры:

- Заполнение карточки товара
- Ответ на отзыв клиента
- Модерация фотохостинга
- Заполнение форм

Time = money

Adoption: автоматизация работы людей

Adoption: B2B маркетплейс

KPI – маржа с завершенных заказов

Покупатели

Покупатели

Покупатели

Брокер

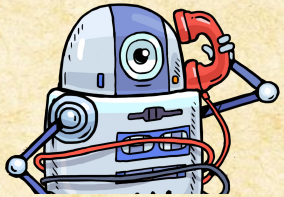
10 заказов / день

Продавцы

Продавцы

Продавцы

Adoption: B2B маркетплейс



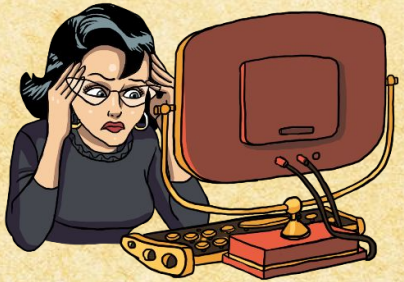
100 заказов / день

AI

Покупатели

Покупатели

Покупатели



Брокер/Агент
10 сложных (рисковых) заказов
для дообучения модели

Продавцы

Продавцы

Продавцы

Для тех, кто все проспал

- Обсудили, что такое типичный ИИ проект,
- Разобрали отличия Delivery проектов от Discovery,
- Прикинули, как планировать, когда ничего не понятно,
- Научились нарезать “пирог” задач,
- Поговорили, как приоритизировать и за что хвататься,
- Поделились кейсами с проблемами Adoption

Узнать больше:

Lean DS группа

t.me/leands_chat

Артемий Малков

t.me/asmalkov am@datamonsters.com

Юлия Рубцова

t.me/yuliya_rubtsova yrubtsova@datamonsters.com

Спасибо за внимание

Artemy Malkov, PhD

Yuliya Rubtsova, PhD