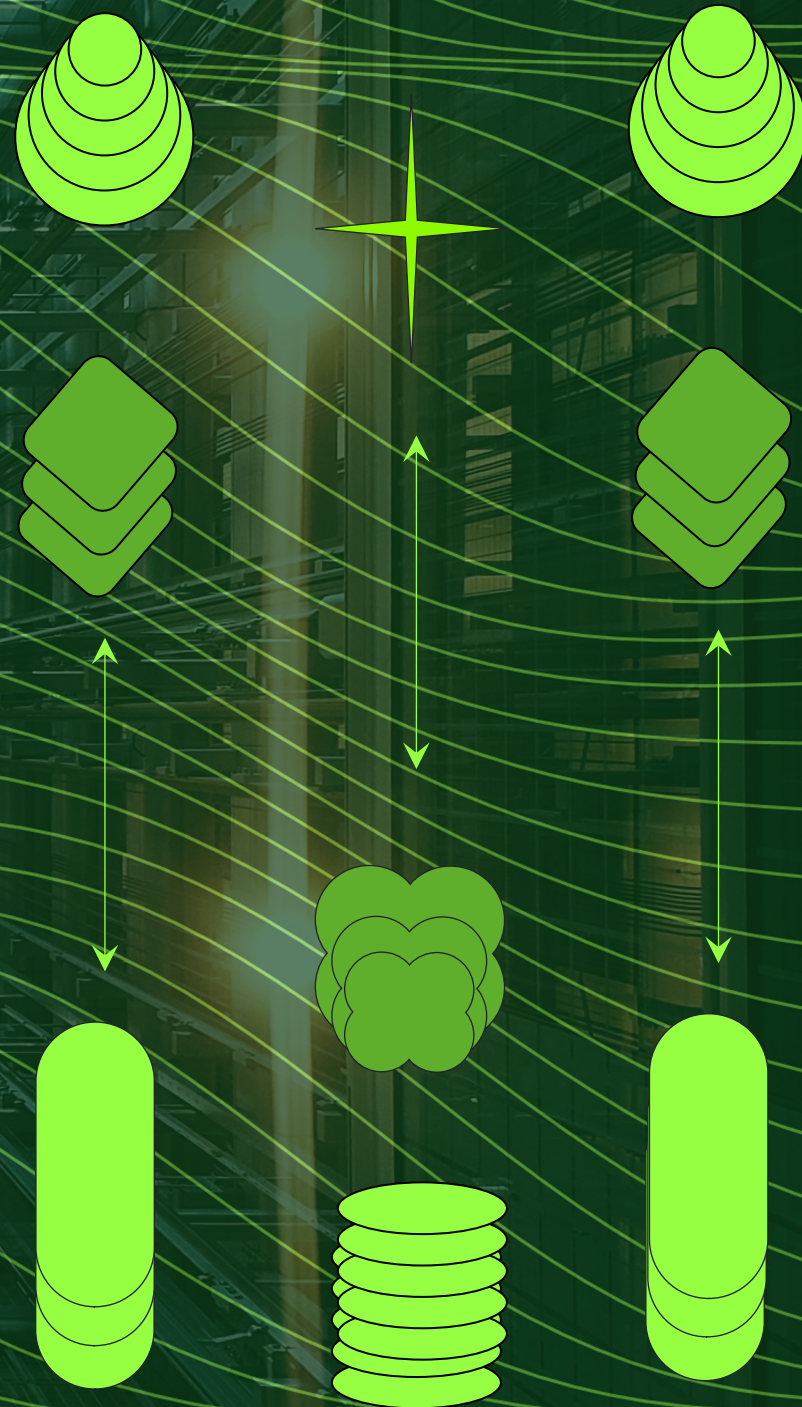


QA в Big Data без прекарас



Обо мне

Виктория Дежкина

Руководитель команды функционального тестирования X5 Tech

Работаю 13 лет в IT, 5 лет в X5 Group



О чем будем говорить?

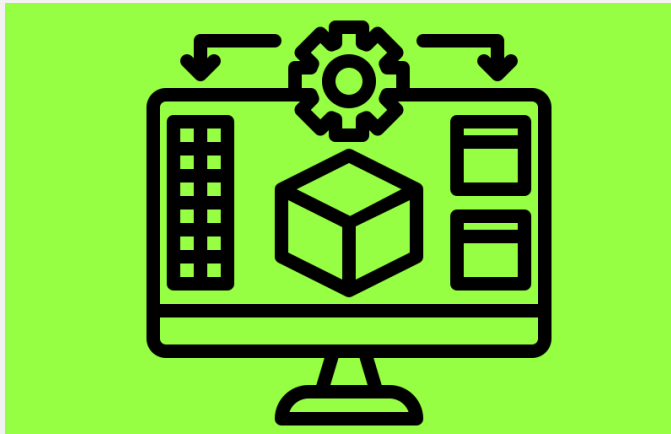
Что такое тестирование Big Data
и в чем отличие от классического тестирования?



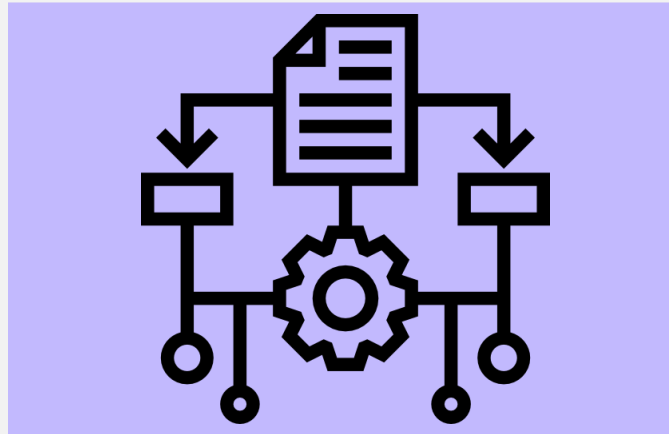
Big Data

Большие данные – это набор больших наборов данных, которые не могут быть обработаны с использованием традиционных вычислительных технологий

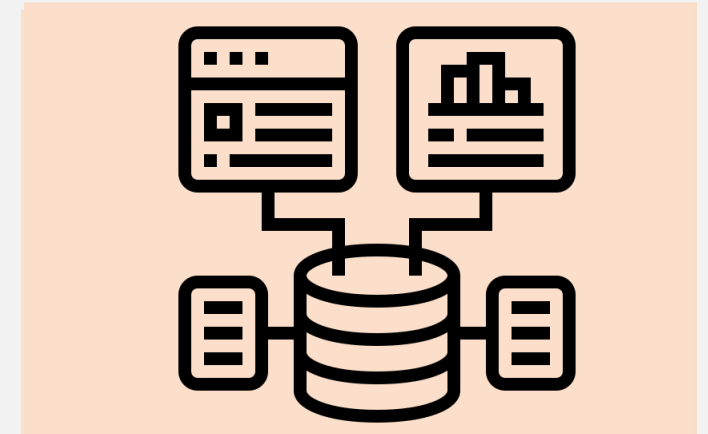
Что это значит для тестировщика?



Фокус на
Архитектуру



Логика обработки
расчетов



Тестовые данные

В чем разница?

Почему возникает необходимость
в отдельной стратегии для больших данных?

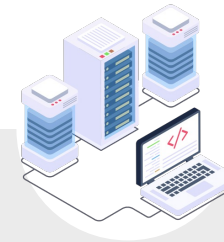


Разница между тестированиями



ТЕСТИРОВАНИЕ ТРАДИЦИОННОЕ

- ✓ Использует структурные данные
- ✓ Не требует специальной среды
- ✓ Тестировщик может выбрать ручное или автотестирование



ТЕСТИРОВАНИЕ БОЛЬШИХ ДАННЫХ

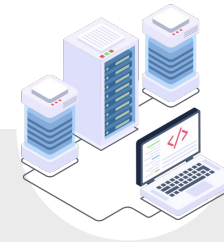
- ✓ Может иметь все типы данных
- ✓ Требуется специальная среда
- ✓ Выбор стратегии зависит от выборки

Разница между тестированиями



ТЕСТИРОВАНИЕ ТРАДИЦИОННОЕ

- ✓ Использует UI-средства автоматизации
- ✓ Требуется базовые знания по инструментам



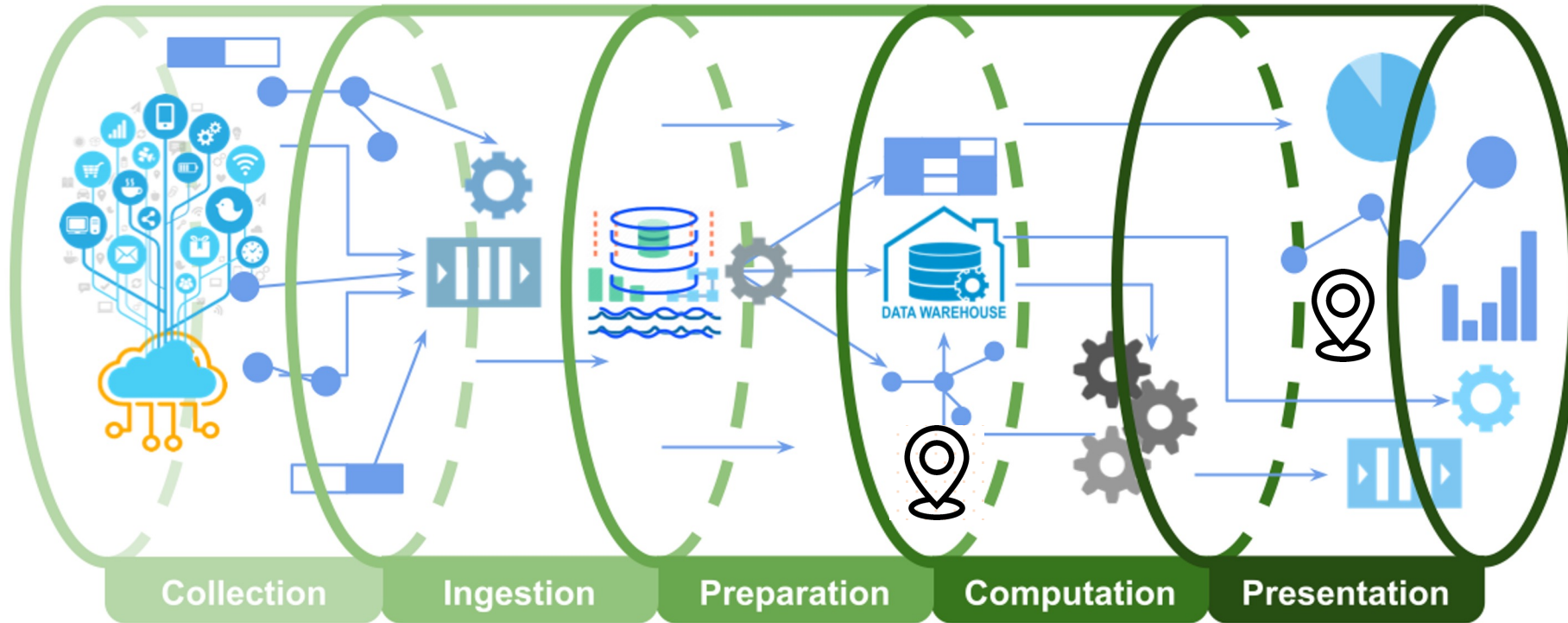
ТЕСТИРОВАНИЕ БОЛЬШИХ ДАННЫХ

- ✓ Необходимы свои инструменты (MapReduce и HIVEQL.)
- ✓ Необходим определенный набор навыков и обучения по тестированию инструмента

Архитектура продуктов



Архитектура продуктов в Big Data



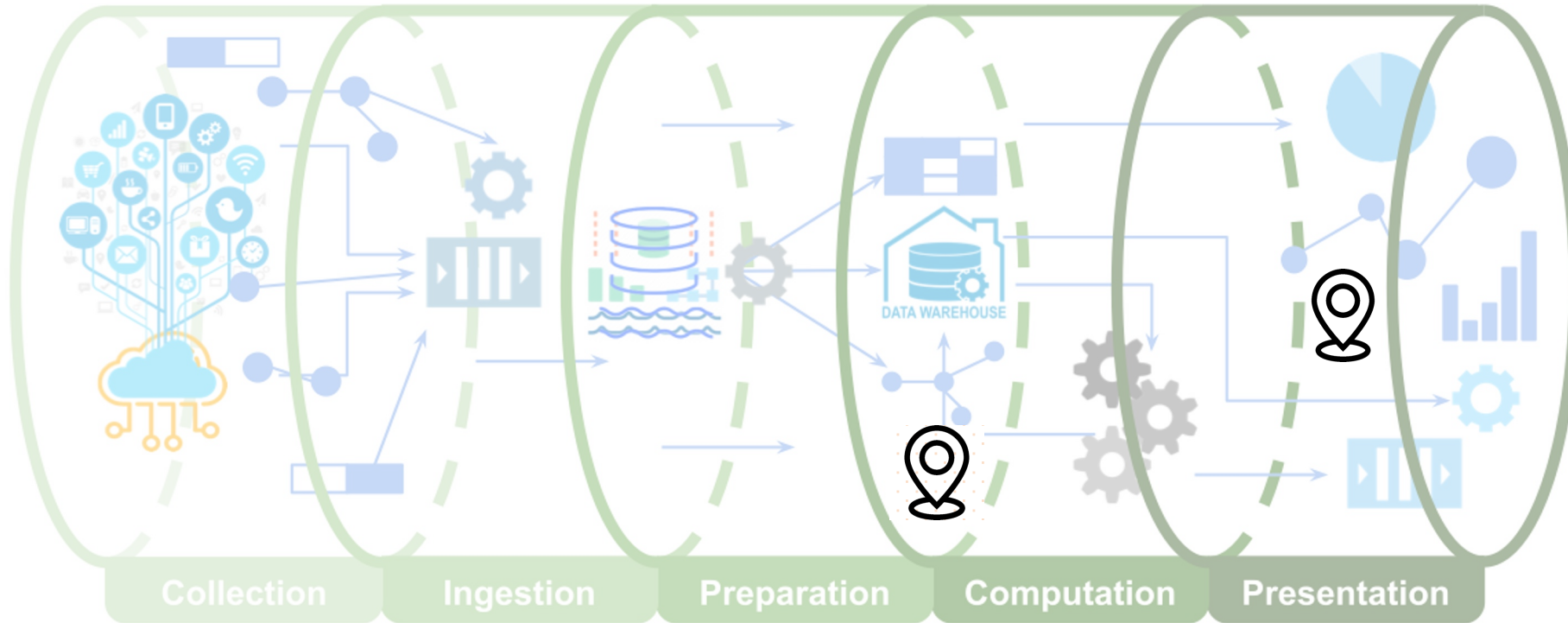
© Satish Chandra Gupta

@scgupta



Место тестировщика

Архитектура продуктов в Big Data



© Satish Chandra Gupta

@scgupta



Место тестировщика

Что помнит тестировщик?

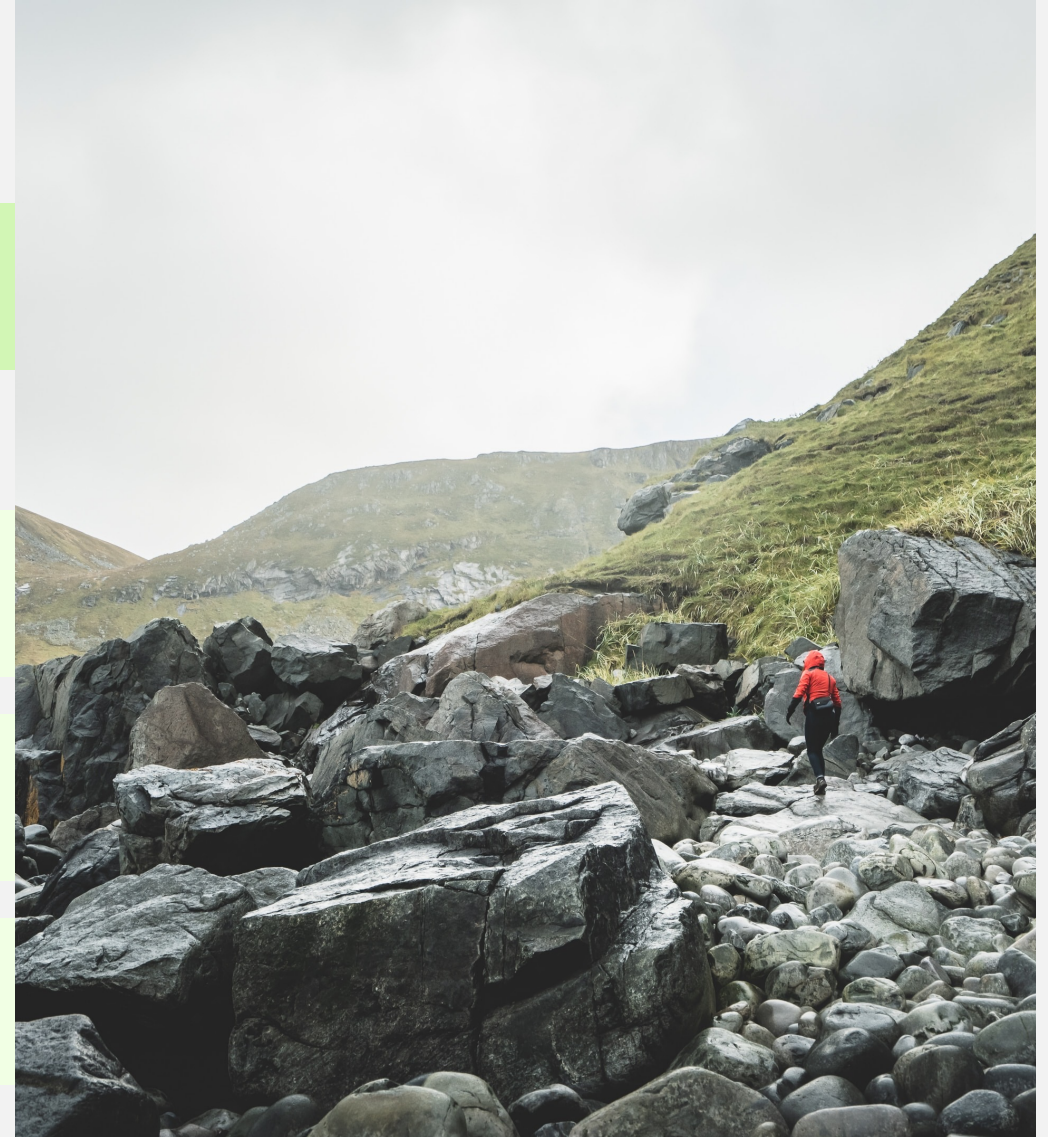
Не доступна часть

Источников Расчетов Данных

Скорость доставки влияет на качество

Наличие соседей

Пользователю не нужны расчеты



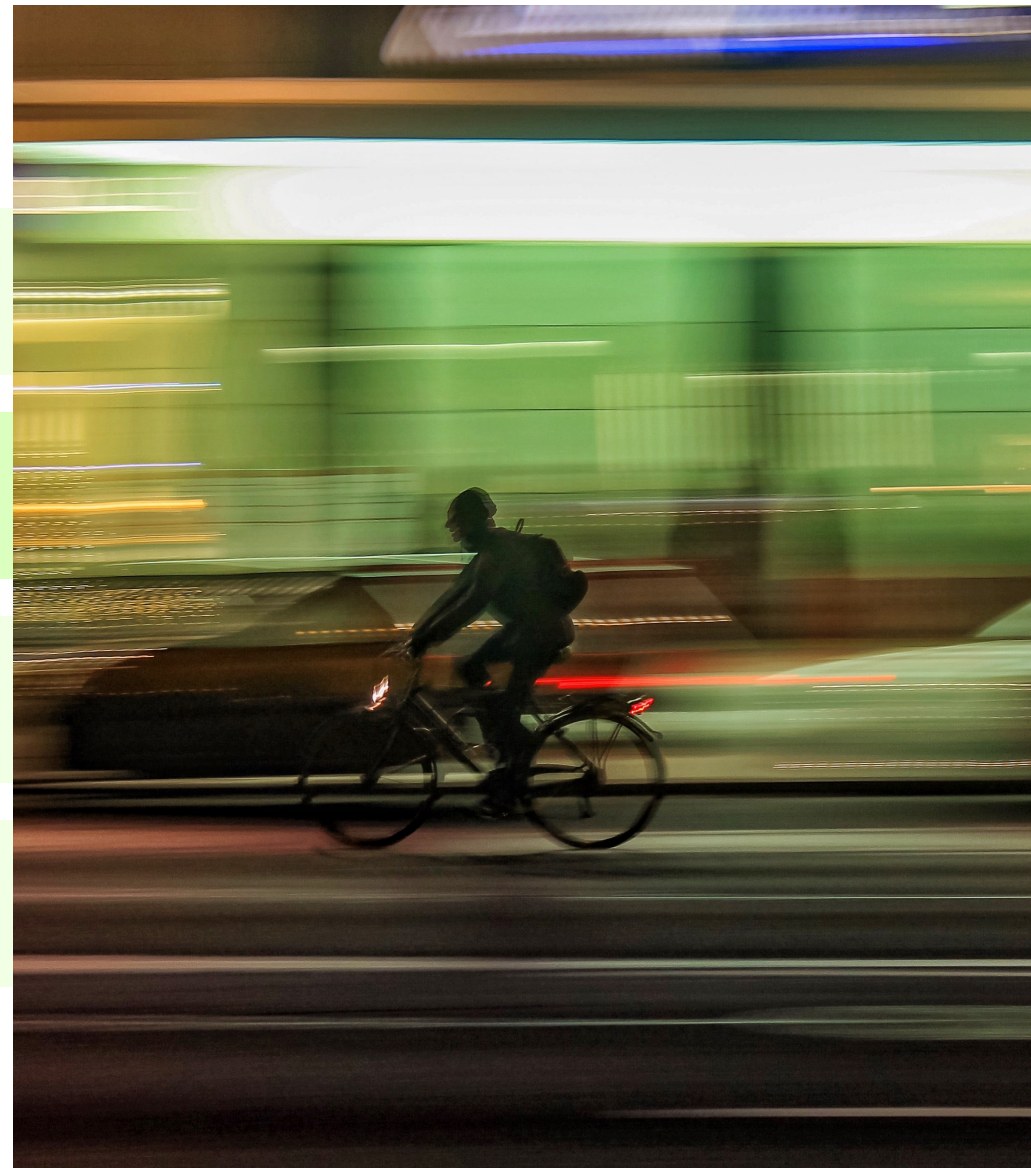
Что помнит тестировщик?

Не доступна часть

Скорость доставки влияет на качество

Наличие соседей

Пользователю не нужны расчеты



Что помнит тестировщик?

Не доступна часть

Скорость доставки влияет на качество

Наличие соседей

Пользователю не нужны расчеты



Что помнит тестировщик?

Не доступна часть

Скорость доставки влияет на качество

Наличие соседей

Пользователю не нужны расчеты



Коротко о главном

Ошибки в ВД динамичны

Тестирование есть на всех слоях,
а тестировщики - нет

Фокус на последние слои



Как тестировать данные?



Немного теории

Выборка

Характеристики выборки

Ошибки 1-2 рода



Немного теории

Выборка

Характеристики выборки

Ошибки 1-2 рода

Объем

Репрезентативность

Механика построения

Немного теории

Выборка

Характеристики выборки

Ошибки 1-2 рода

Он виновен

H

Виноват не
виновный

1

Не виноват
виновный

2

Отличайте причину и следствие



Логика расчётов

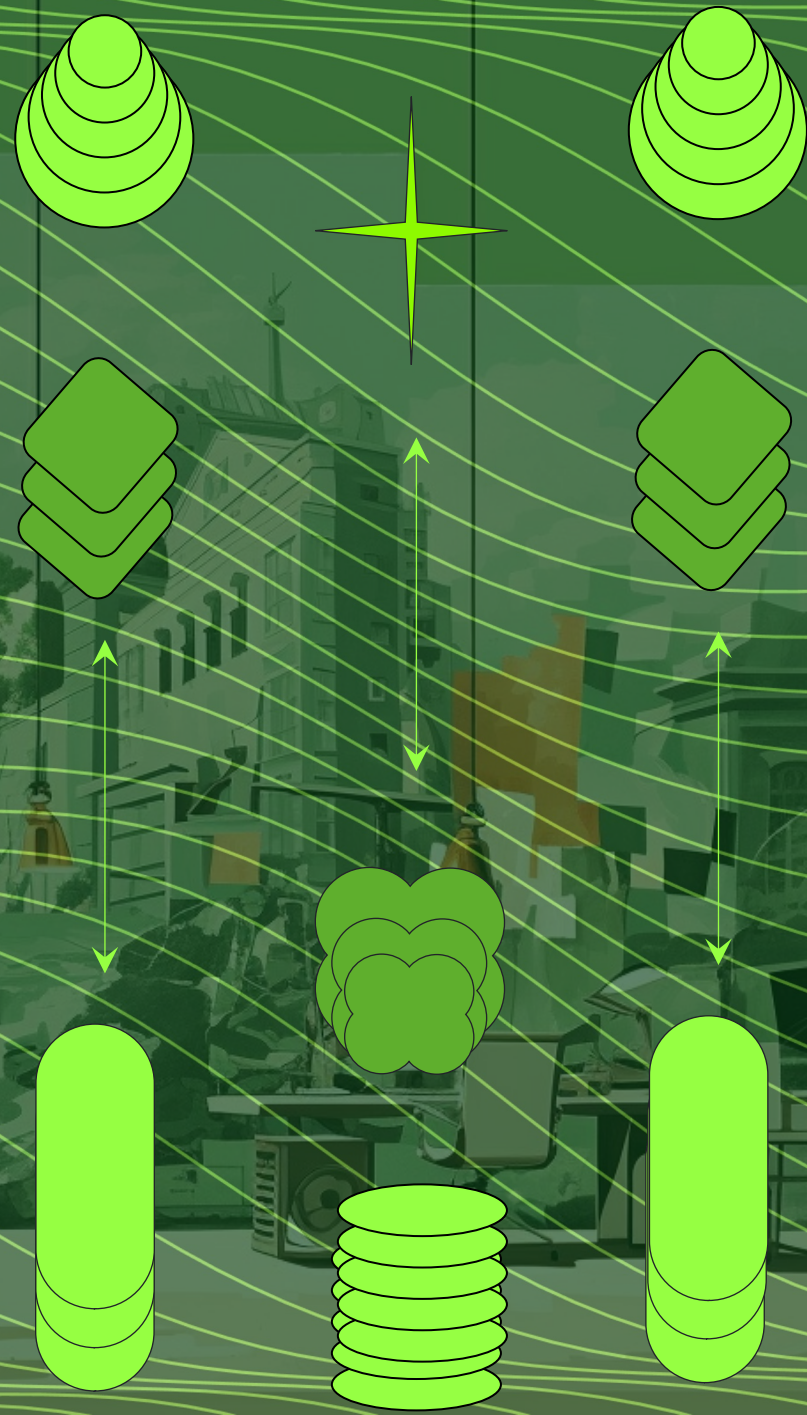


Продукты с ML -
моделями



Продукты с
традиционными
расчётами

Продукты с ML моделями



Что помнит тестировщик?

Доверие к результатам

Применимость модели в реальных условиях

Информативность для бизнес-процесса

Прогноз или
Корректировка?

Прозрачность
«слабых мест»

Для
команды

Для
пользователя

Что помнит тестировщик?

Доверие к результатам

Применимость модели в реальных условиях

Информативность для бизнес-процесса

Мониторинг

Реальный мир

Пользовательский
портрет

Что помнит тестировщик?

Доверие к результатам

Применимость модели в реальных условиях

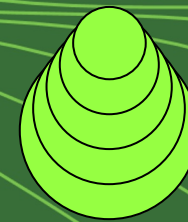
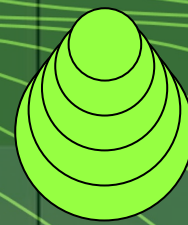
Информативность для бизнес-процесса

Push to action

Кастомизация результата

Устойчивость к adversarial атакам

Продукты с Традиционными расчетами



Что помнит тестировщик?

Оркестраторы

Классическое тестирование



Выбор тестовых данных



Что помнит тестировщик?

Проверка на стабах/моках

Проверка на реальных данных

Комбинация двух подходов



Немного теории

Тестирование на синтетике

Реальные выборки

Вектор данных

Нулевой вектор

Единичный вектор

Немного теории

Тестирование на синтетике

Реальные выборки

Прошлые данные

Обезличивание

Набор для бизнеса

Выводы



Выводы

Не берите на себя слишком много

Дефекты динамичны

Фокусируйтесь на причине, а не следствии

Вам придется много спрашивать

Формируйте верные ожидания!



Выводы

Не берите на себя слишком много

Дефекты динамичны

Фокусируйтесь на причине, а не следствии

Вам придется много спрашивать

Формируйте верные ожидания!



Выводы

Не берите на себя слишком много

Дефекты динамичны

Фокусируйтесь на причине, а не следствии

Вам придется много спрашивать

Формируйте верные ожидания!



Выводы

Не берите на себя слишком много

Дефекты динамичны

Фокусируйтесь на причине, а не следствии

Вам придется много спрашивать

Формируйте верные ожидания!



Выводы

Не берите на себя слишком много

Дефекты динамичны

Фокусируйтесь на причине, а не следствии

Вам придется много спрашивать

Формируйте верные ожидания!



Вопросы?



Виктория Дежкина

Руководитель команды функционального тестирования

