



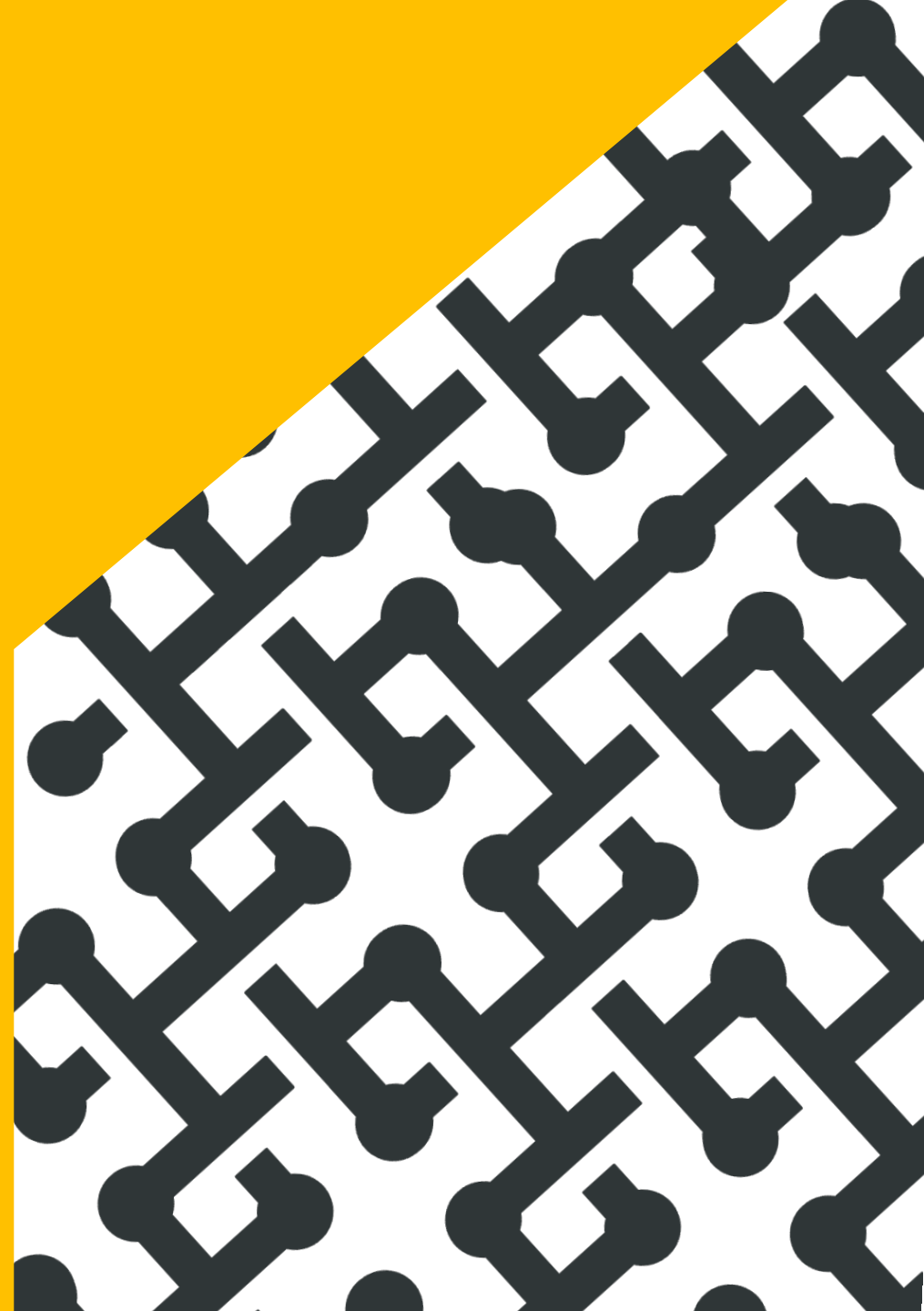
ЛЕМАНА
ПРО



DATA MESH: ОЖИДАНИЯ VS РЕАЛЬНОСТЬ



× SmartData 2024



ОБО МНЕ



- Руководитель практики инженерии данных в Лемана ПРО (ex. Леруа Мерлен)
- Работаю с хранилищами данных более 5 лет
- Выпускник Мехмата МГУ и Бизнес-Информатики НИУ ВШЭ

**Петр
Гуринов**

 petr.gurinov@lemanapro.ru

 t.me/vonirug

ОБО МНЕ



**Петр
Гуринов**



petr.gurinov@lemanapro.ru



t.me/vonirug

	Монолит	Data Mesh
В продукте	X	X
В платформе данных	X	X

СОДЕРЖАНИЕ

Data Mesh

Как у нас

Путь трансформации

Предыстория

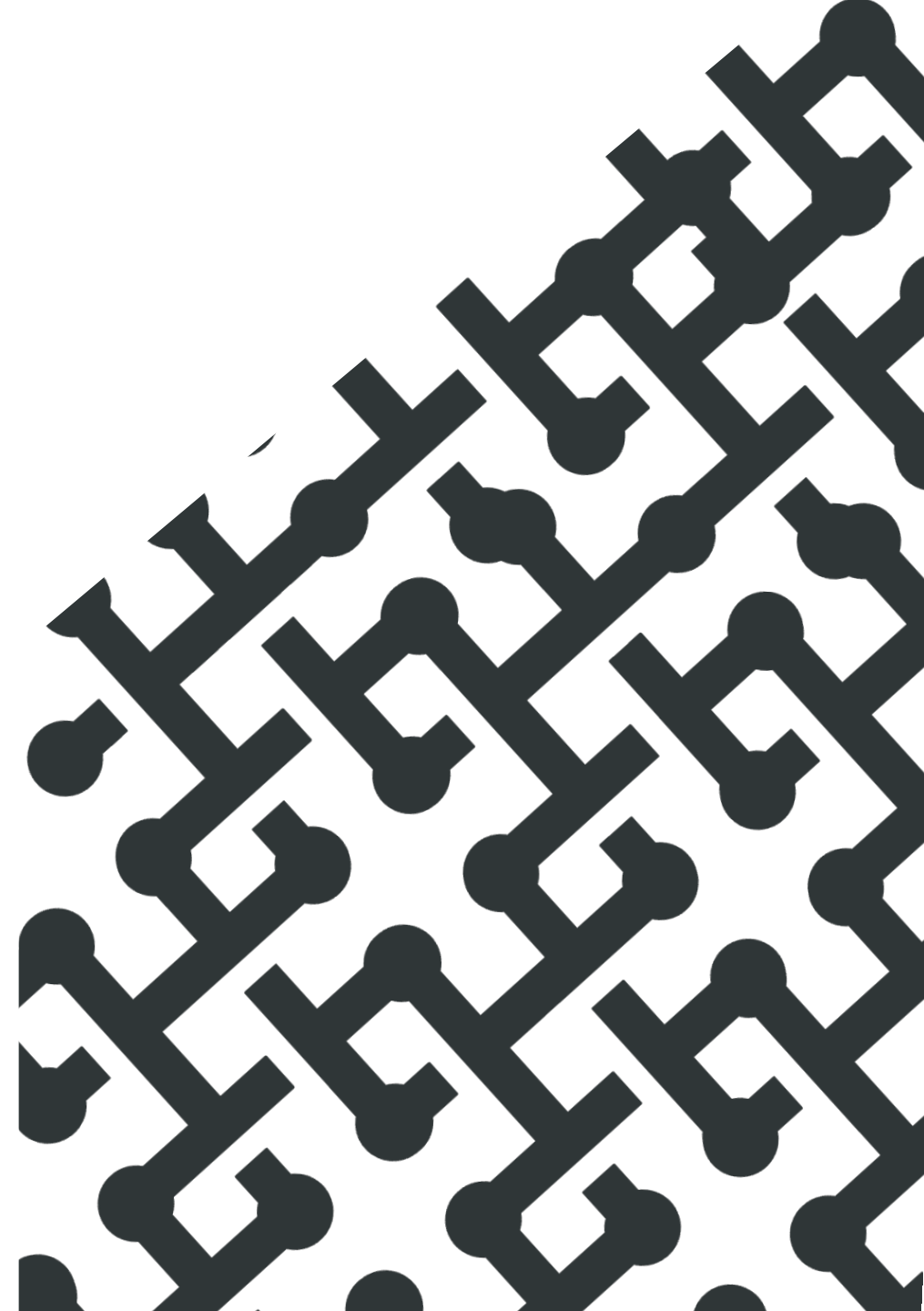
Организационная структура

Принципы и процессы

Технологии и решения

Советы себе 5 лет назад

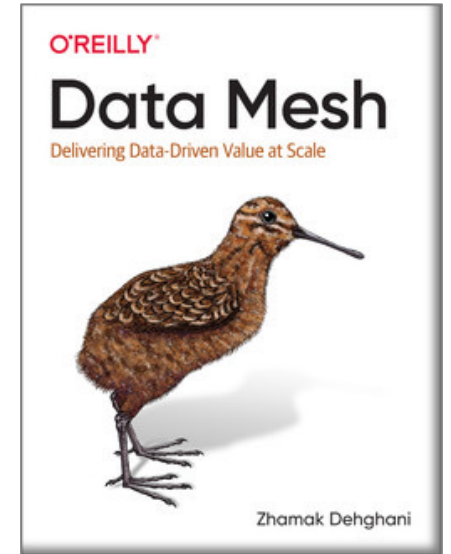
Заключение



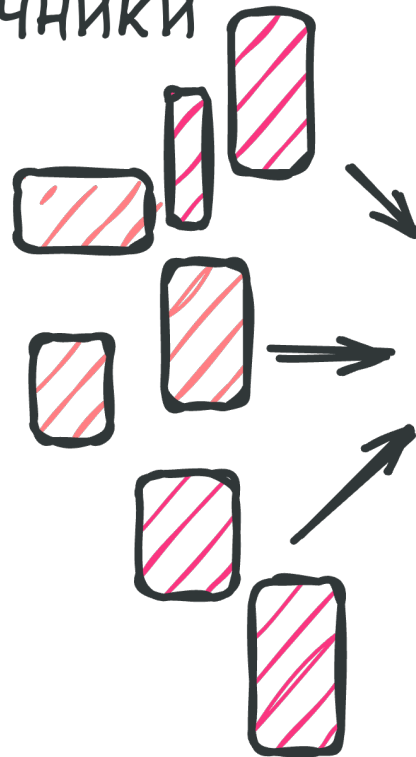
DATA MESH

DATA MESH

- Статья – **20 Мая 2019**
- Децентрализованное управление данными
- Есть рекомендации, но нет готового рецепта

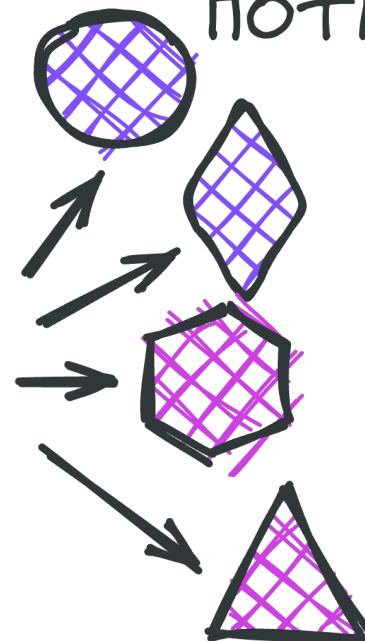


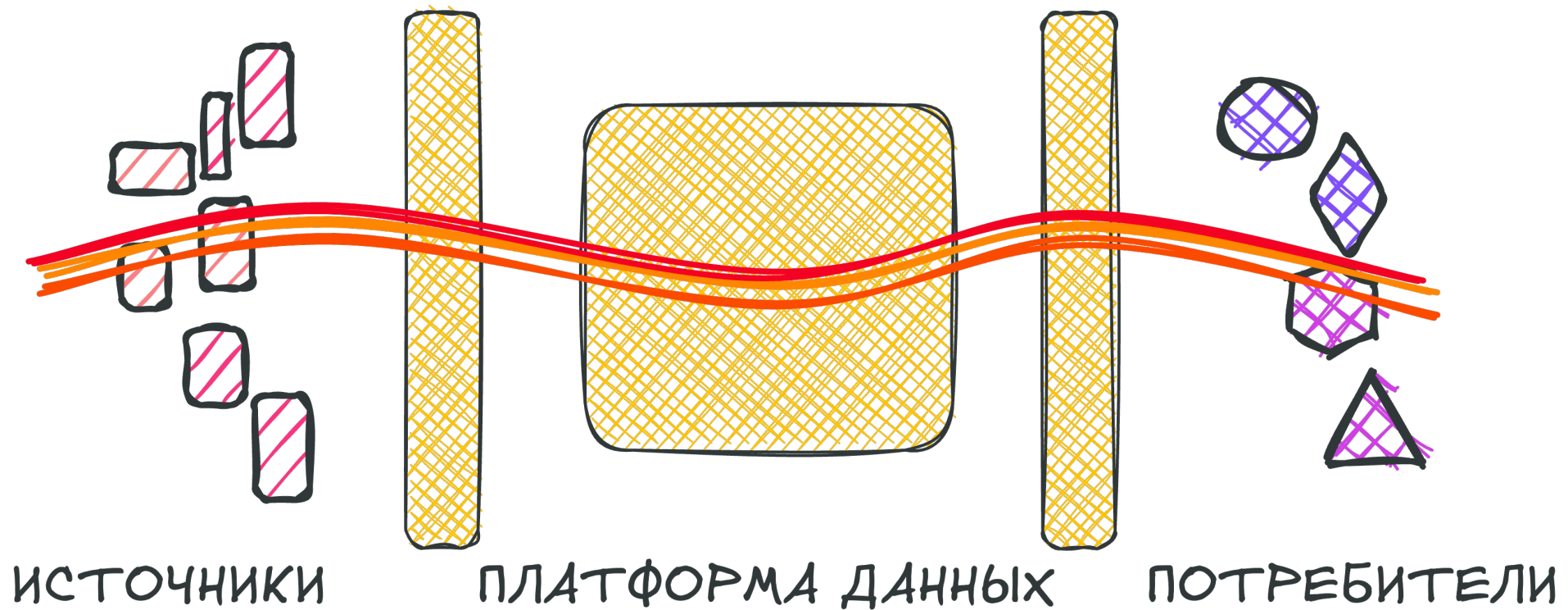
ИСТОЧНИКИ



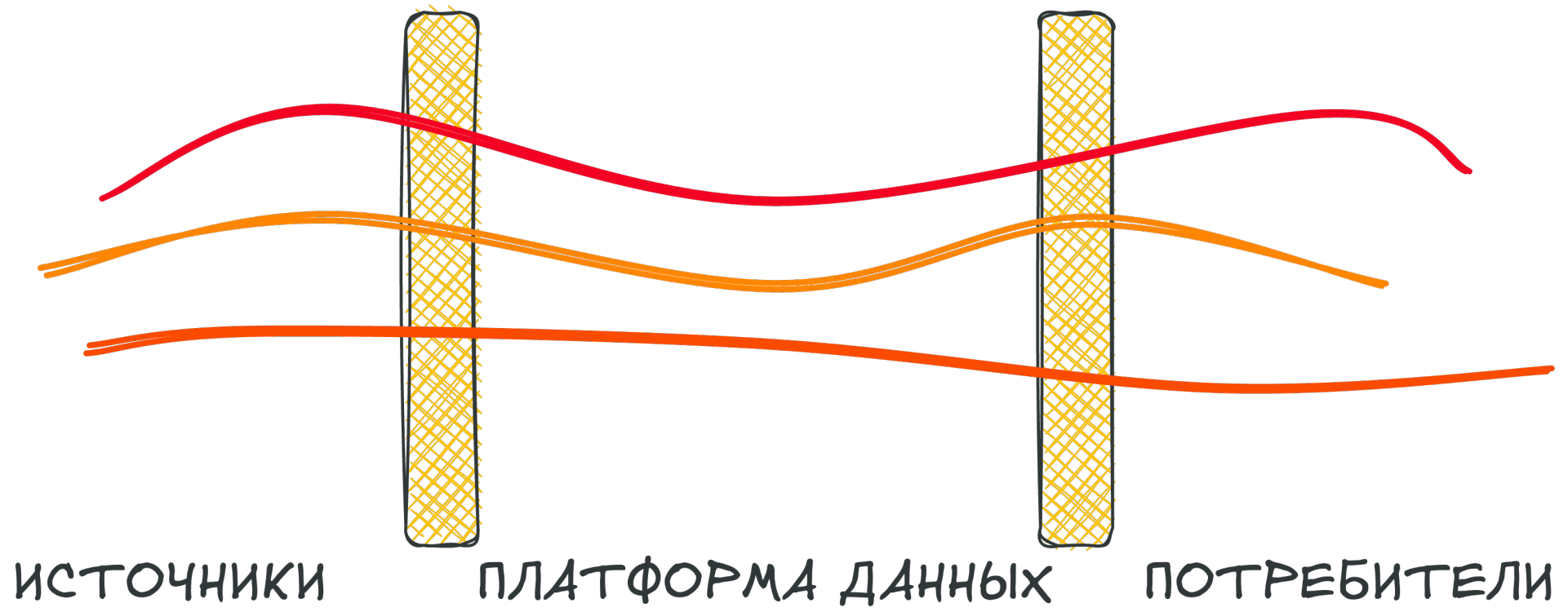
Платформа
данных

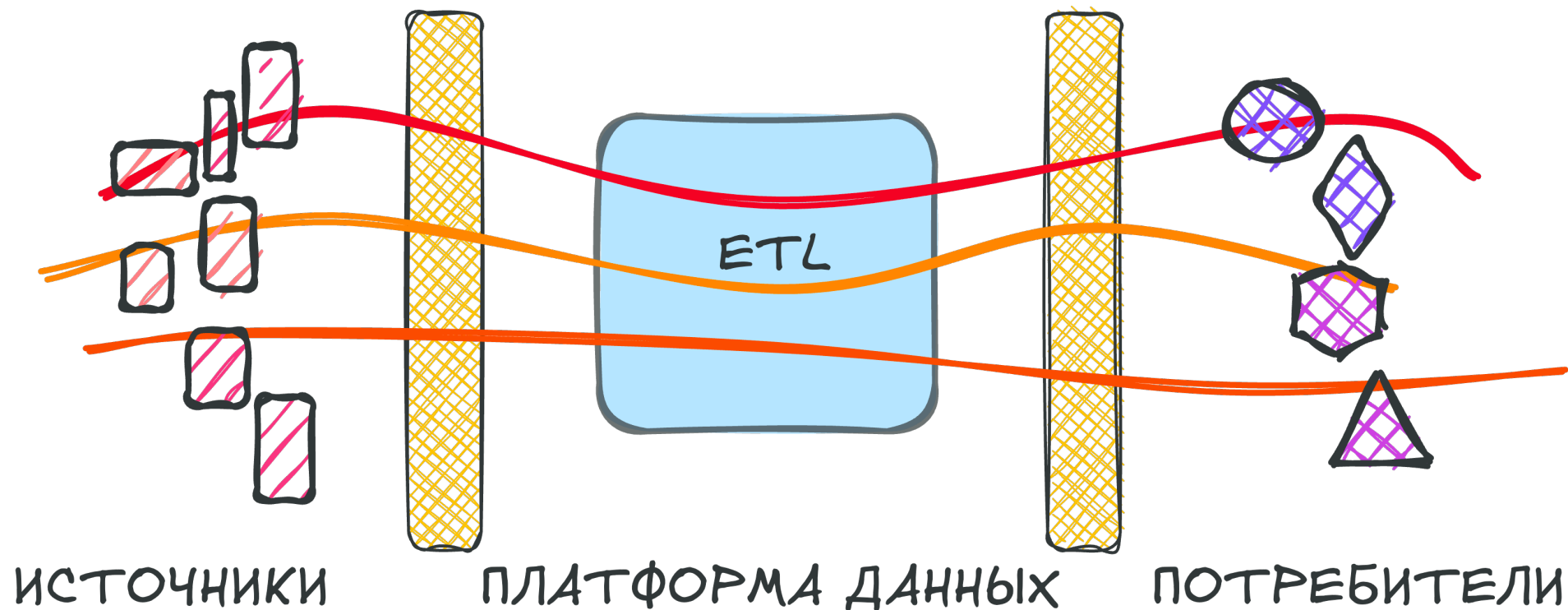
ПОТРЕБИТЕЛИ



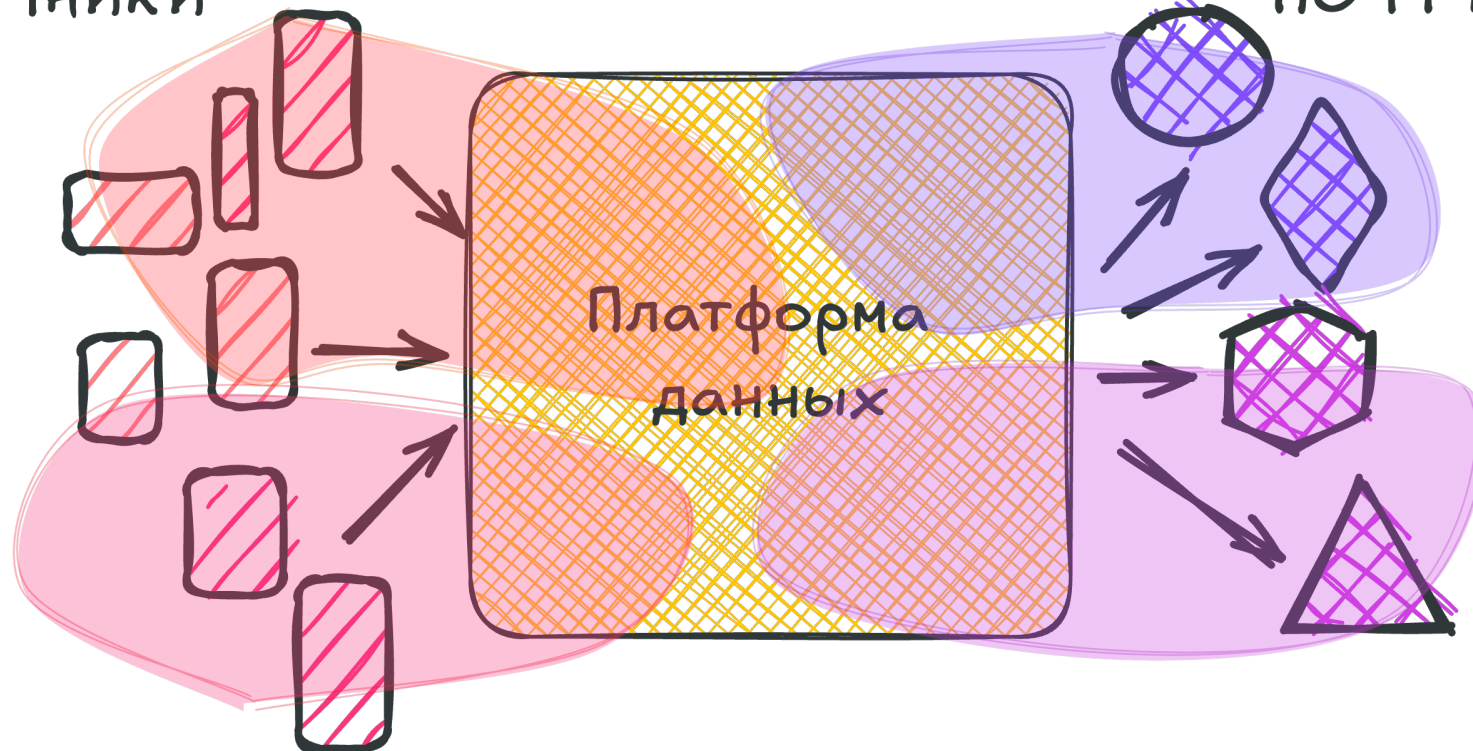




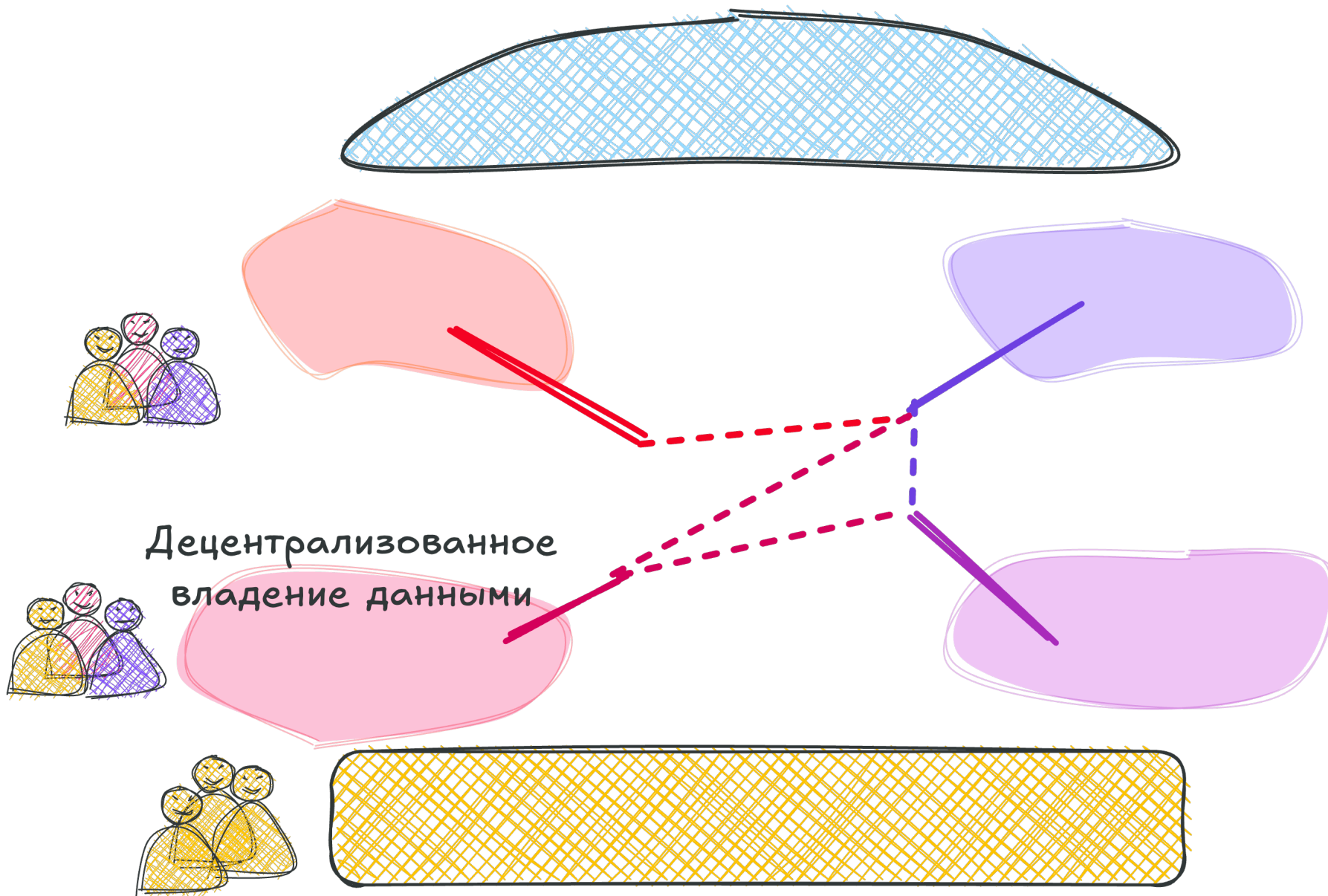


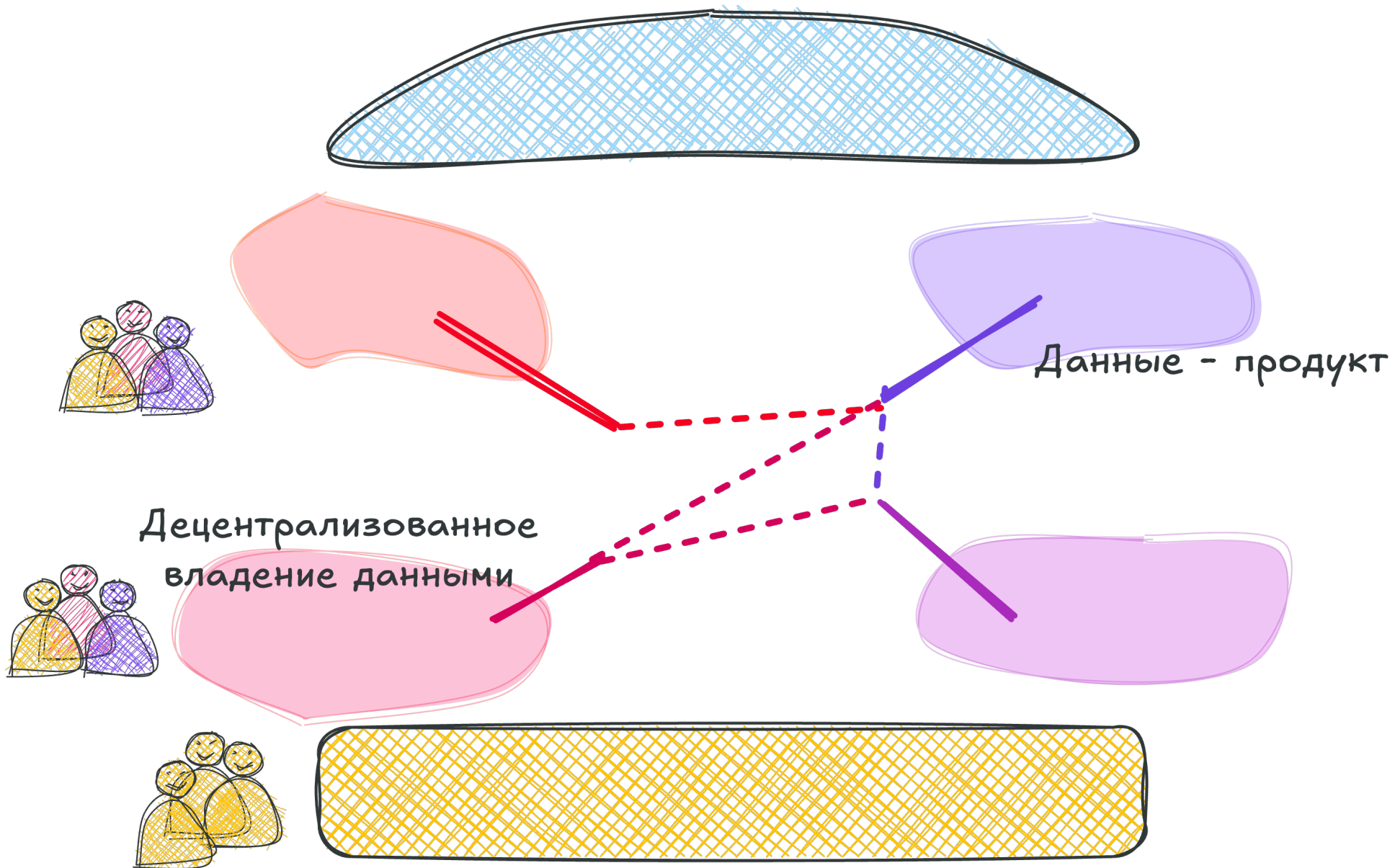


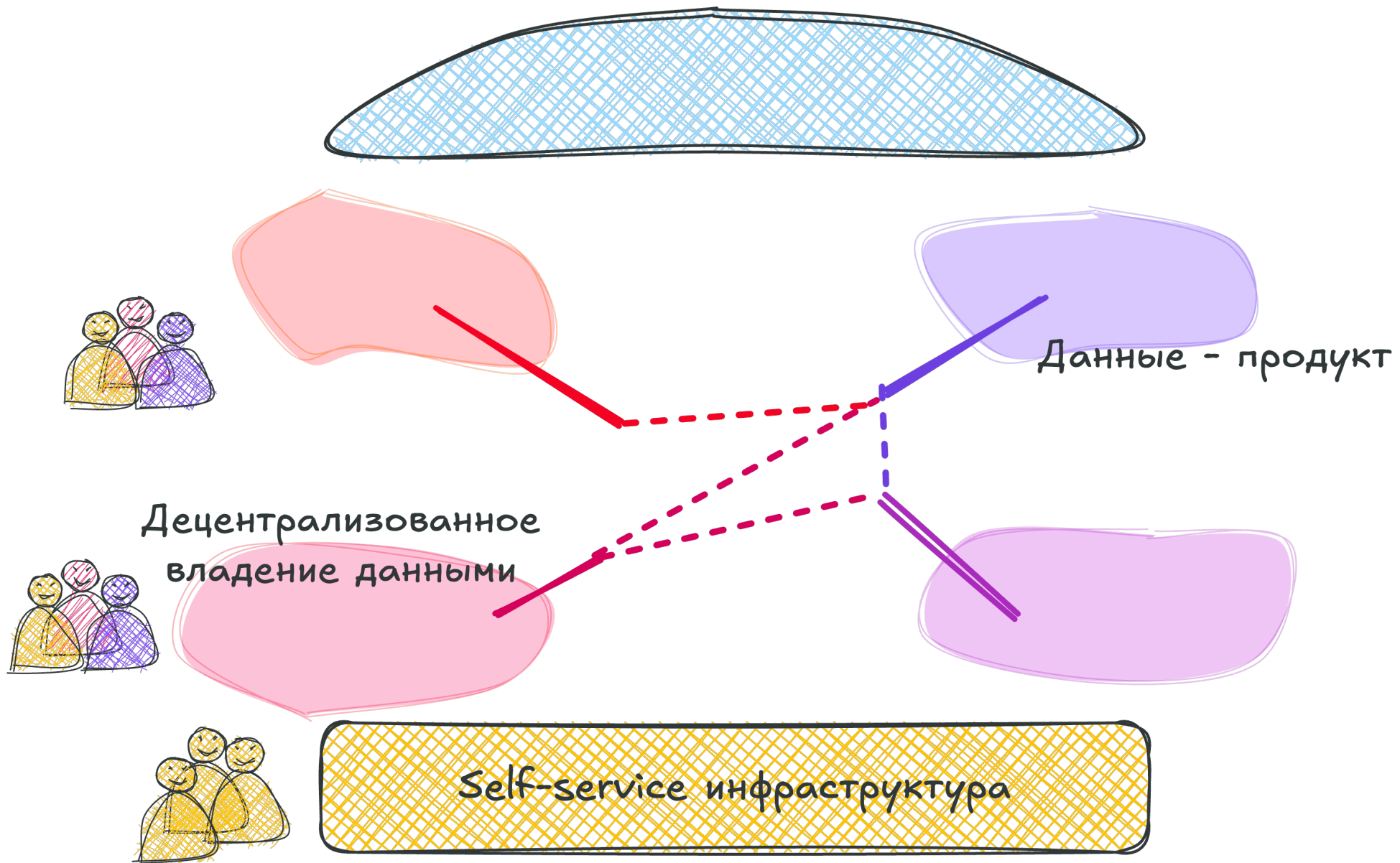
ДОМЕНЫ
ИСТОЧНИКИ

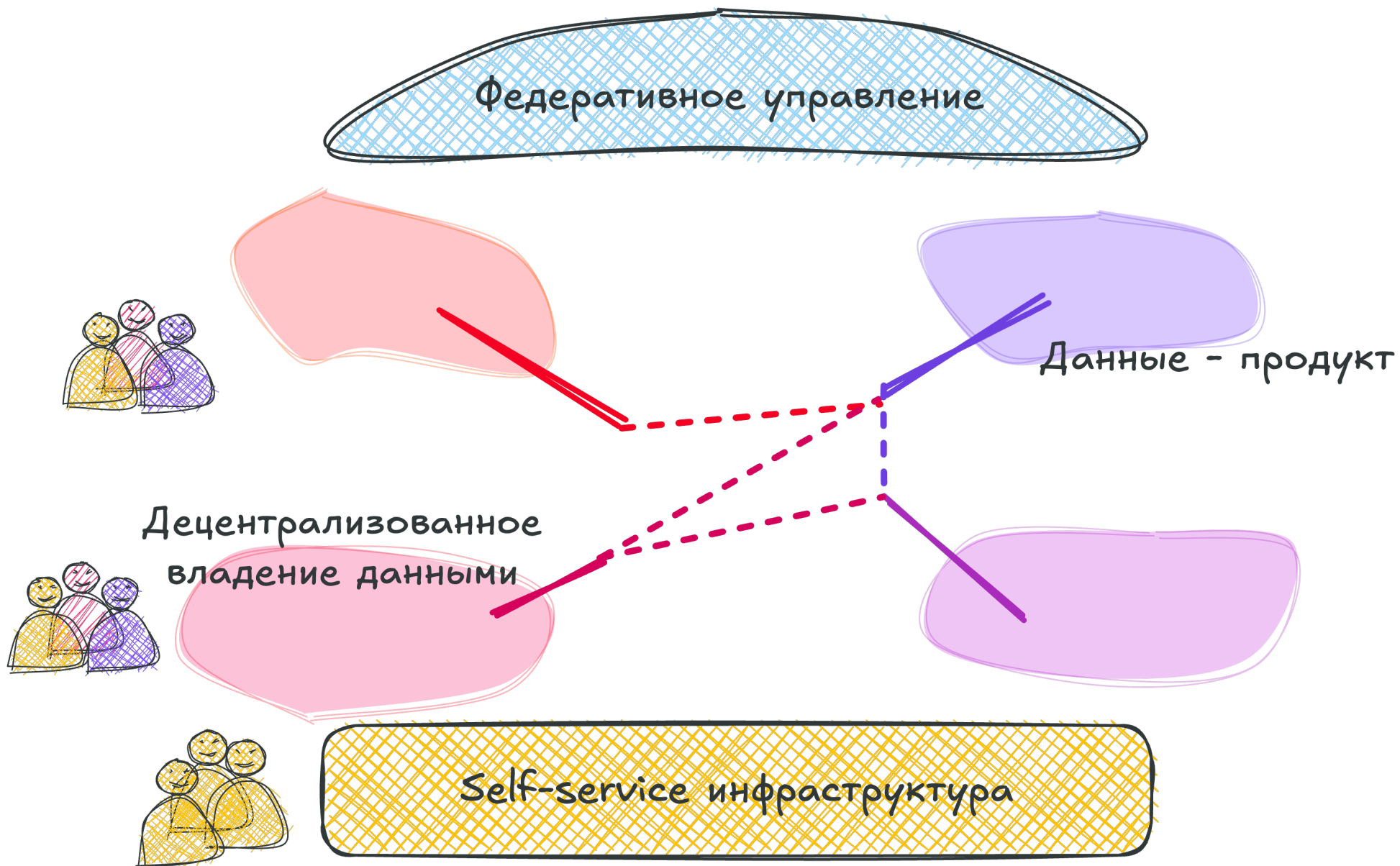


ДОМЕНЫ
ПОТРЕБИТЕЛИ









DATA MESH ДЛЯ НАС

1

Данными владеют
и управляют доменные
команды

3

Коммунальная
платформа данных

2

Домены создают
дата-продукты

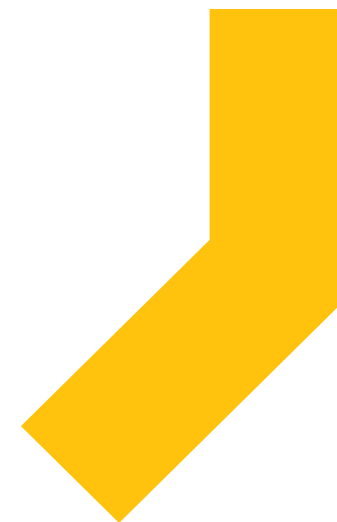
4

Единые правила
в работе с данными

ТРАНСФОРМАЦИЯ КОМПАНИИ



ПРЕДЫСТОРИЯ



КОМПАНИЯ В 2019

- Деление на бизнес и ИТ подразделения
- Подразделение ИТ около 250 человек
- ИТ ландшафт – коробочные решения

DATA В 2019

- Центральная команда DWH из 8 человек + аутстаф
 - Отчет разрабатывается ~6 месяцев
 - ~300 отчетов, DAU 6 000
 - Выгрузки из бизнес-приложений
- Подход – Excel Driven Development

УПРАВЛЕНИЕ ДАННЫМИ В 2019

- Команда DWH сопровождает Datablender (MSSQL) интеграция через dblink
- Огромная любовь к SSRS
- Qlikview – не только BI, но и ETL через csv выгрузки

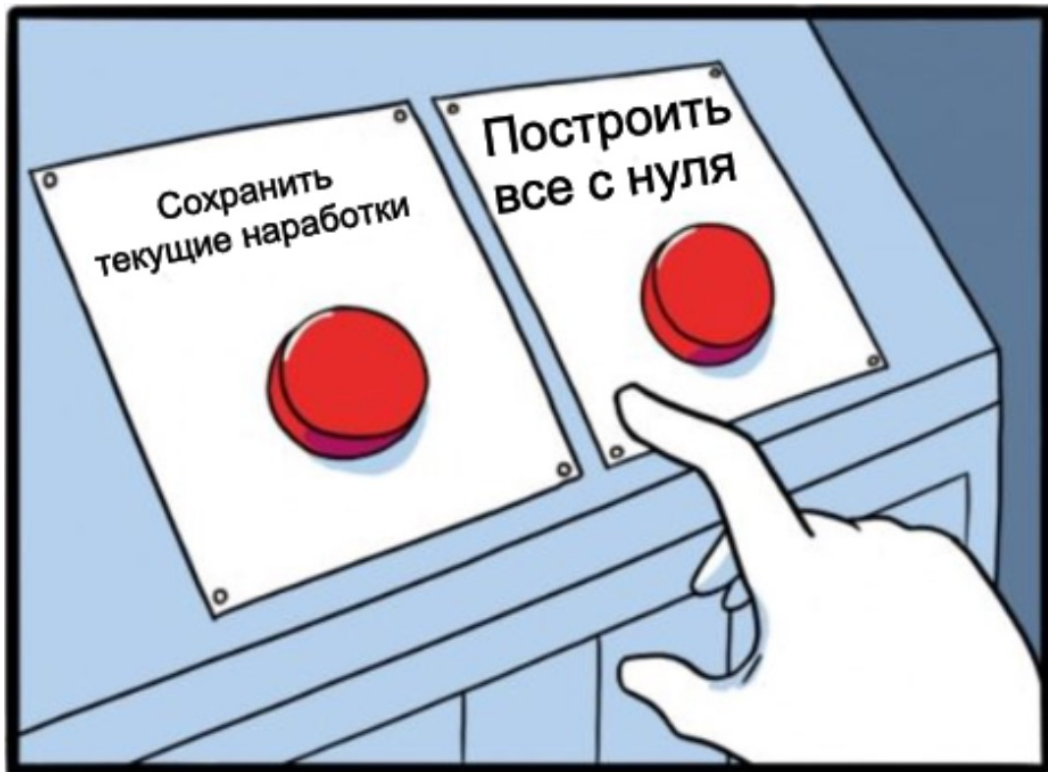
ИТОГ

- У команды DWH увеличивался бэклог, кол-во сопровождаемых отчетов
- Бизнес не понимает, какие данные есть
- Решения принимаются на экспертном уровне

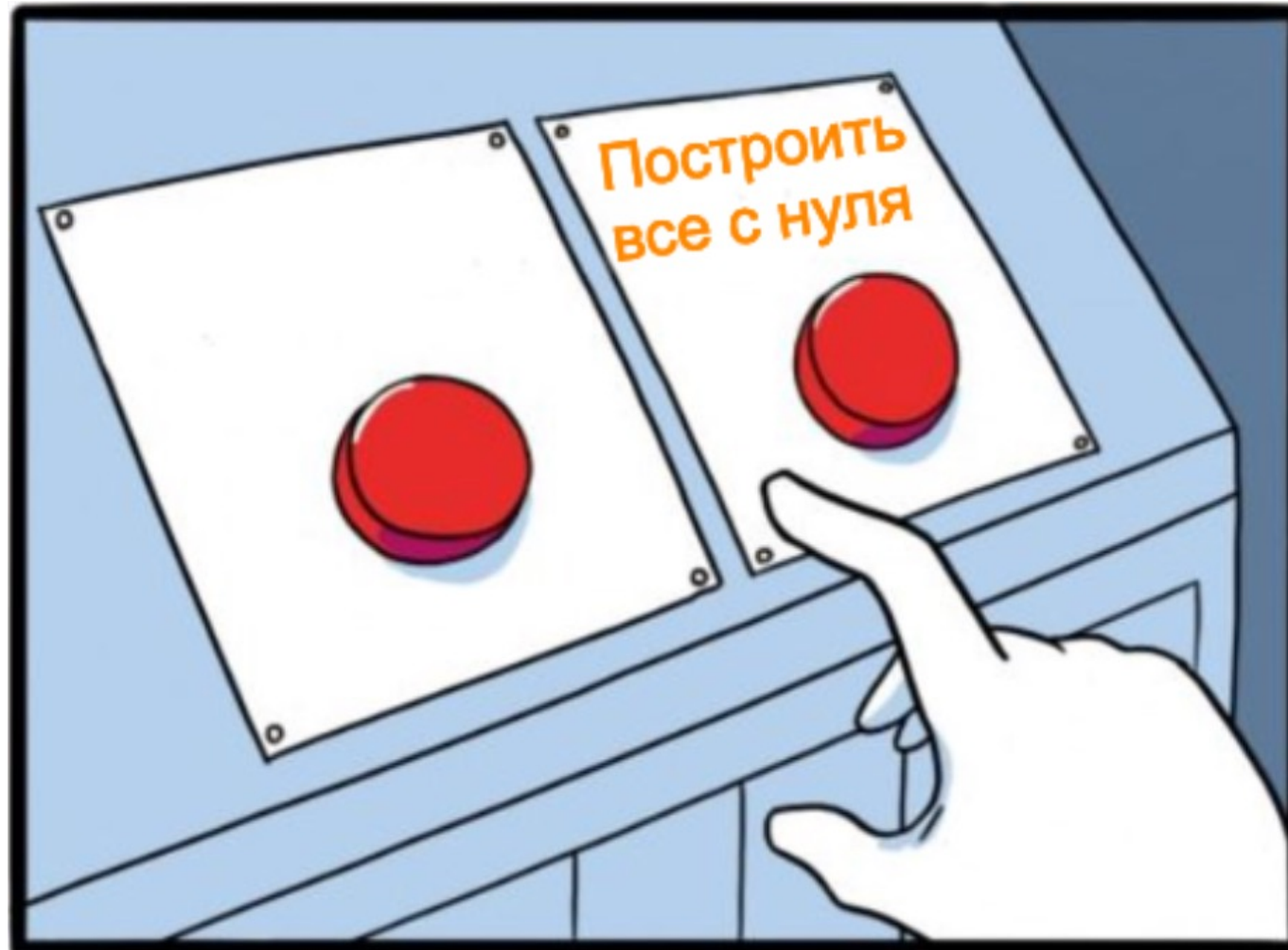
ТРАНСФОРМАЦИЯ КОМПАНИИ

- Изменение в организационной структуре и переход на доменную модель с операционными и продуктовыми командами
- Переход в Data Mesh

ТРАНСФОРМАЦИЯ В DATA MESH



ТРАНСФОРМАЦИЯ В DATA MESH





ТРАНСФОРМАЦИЯ В DATA MESH



НАШ ПЕРЕХОД В DATA MESH

шаги

- Собрать команду платформы данных*
- Сформировать регламенты по работе с DP
- Нанять специалистов в домены
- Интегрировать источники в DP
- Внедрить KPI по доступности источников в DP

* Платформа данных – Data platform – DP

НАШ ПЕРЕХОД В DATA MESH

области

- Организационная структура
- Принципы и процессы
- Технологии и решения



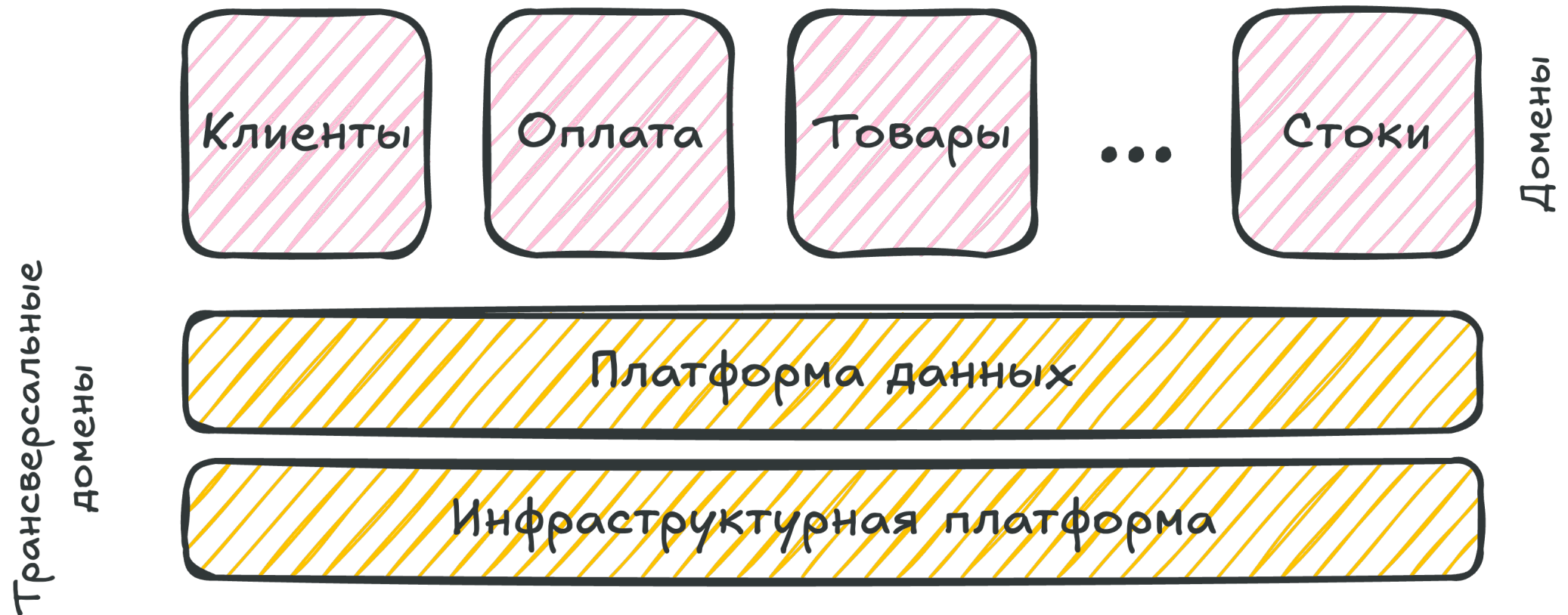
ОРГАНИЗАЦИОННАЯ СТРУКТУРА



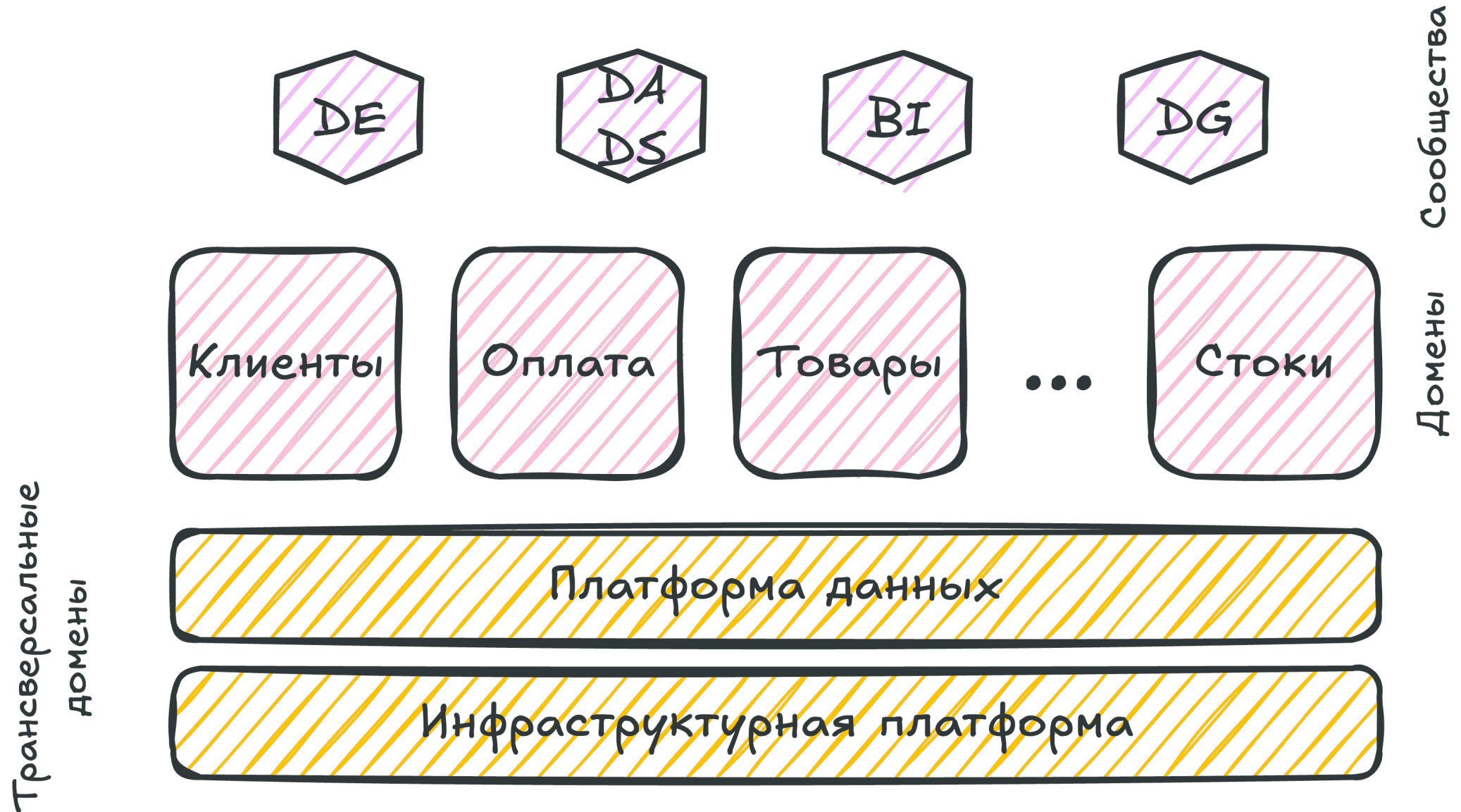
ОРГАНИЗАЦИОННАЯ СТРУКТУРА



ОРГАНИЗАЦИОННАЯ СТРУКТУРА



ОРГАНИЗАЦИОННАЯ СТРУКТУРА



ОРГАНИЗАЦИОННАЯ СТРУКТУРА

Ожидание:

- Домен будет ответственен за бизнес-процесс, приложение и дата-продукты

ОРГАНИЗАЦИОННАЯ СТРУКТУРА

Ожидание:

- Домен будет ответственен за бизнес-процесс, приложение и дата-продукты

Реальность:

- Мастер системы управляются другими командами
- Оказались домены без данных в DP
- Домены без специалистов по работе с данными

ОРГАНИЗАЦИОННАЯ СТРУКТУРА

Выводы:

- Монолиты – ЗЛО
- Нужно обучать не только технологиям, но и другому стилю менеджмента
- Enterprise Architect* – ваши лучшие друзья для поиска правильного решения

* Enterprise Architect – архитектор предприятия



ПРИНЦИПЫ И ПРОЦЕССЫ



ПРИНЦИПЫ ИТ

- Команды в доменах – кроссфункциональны
- You Build it, you Run it
- Короткие релизные циклы
- InnerSource *

- InnerSource — использование подходов open source внутри корпоративных границ
<https://leroymerlin-tech.ru/knowledge/innersource/>

ПРИНЦИПЫ ИТ

Ожидание:

- Домены будут «знать свои данные»
- Домены будут участвовать в развитии DP

ПРИНЦИПЫ ИТ

Ожидание:

- Домены будут «знать свои данные»
- Домены будут участвовать в развитии DR

Реальность:

- Не все команды понимают принцип «данные как продукт»
- Кроссфункциональны настолько, что строят свои велосипеды

ПРИНЦИПЫ ИТ

Выводы:

- Помогать доменам с выявлением потребностей
- Популяризировать DR, много, постоянно
- DR – такой же продукт

ПРОЦЕССЫ

- Определение потребностей по специалистам в домене
- Найма
- Онбординга
- Работа с сообществами
- Разработки / Ревью
- Аудита

ПРОЦЕССЫ

Ожидание:

- Один человек из Центра Компетенций сможет настроить все процессы

ПРОЦЕССЫ

Ожидание:

- Один человек из Центра Компетенций сможет настроить все процессы

Реальность:

- Часть процессов остается без внимания

ПРОЦЕССЫ

Выводы:

- Привлечь специалистов из доменов для выстраивания регламентов



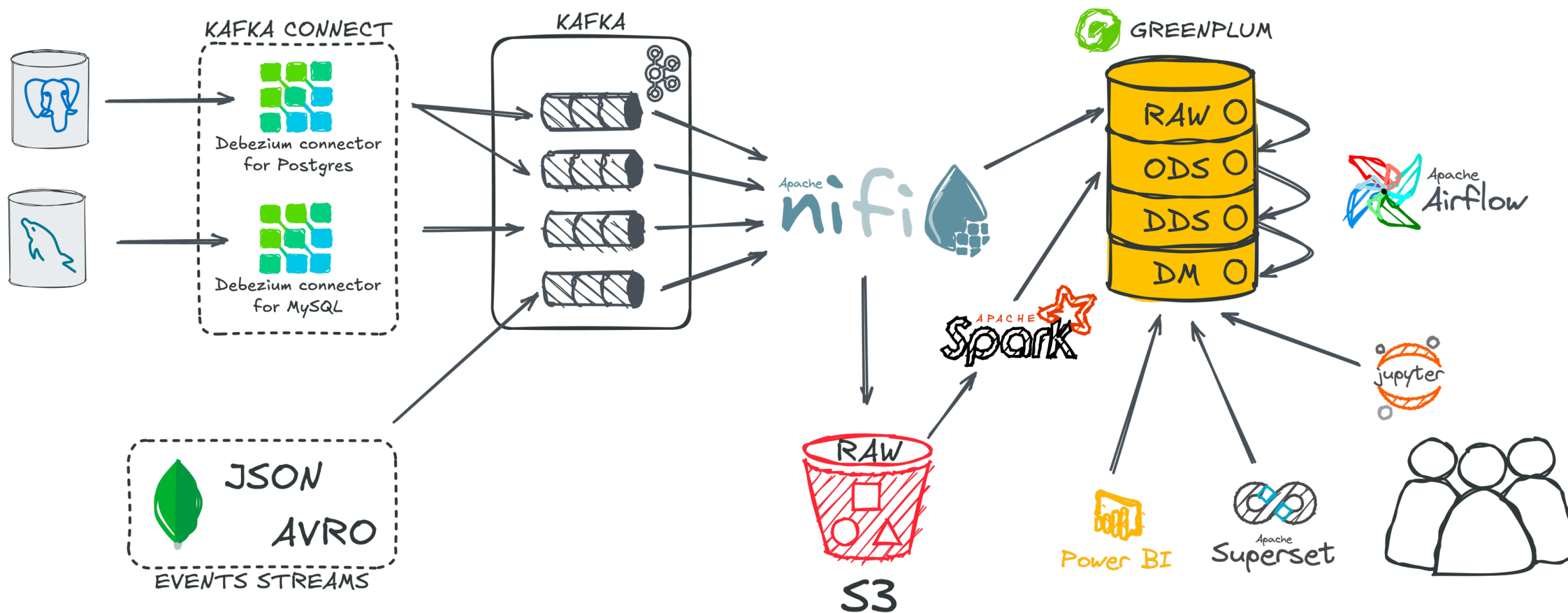
ТЕХНОЛОГИИ И РЕШЕНИЯ



ТЕХНОЛОГИЧЕСКИЕ ПРИНЦИПЫ

- Масштабируемость
- Cloud ready / Cloud native / Cloud Agnostic
- Open Source
- Внутренняя разработка сервисов

КОММУНАЛЬНАЯ ПЛАТФОРМА ДАННЫХ



ПЛАТФОРМЕННЫЕ СЕРВИСЫ

- Описание данных
- Качество данных
- Аудит доступа к платформе данных
- Единообразная интеграция данных
- Управление справочниками
- «Отстрел» зависших сессий
- Создание семплов данных для разработки

ТЕХНОЛОГИИ И РЕШЕНИЯ

Ожидание:

- Облака – boost технологического развития
- Сокращаем время сопровождения через self-service сервисы

ТЕХНОЛОГИИ И РЕШЕНИЯ

Ожидание:

- Облака – boost технологического развития
- Сокращаем время сопровождения через self-service инструменты

Реальность:

- Упираемся в локальную инфраструктуру
- Open Source ломается в самый неподходящий момент

ТЕХНОЛОГИИ И РЕШЕНИЯ

Выводы:

- Учиться «готовить» Open Source самостоятельно
- Fullstack разработчики помогут быстро разрабатывать внутренние сервисы
- Команда платформы данных не менее 20% от общего кол-ва дата-специалистов



ГДЕ МЫ СЕЙЧАС




DATA В 2024


- Более 100 специалистов по работе с данными в компании + аутстаф
- Интеграция данных от 2 недель
- Более 1 500 отчетов, DAU 12 500
- Домены самостоятельно создают дата-продукты
- Подход – Data Informed

СОВЕТЫ СЕБЕ 5 ЛЕТ НАЗАД

- Трансформации позволяют развиваться быстрее
- Заложить время на коммуникации и взаимодействие
- Активнее развивать self-service инструменты
- Регулярно проводить обучения
- Привлекать дата-специалистов из доменов



**Централизованный дата офис
может быть эффективен для
небольших компаний с
ограниченным объемом данных,
тогда как Data Mesh подходит для
крупных организаций с
множеством доменов**



ССЫЛКИ

- [How to Move Beyond a Monolithic Data Lake to a Distributed Data Mesh](#)
- [Data Mesh Principles and Logical Architecture](#)
- [Data Mesh by Zhamak Dehghani](#)

- [Introduction to Data Mesh with Zhamak Dehghani](#)
- [Дмитрий Шостко: Как построить Data Mesh в организации](#)
- [Евгений Ермаков: Как с помощью Data Mesh разломать ваше DWH](#)



18:30 – 19:30



ОБСУЖДЕНИЕ

Offline only

~~REC~~

Зал 2

Круглый стол: Data Mesh



Наталья Хапаева LamodaTech, МТС



Дмитрий Руднев Т-Банк



Дмитрий Ибрагимов Лемана ПРО (Л...



Петр Гуринов Лемана ПРО (Леруа Ме...



Сергей Муханов Райффайзен Банк



Дмитрий Зуев Positive Technologies

RU

ВОПРОСЫ

Петр Гуринов

 petr.gurinov@lemanapro.ru

 t.me/vonirug