

Исследование состояния DATA в России

a.k.a The-State-of-Data Russian edition

Олег Кочергин

whoami

- software development background since '00
- in data since '16
- dev → data engineer → head of data → cdo
- Positive Technologies
ex. Alfa-Bank, ex. SberHealth
- SmartData resident



**Собственно,
про исследование**



Цели исследования

- Понять что из технологий актуально / модно / молодежно / устарело
- Понять интересы аудитории будущих конференций
- Сориентировать новичков и старичков
- Да и это просто интересно :)

Методология

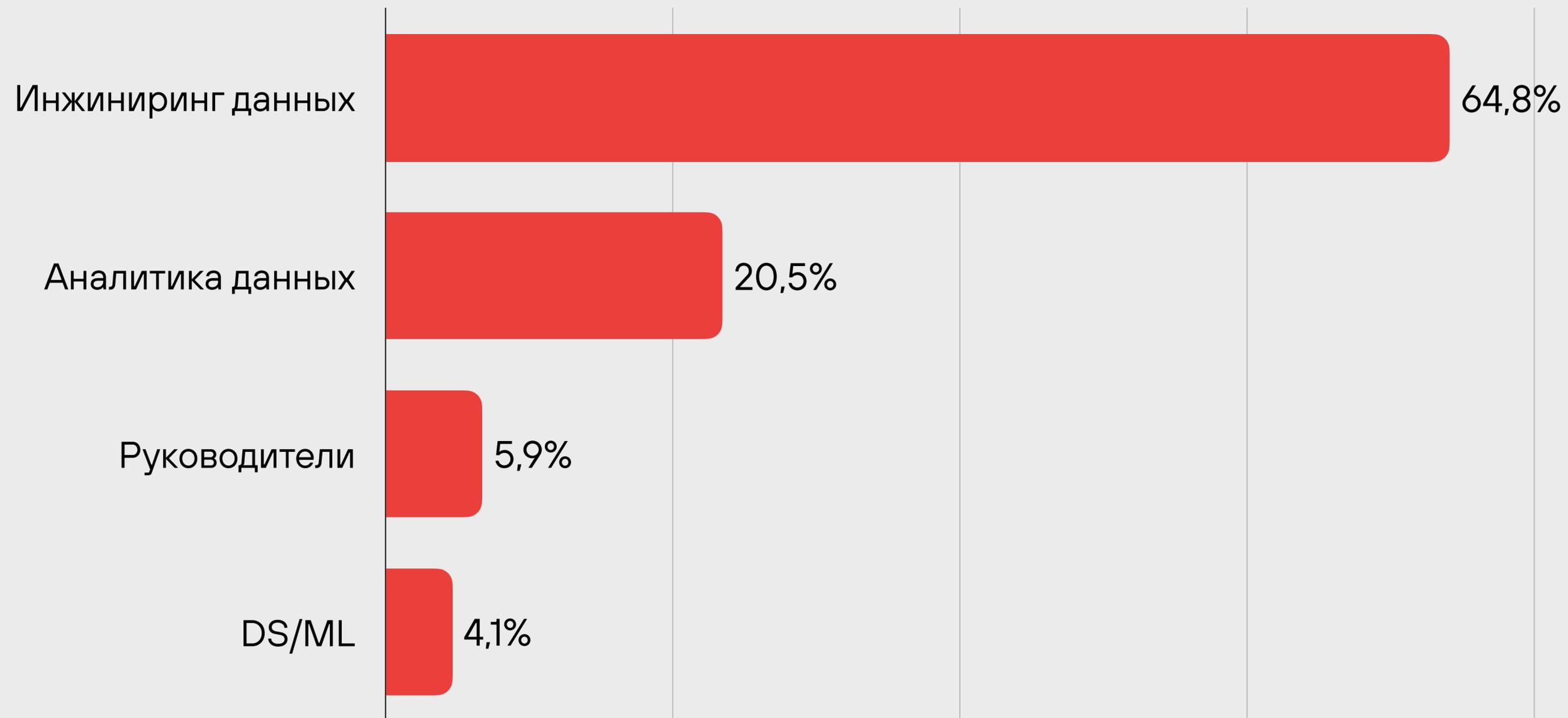
Опрос в гуглоформах из 4 частей:

- Базовый соцдем (специализация, опыт, контекст компании итп)
- Тулинг вокруг задач DE
- Тулинг вокруг задач BI и прочей аналитики
- Тулинг вокруг задач MLOps



Портрет респондента

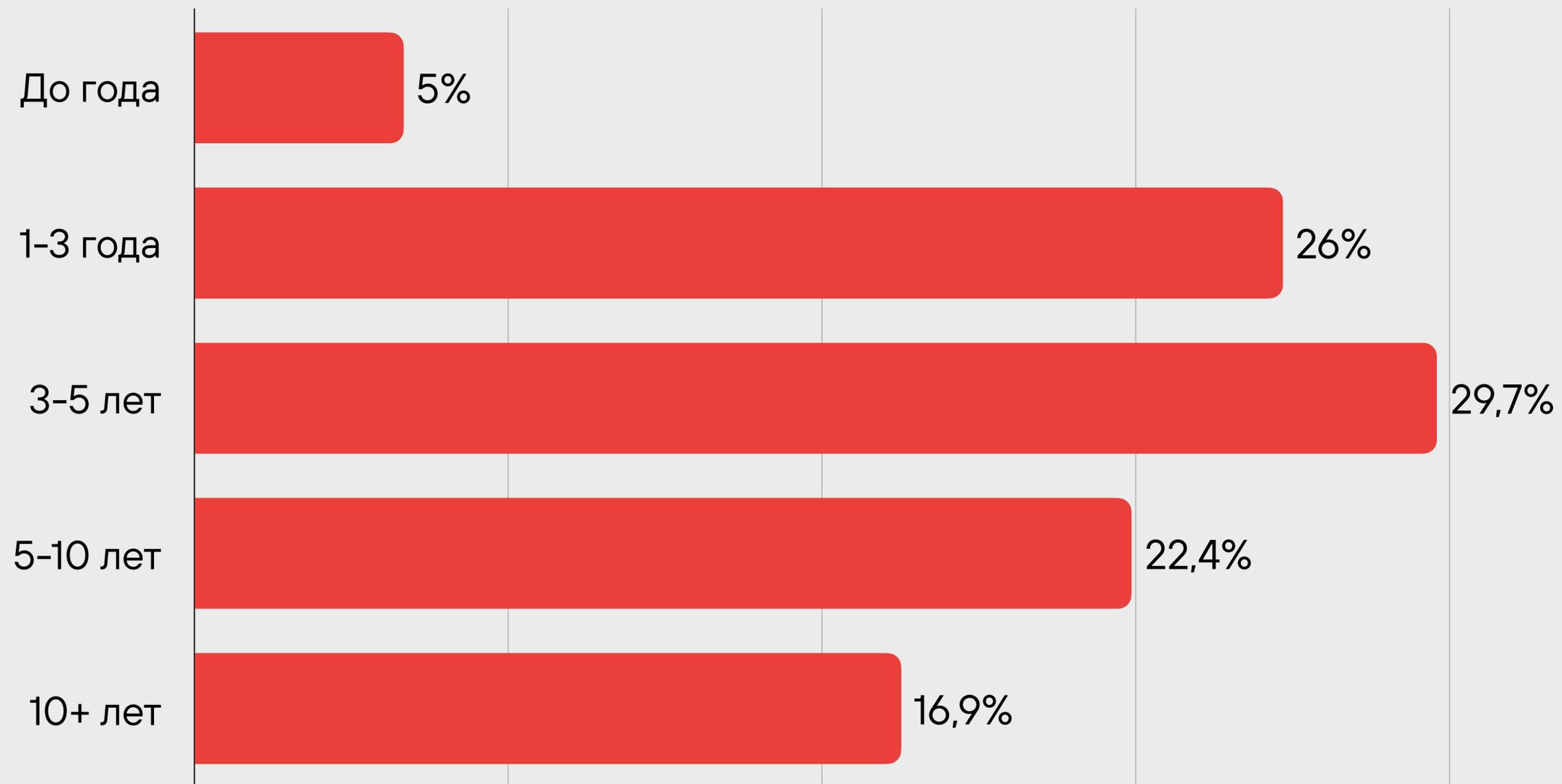
Распределение по специализациям





Портрет респондента

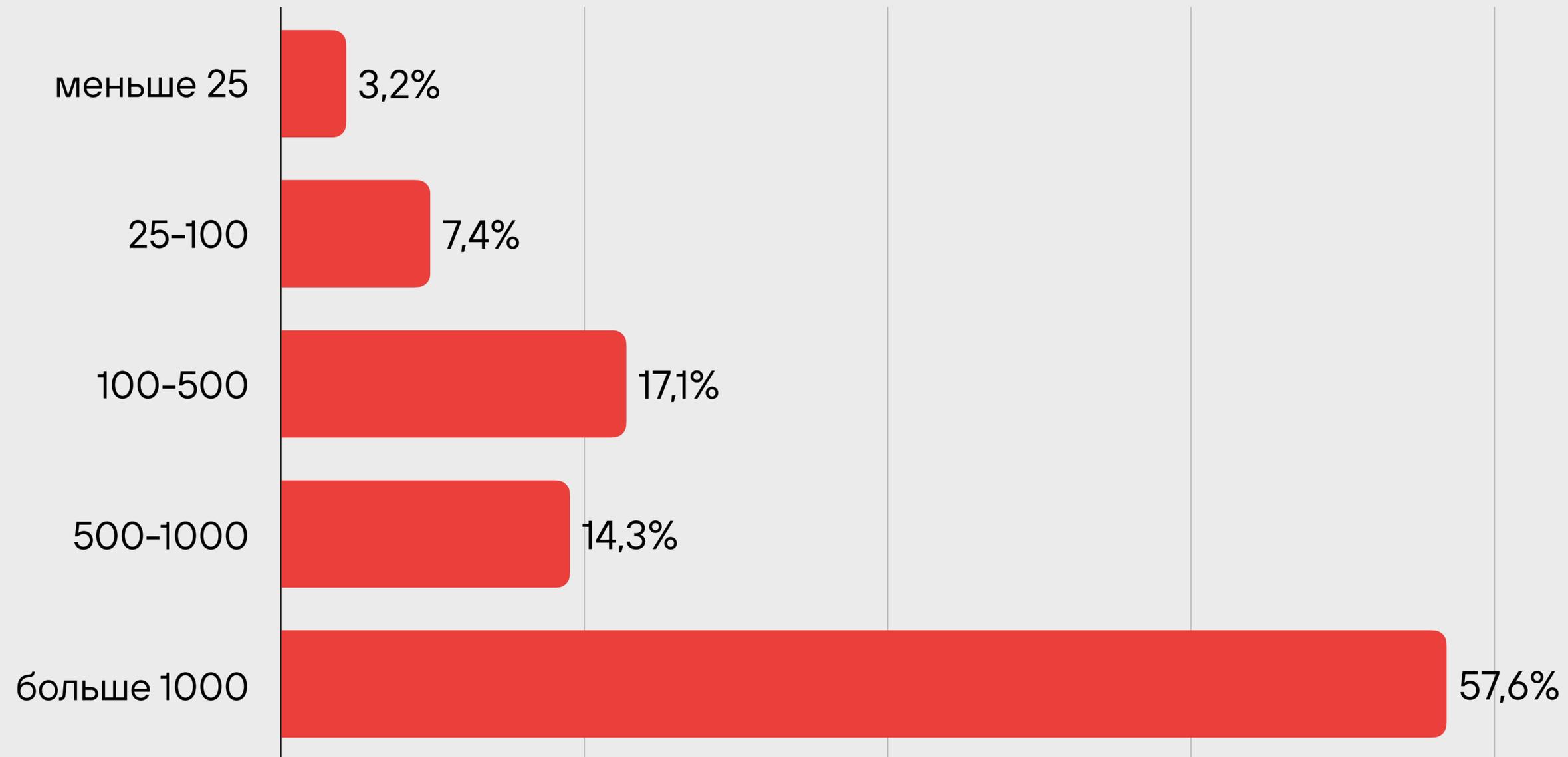
Распределение по опыту в отрасли





Портрет респондента

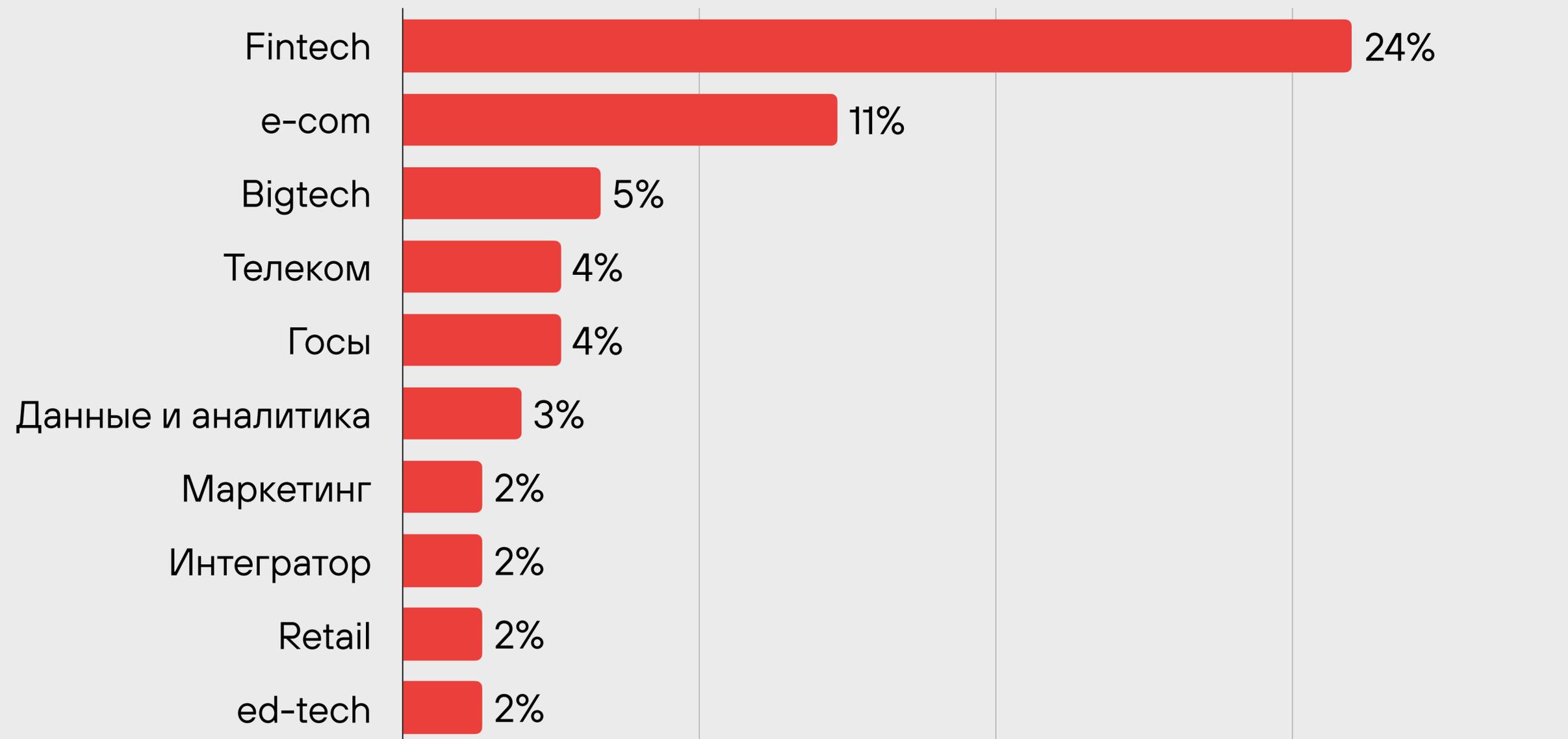
Распределение по размеру компаний



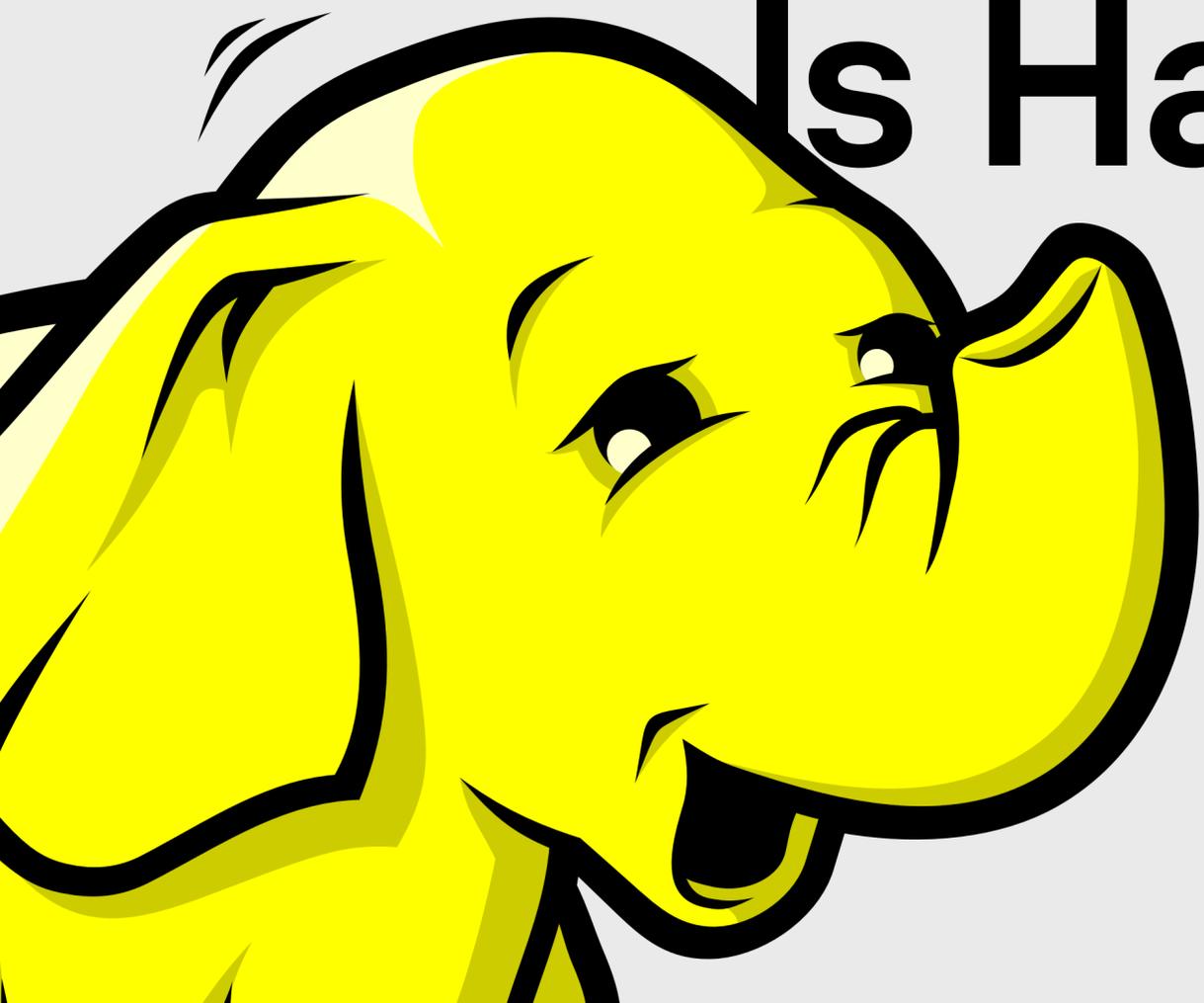


Портрет респондента

Распределение по индустриям компаний (топ10)



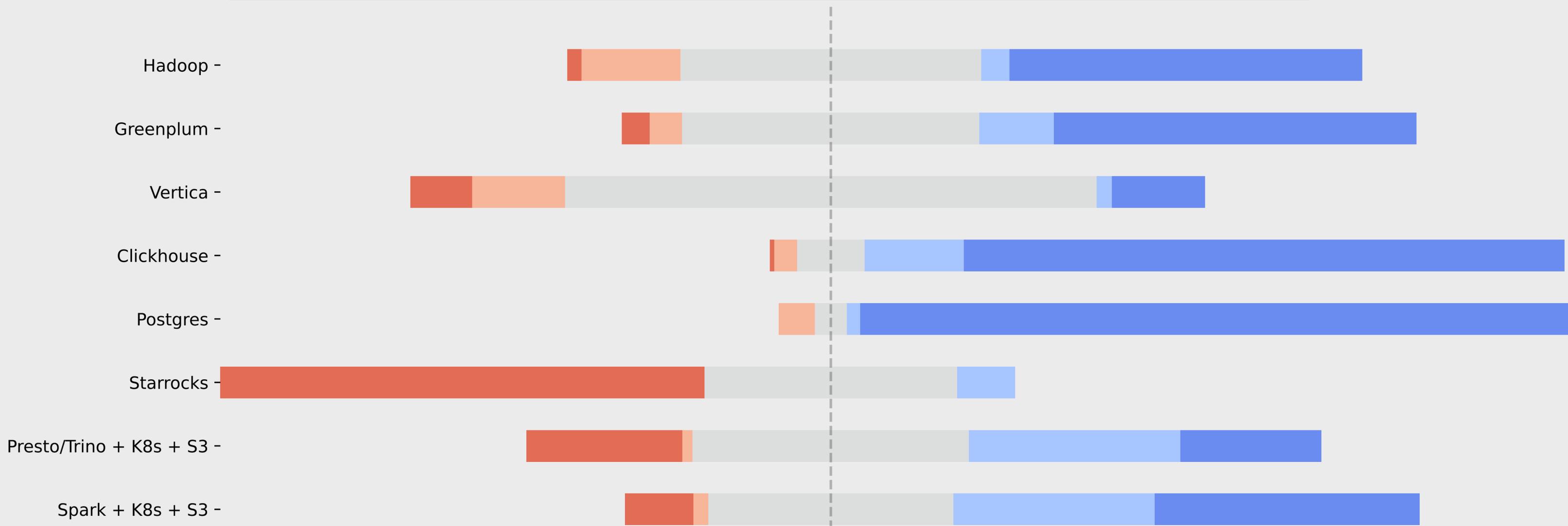
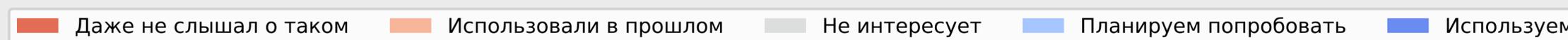
Is Hadoop dead?



Начнем с базы



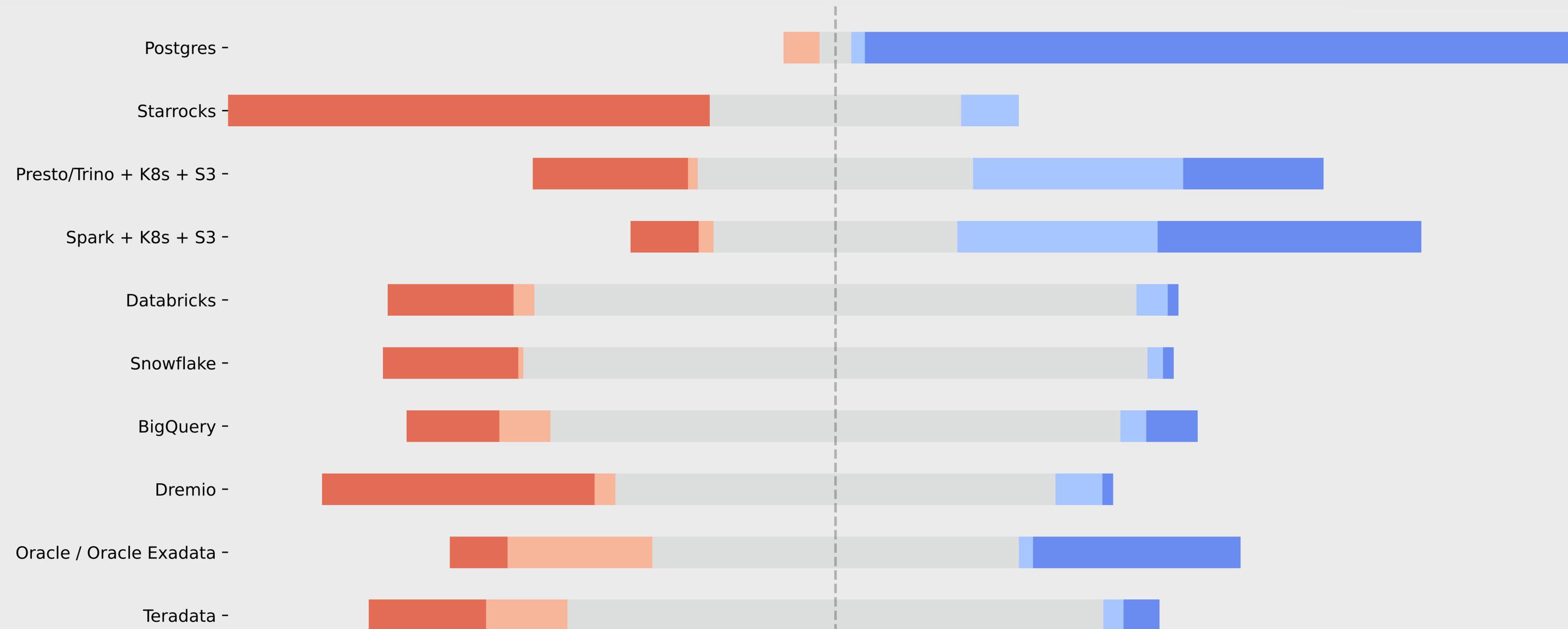
что там с хранением данных



Начнем с базы



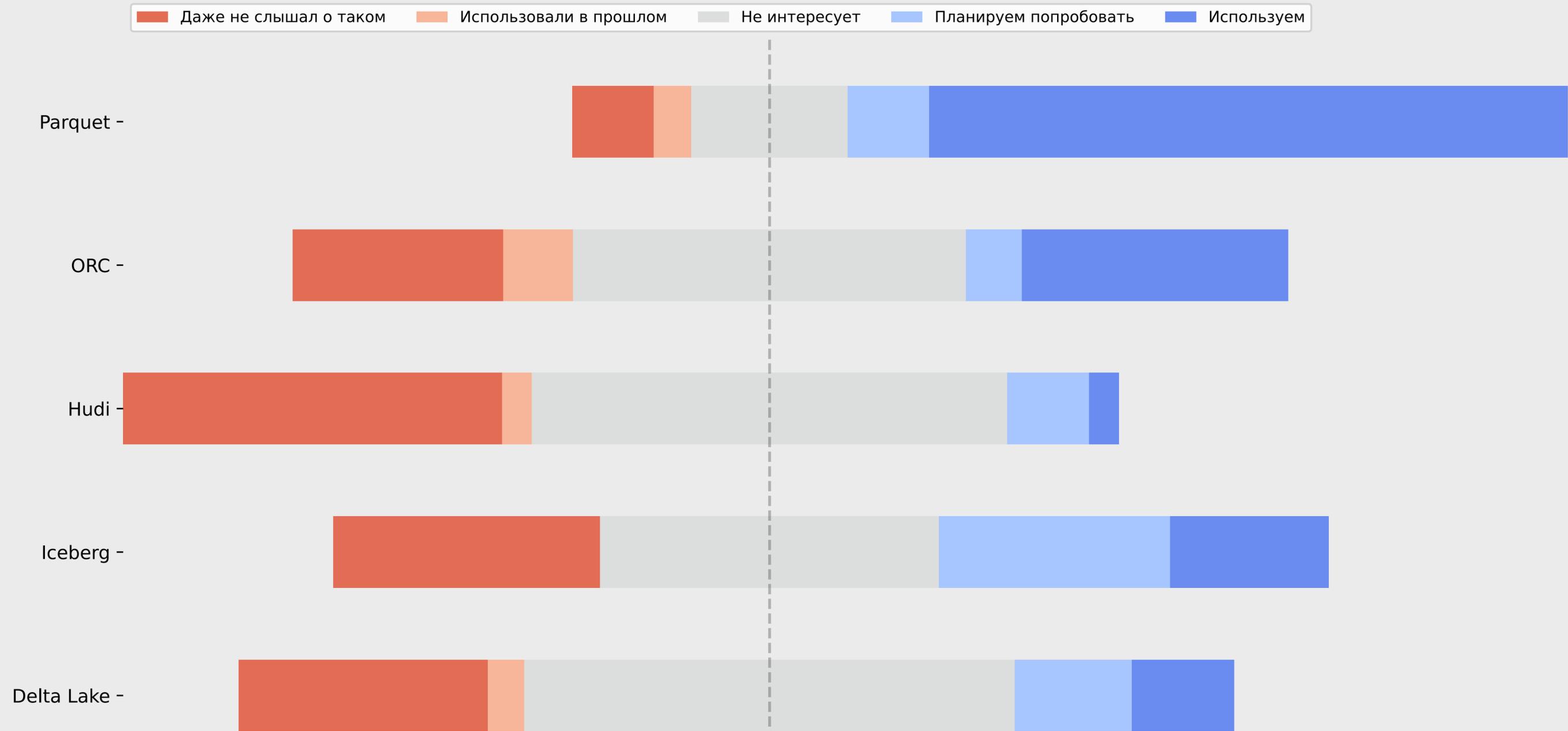
что там с хранением данных





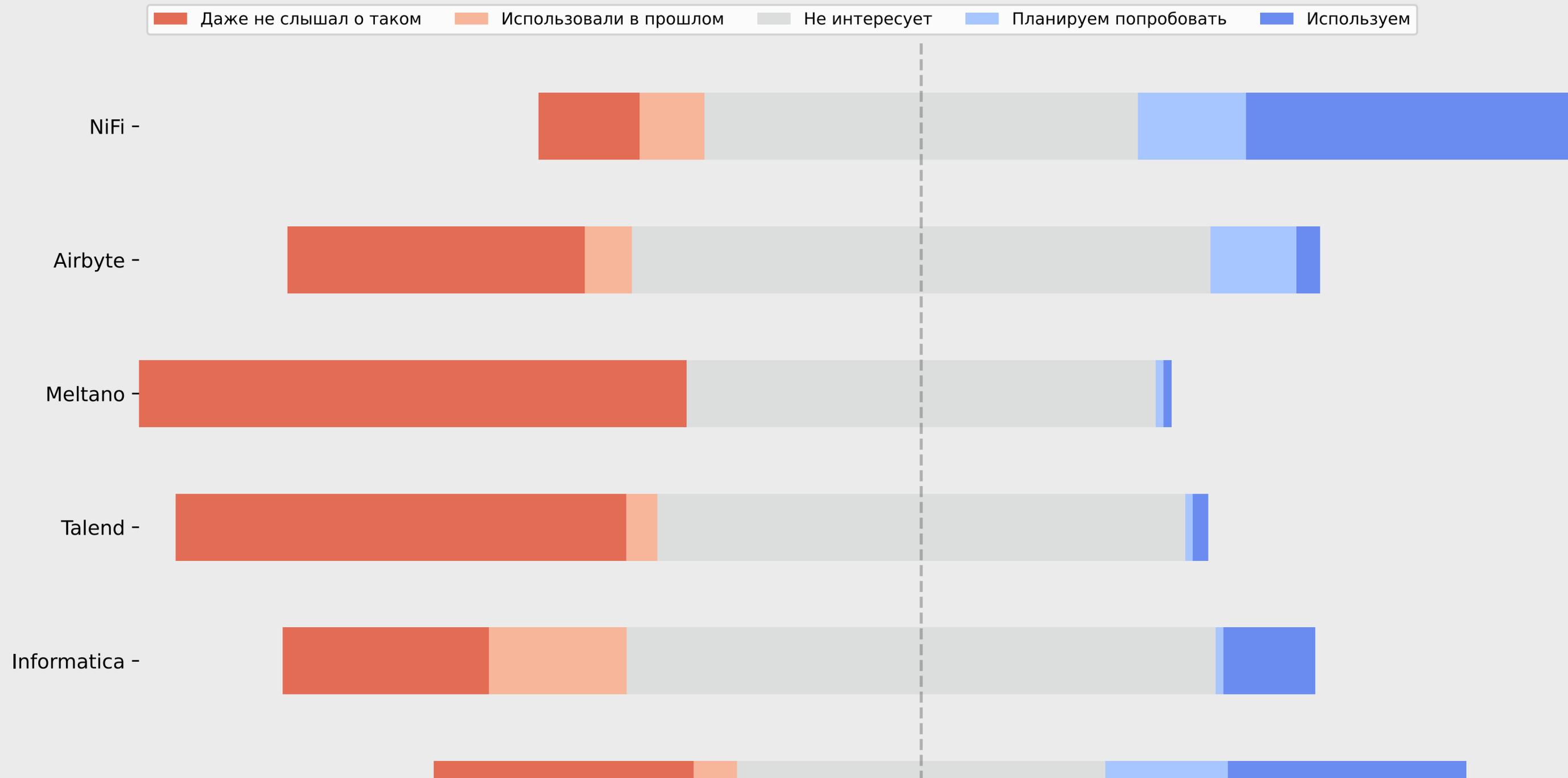
Если это S3/Hadoop

то как данные хранятся?



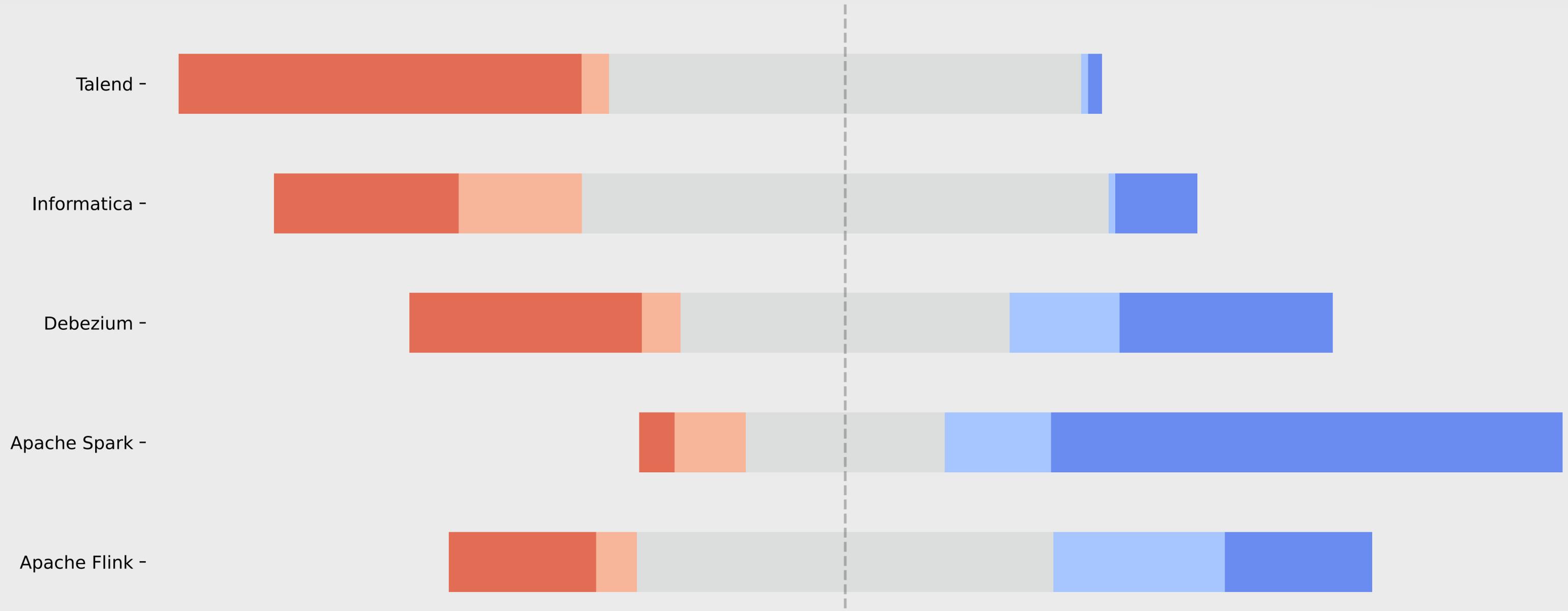
С хранением разобрались

Как грузим?



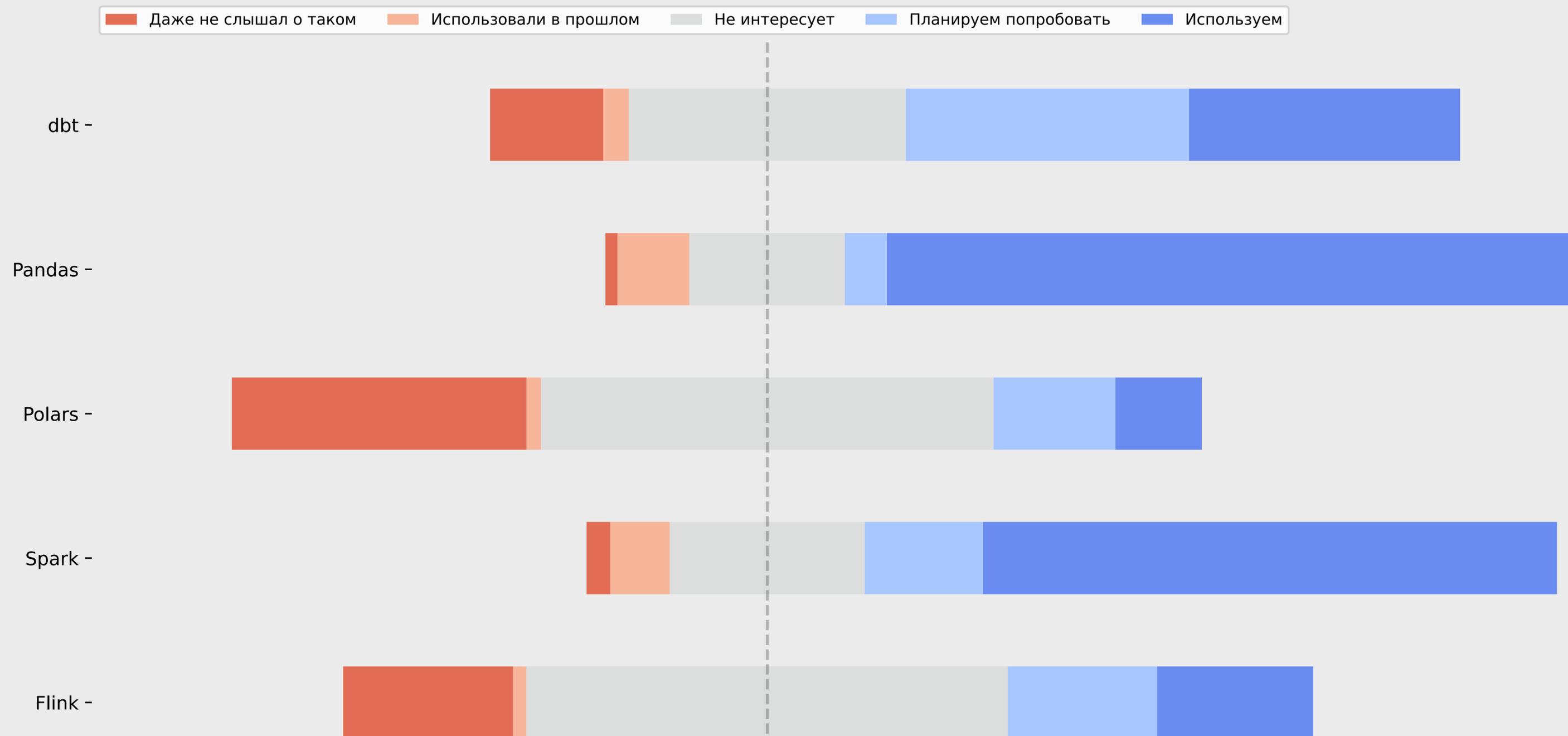
С хранением разобрались

Как грузим?



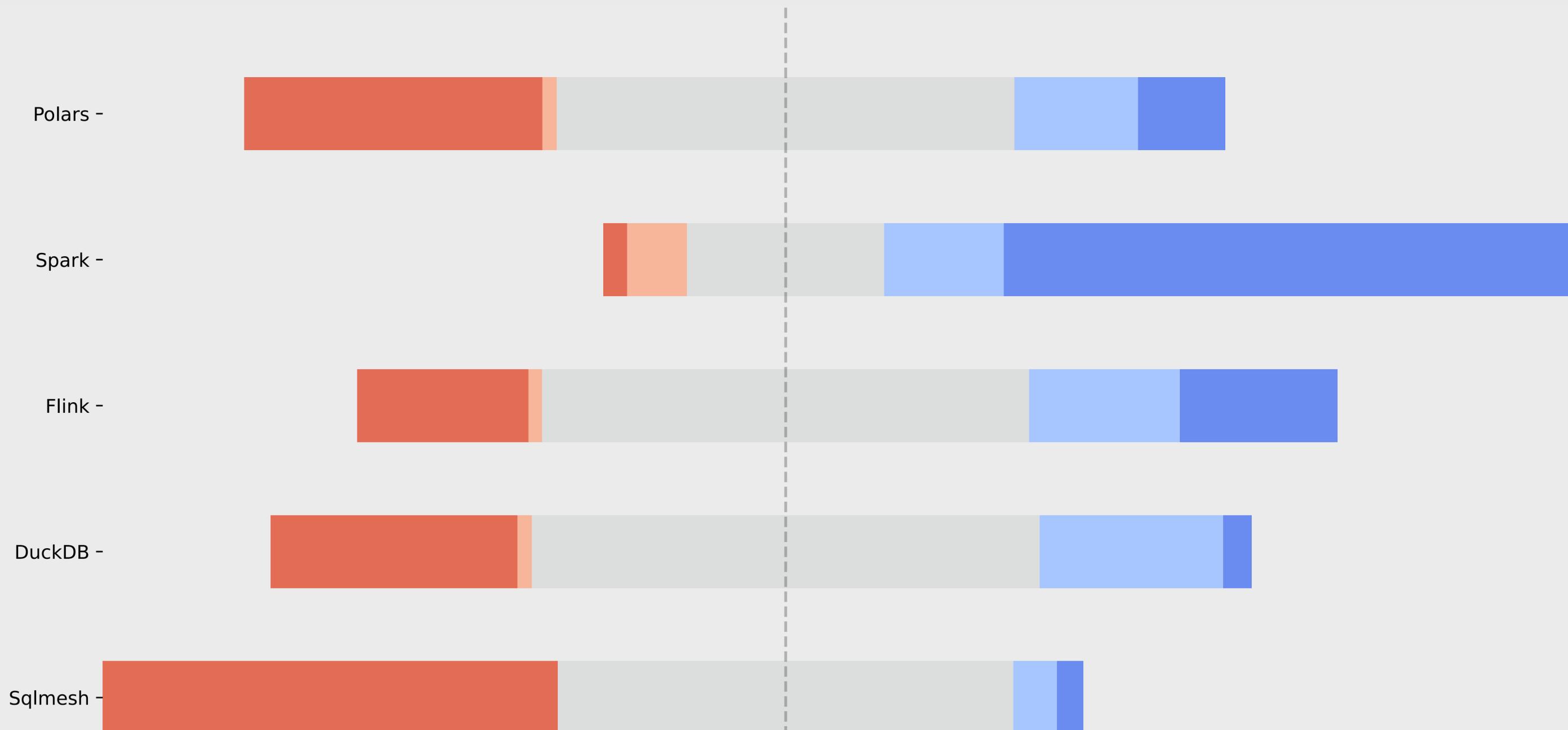
Загрузили!

Пришло время трансформировать

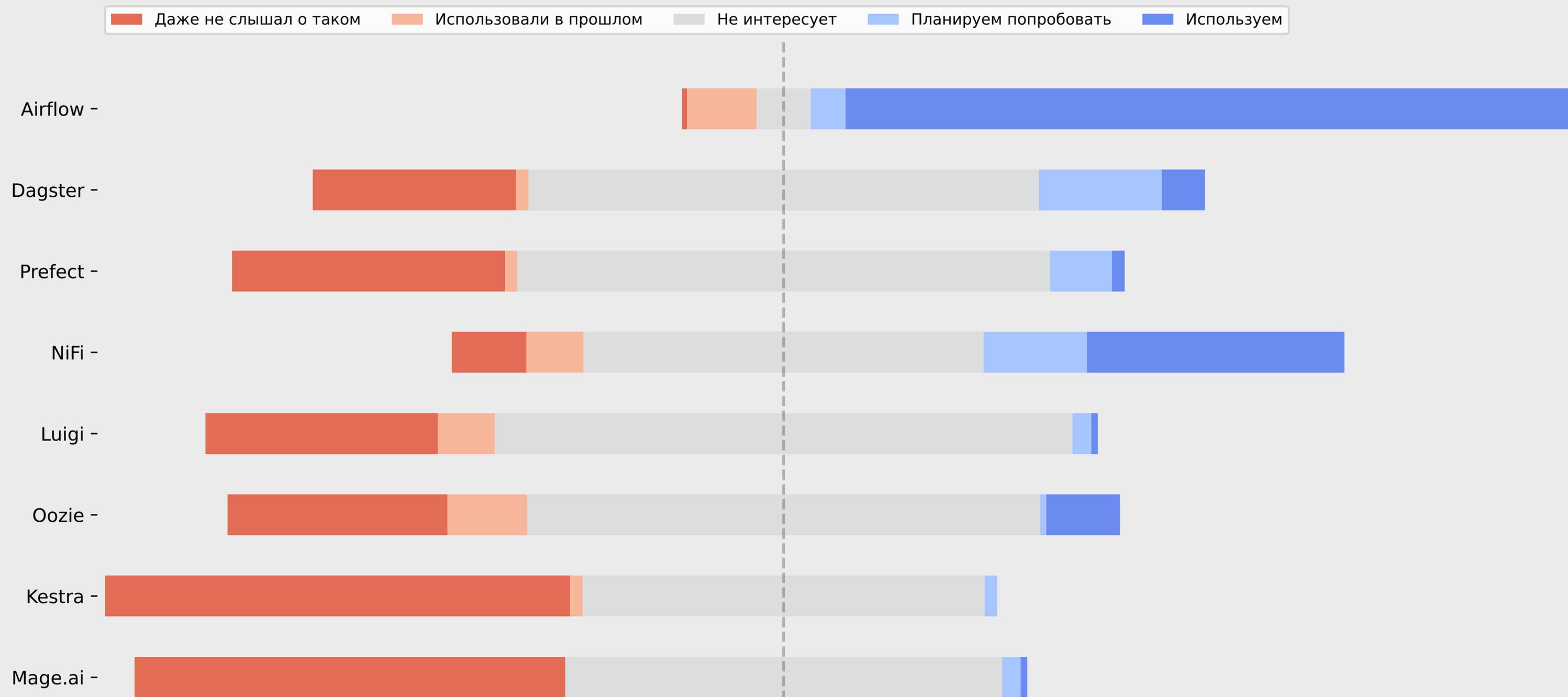


Загрузили!

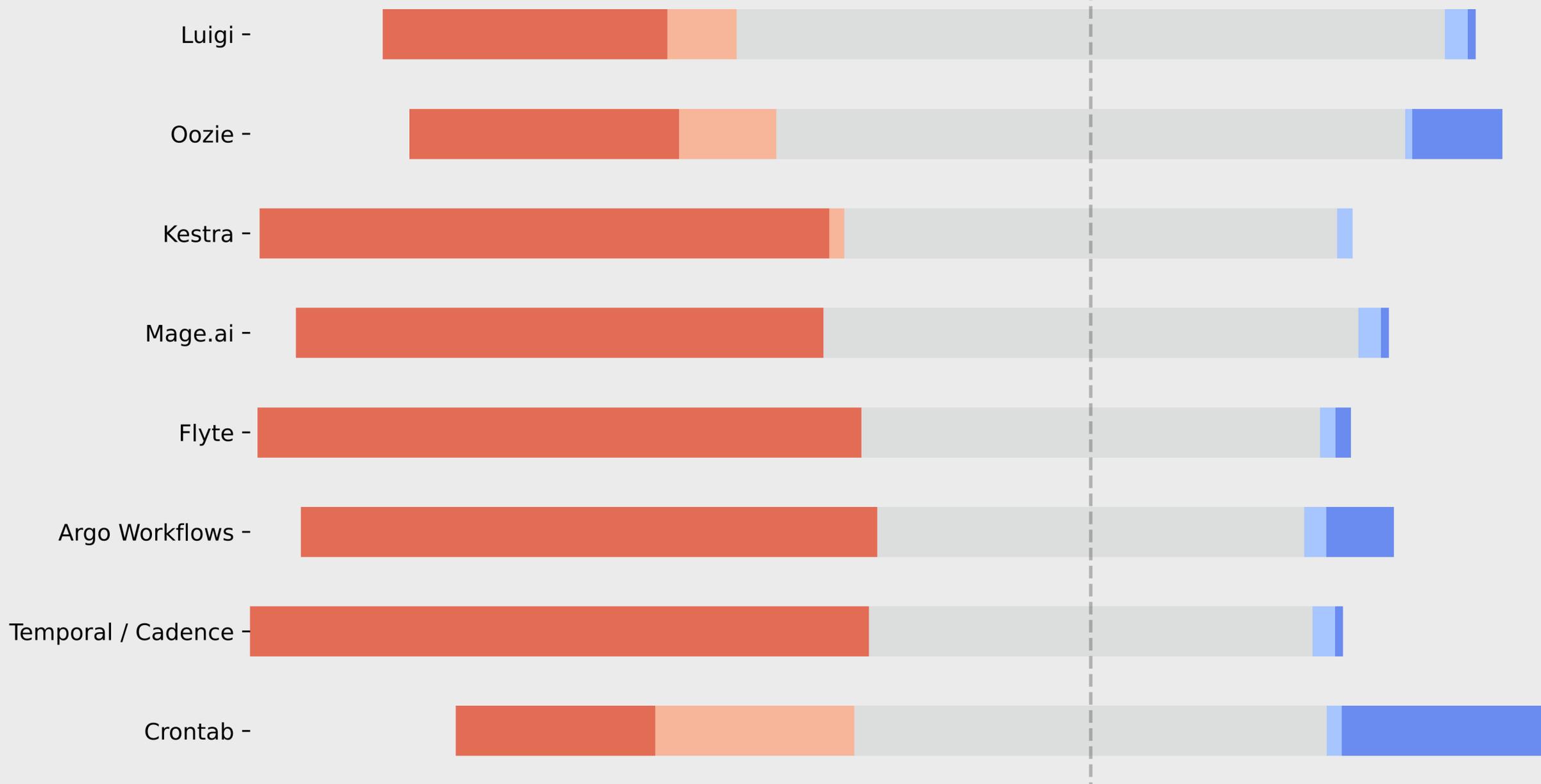
Пришло время трансформировать



Оркестрация

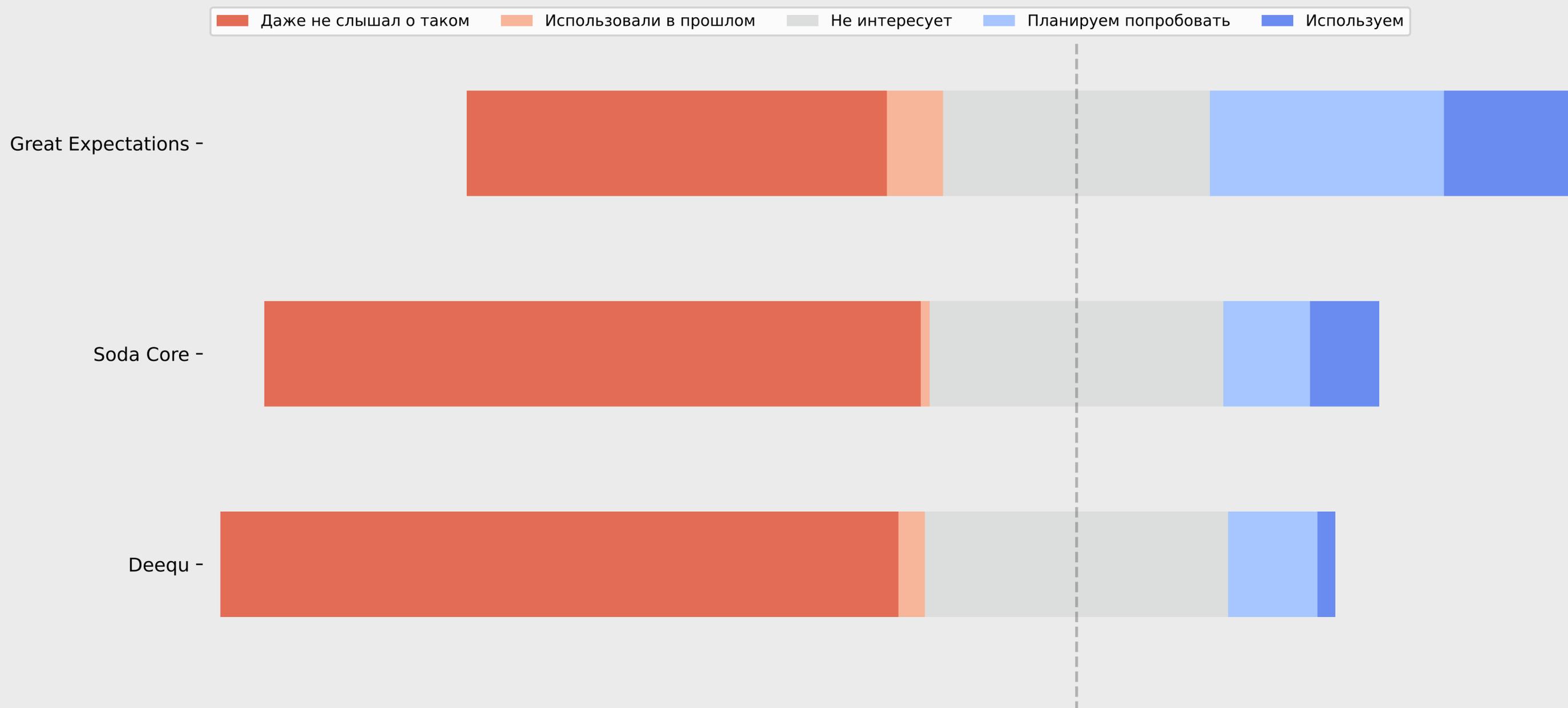


Оркестрация



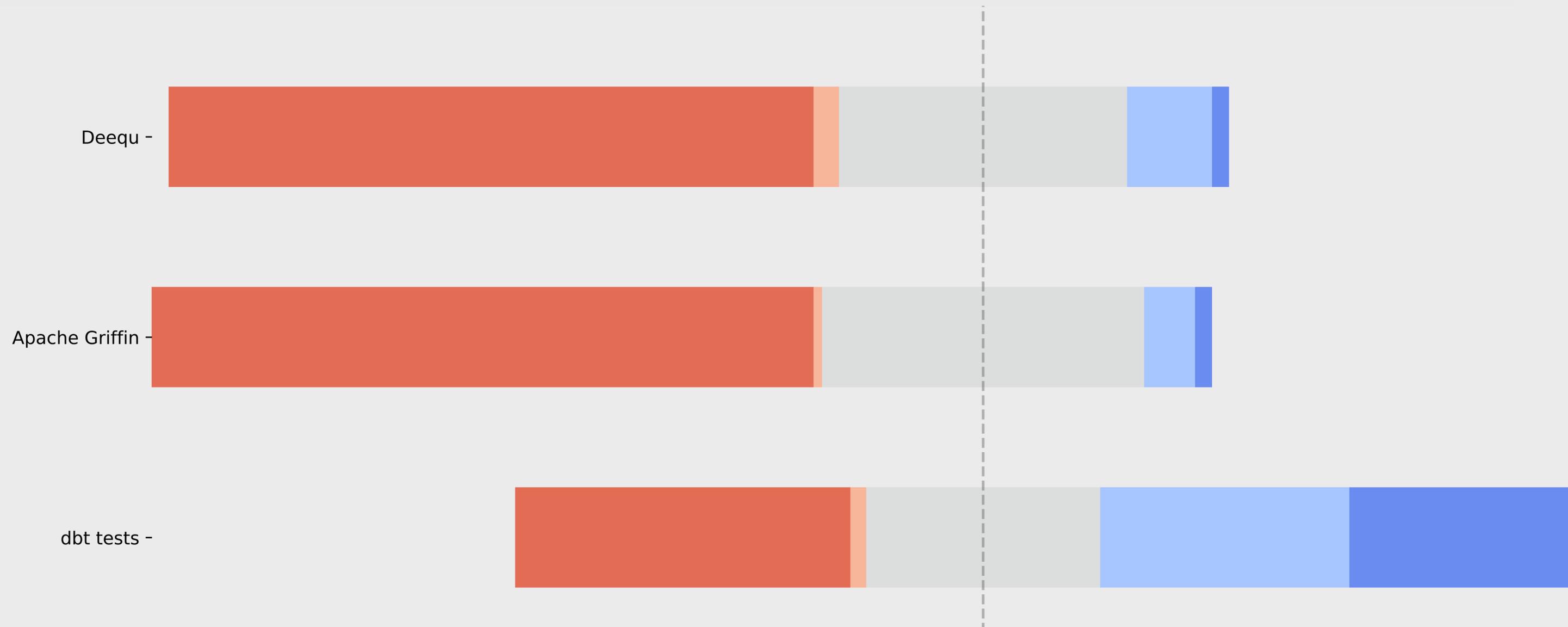
Data governance!

и качество данных



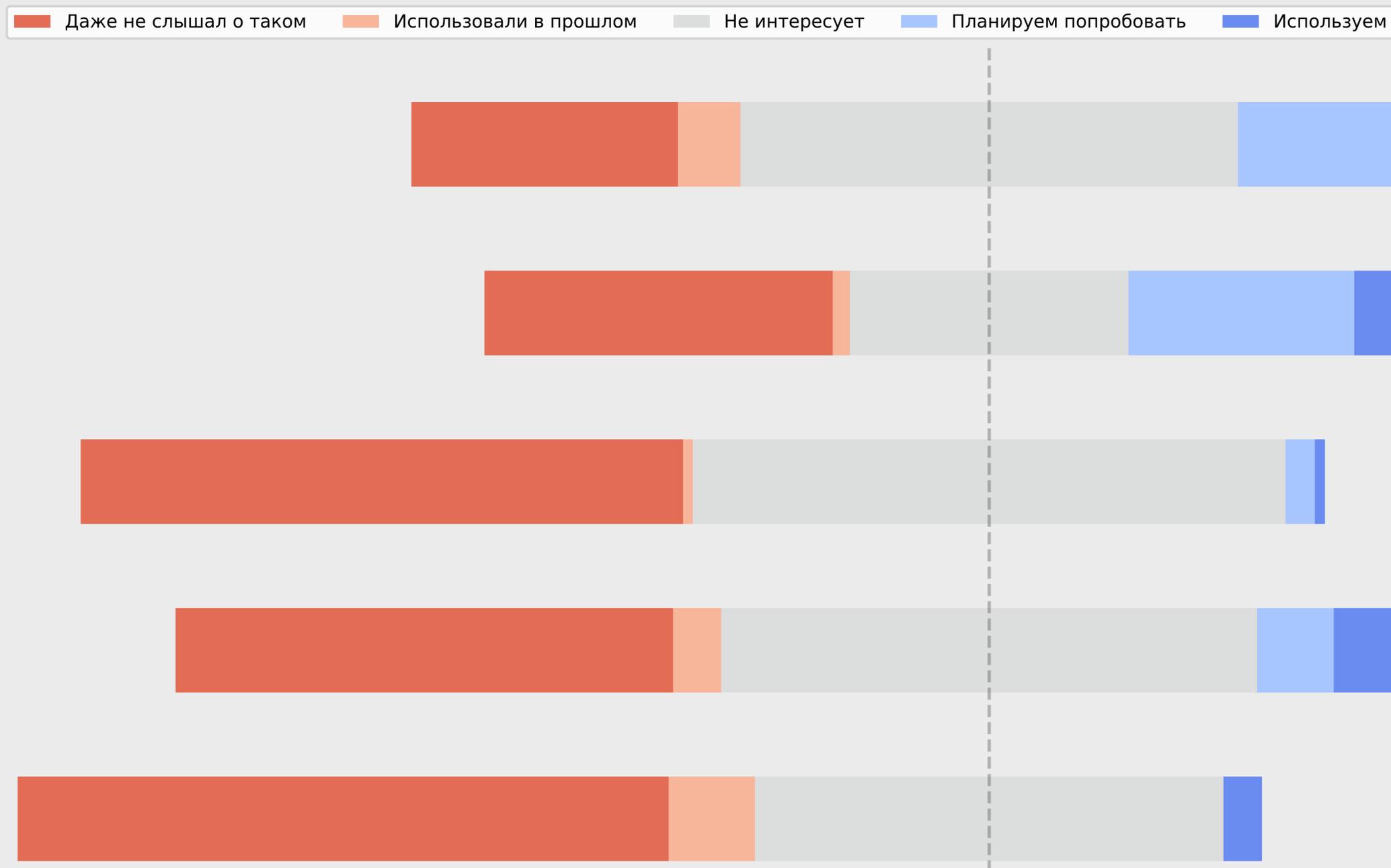
Data governance!

и качество данных



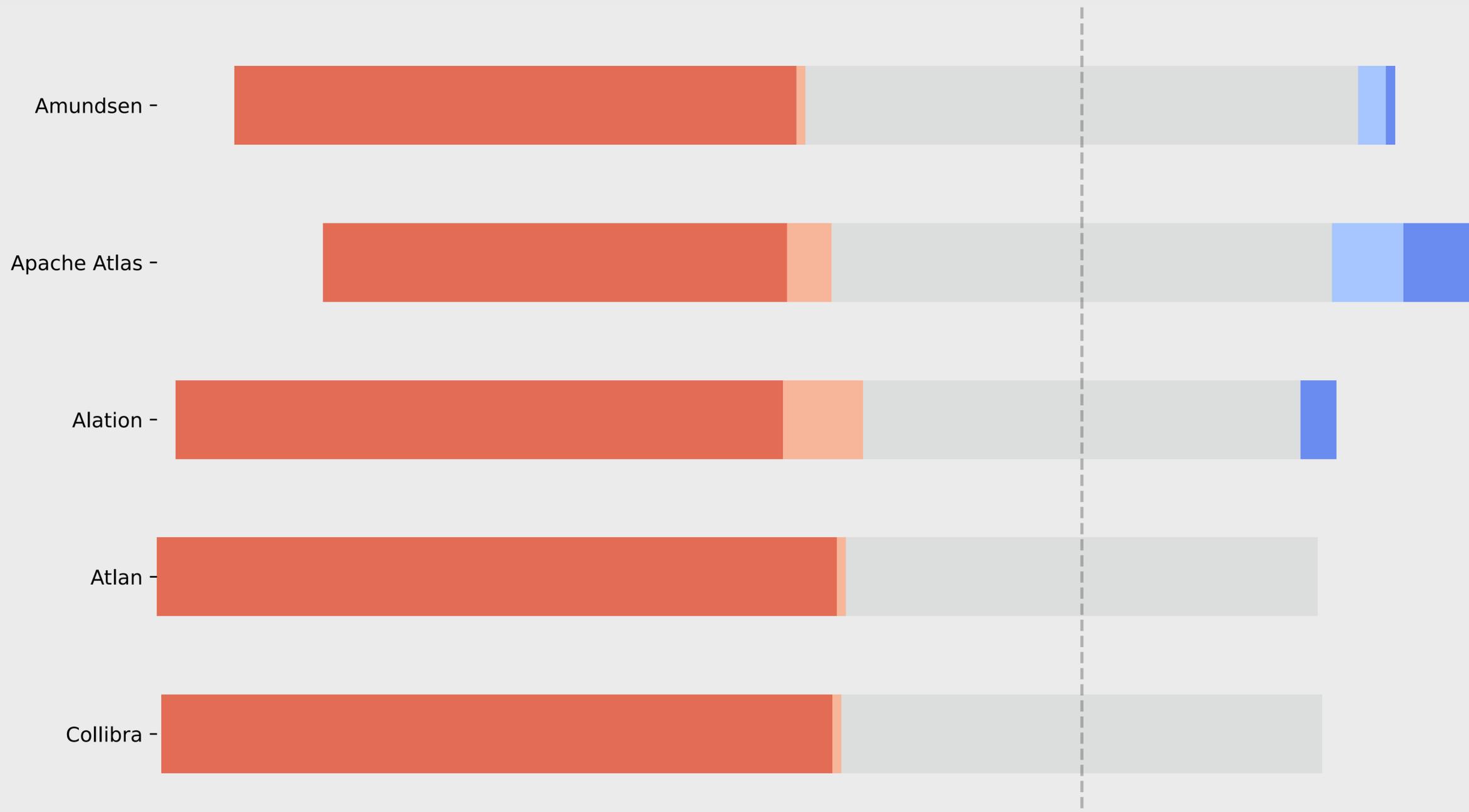
Data governance!

и каталоги данных



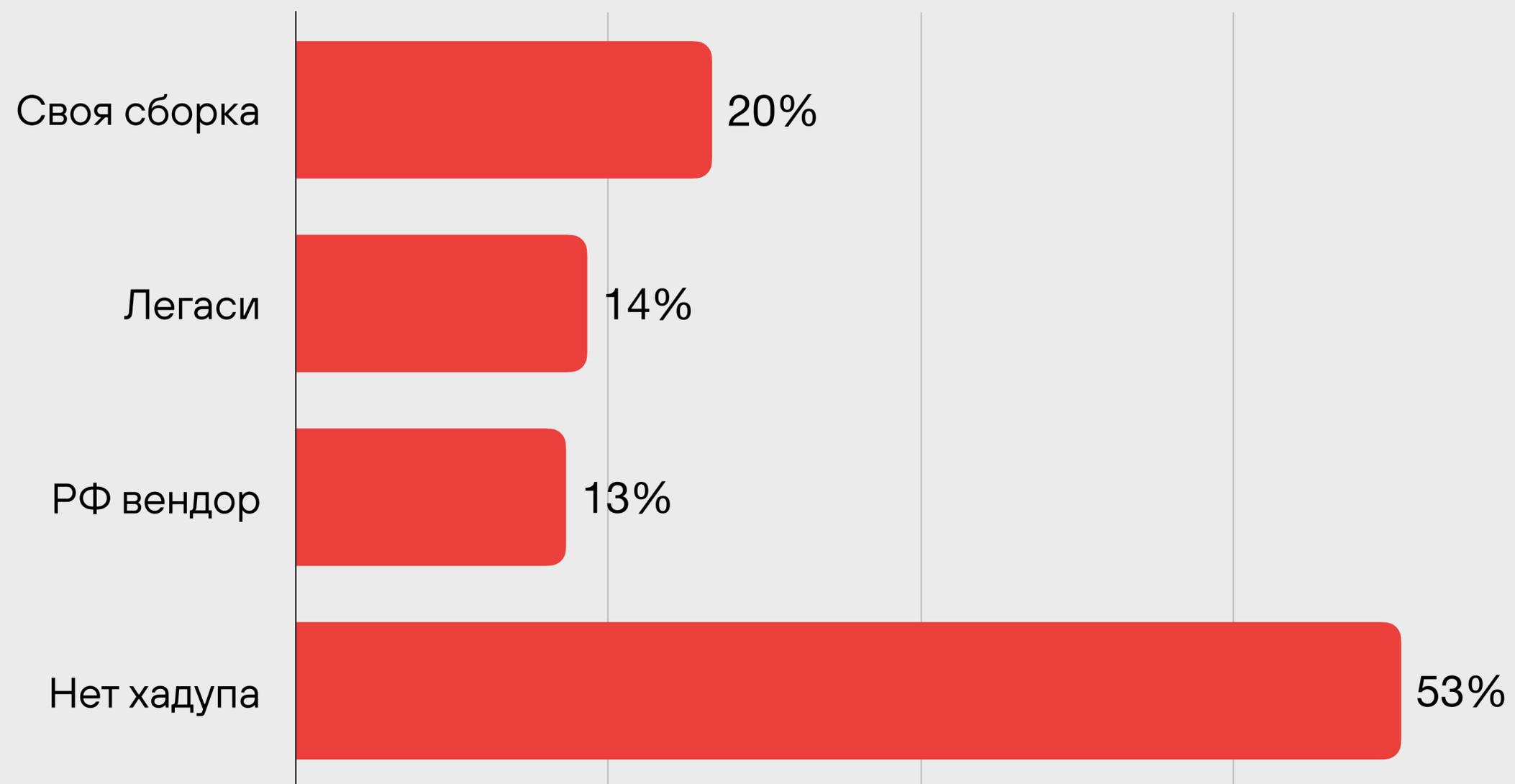
Data governance!

и каталоги данных



Возвращаясь к Хадупу

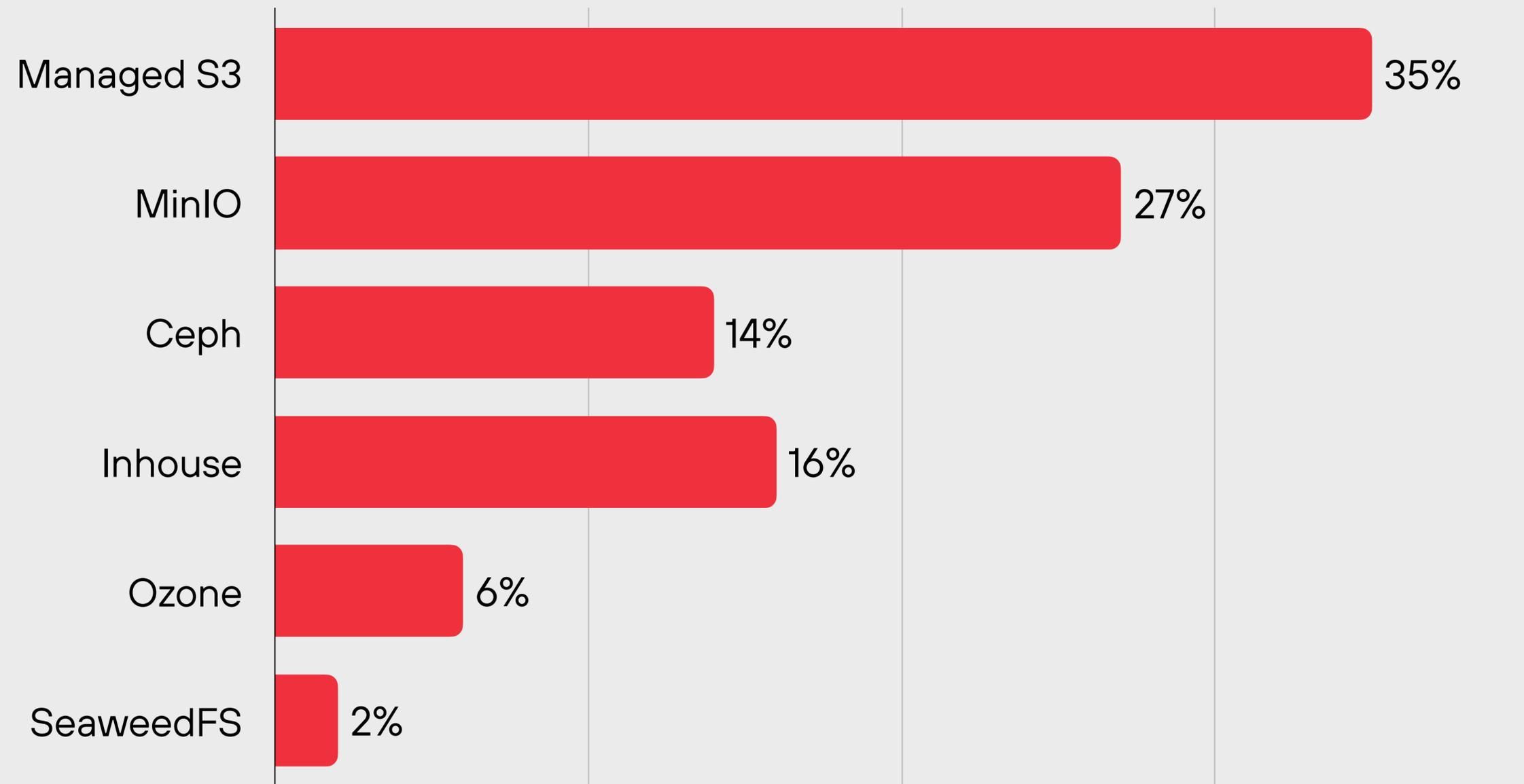
Чей хадуп?



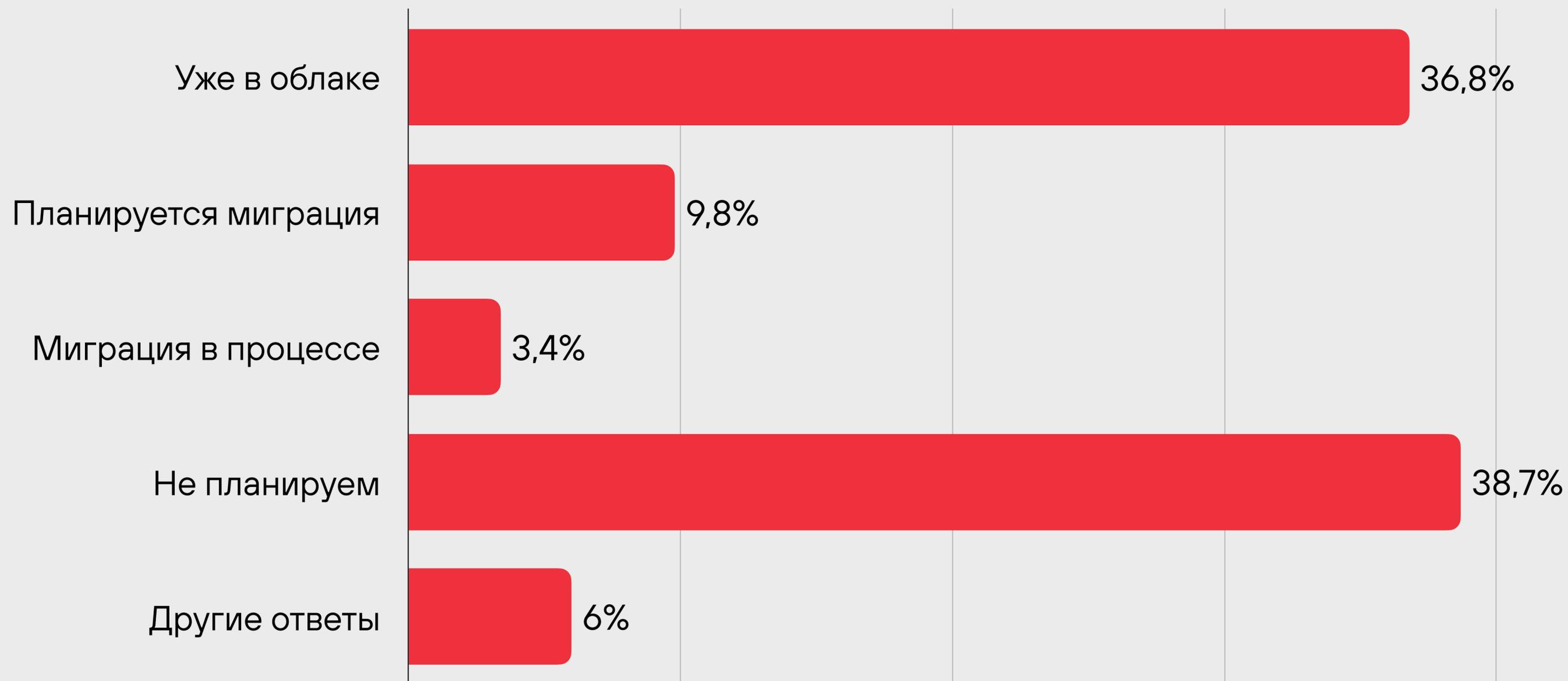
Использование S3



S3 ТЕХНОЛОГИИ



Облака

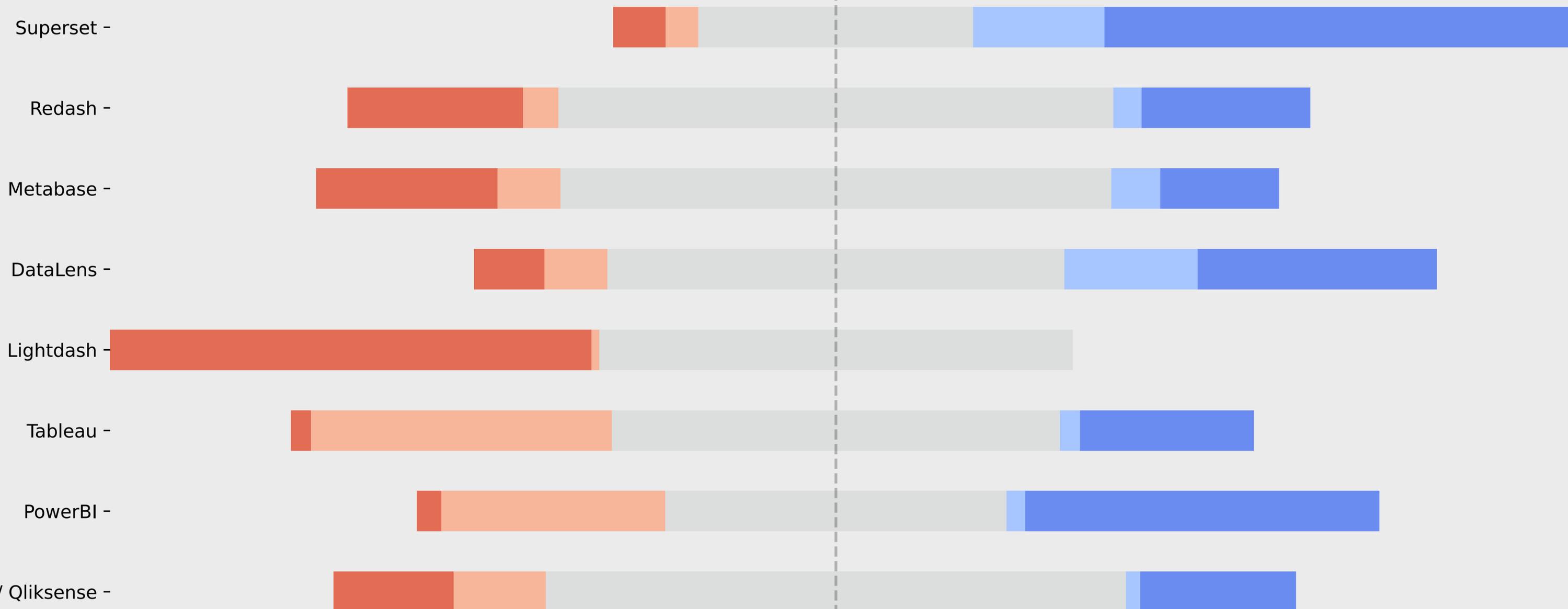


Совсем чуть-чуть про VI и ML

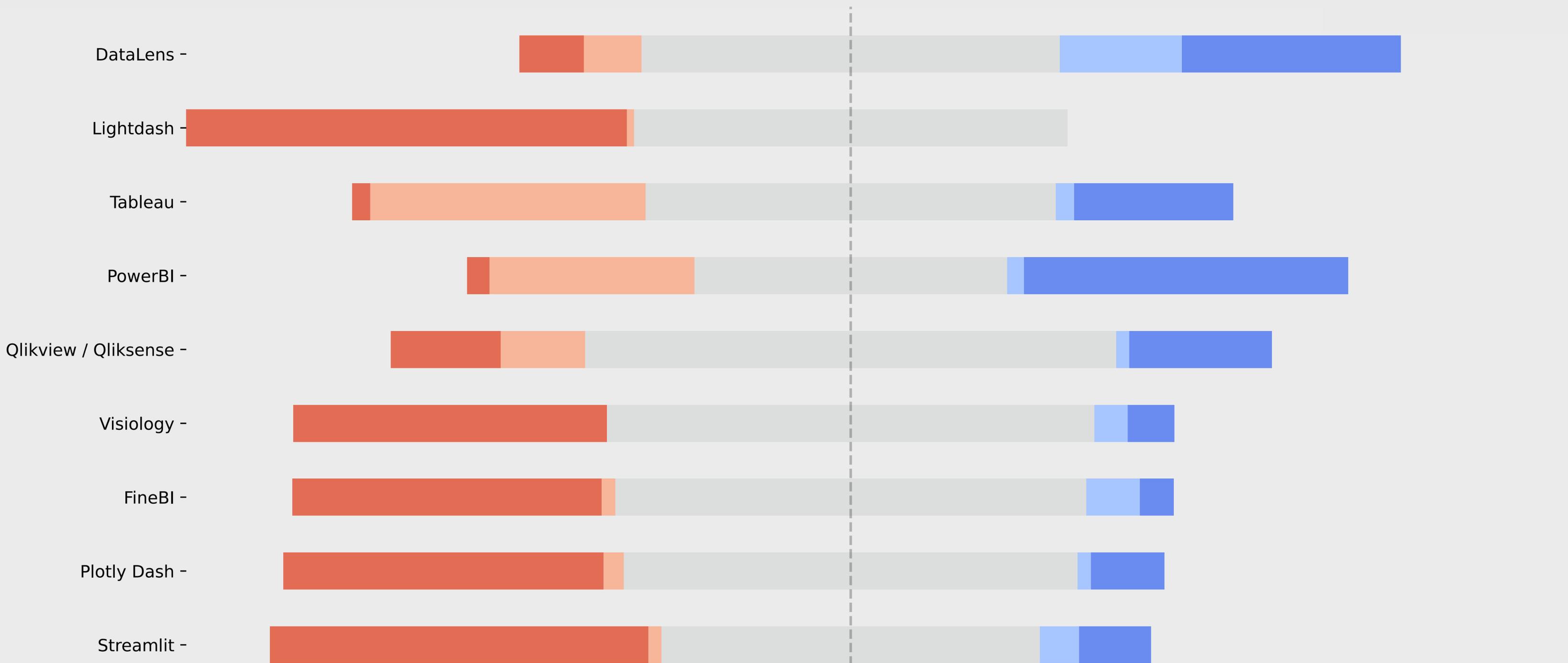
BI инструменты



Даже не слышал о таком Использовали в прошлом Не интересуется Планируем попробовать Используем



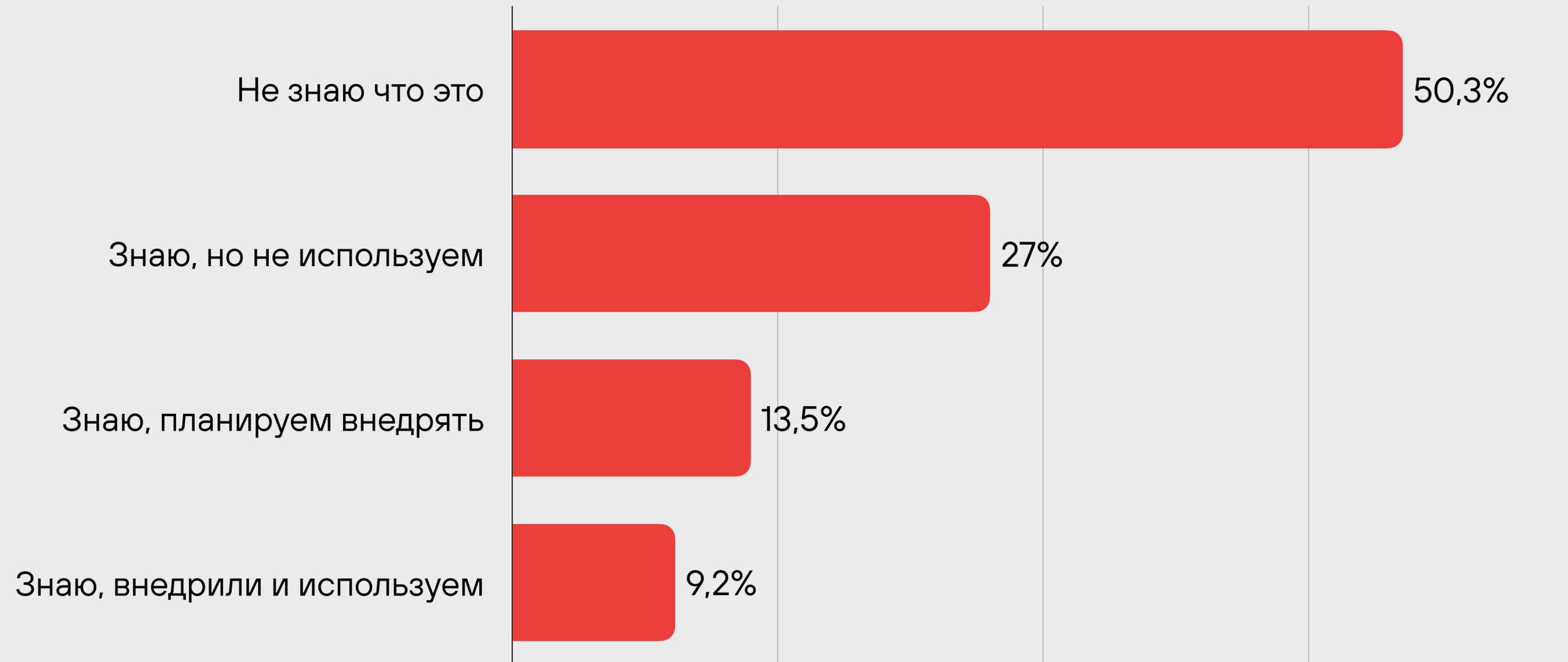
BI инструменты



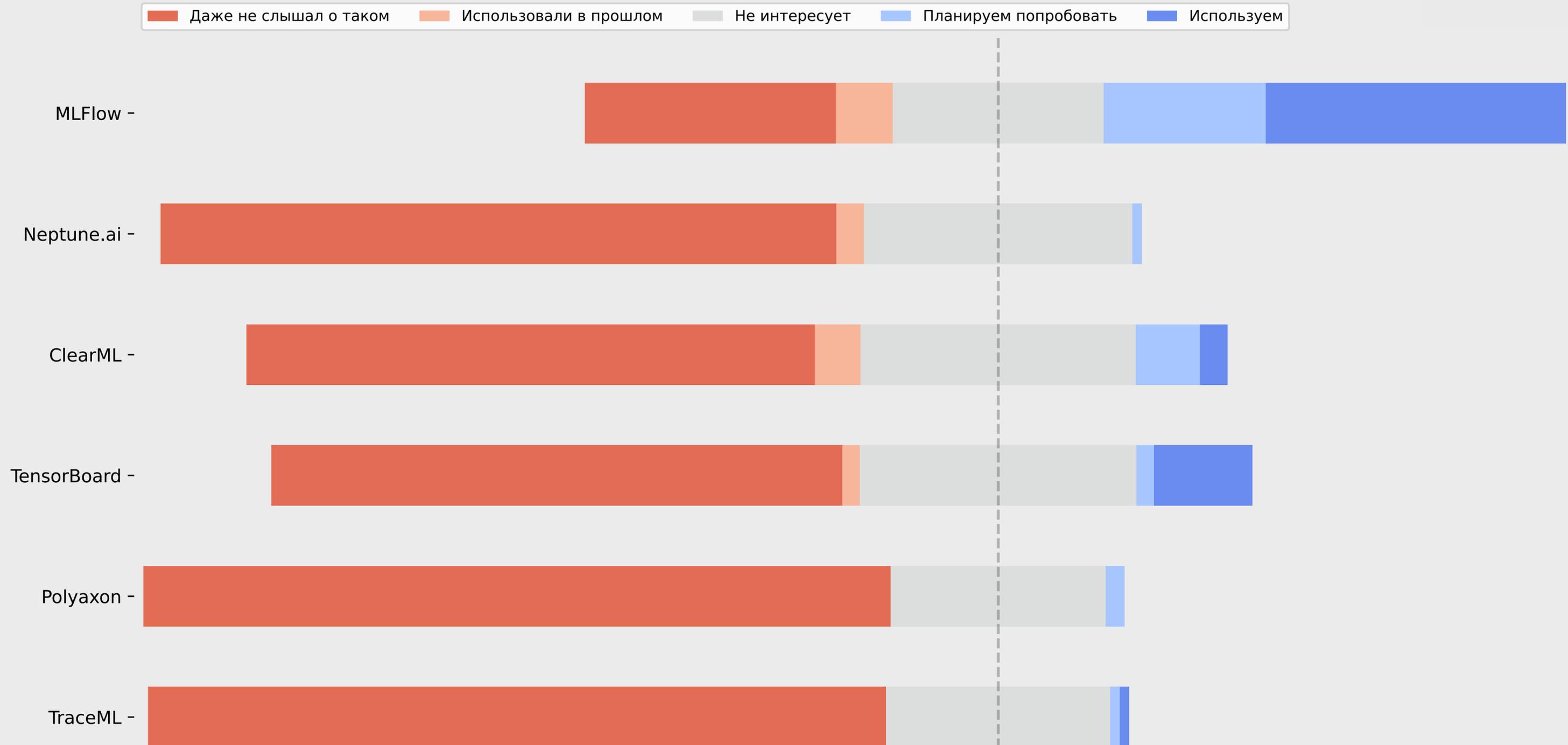
Semantic layer



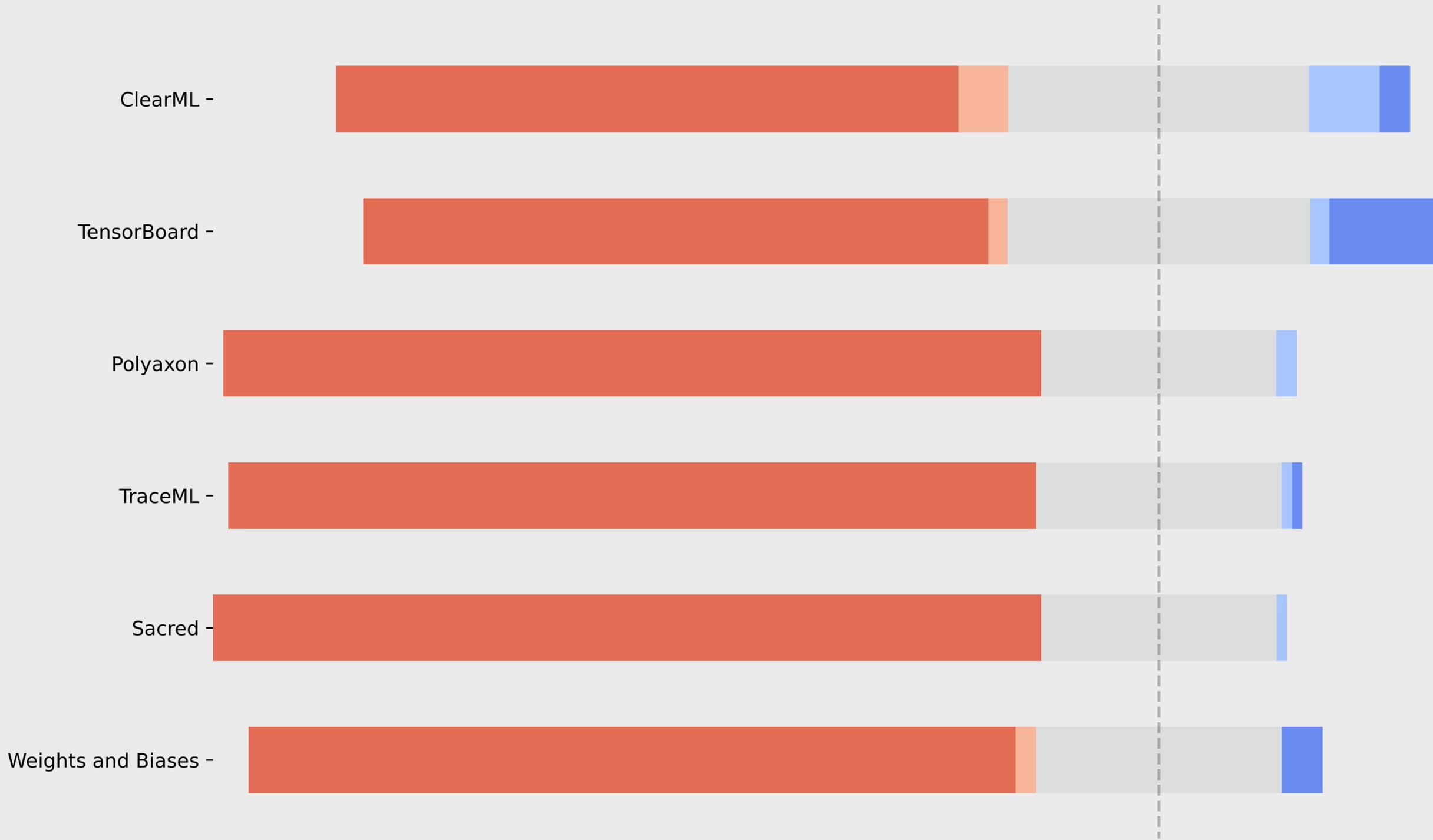
Используете ли вы подход/инструменты Semantic Layer?



ML Model Management



ML Model Management



**А что там на
западе?**

В качестве заключения



- Embrace your natural curiosity

В качестве заключения



- Embrace your natural curiosity
- У вас есть ориентир по технологиям, которые, потенциально, интересны – мы ждем вас с докладами в следующем сезоне

В качестве заключения

- Embrace your natural curiosity
- У вас есть ориентир по технологиям, которые, потенциально, интересны – мы ждем вас с докладами в следующем сезоне
- Полностью исследование будет опубликовано в декабре '24

В качестве заключения

- Embrace your natural curiosity
- У вас есть ориентир по технологиям, которые, потенциально, интересны – мы ждем вас с докладами в следующем сезоне
- Полностью исследование будет опубликовано в декабре '24
- Призывайте знакомых и коллег, особенно из сфер аналитики / bi / ml



**Вопросы?
Набросы?!**

