

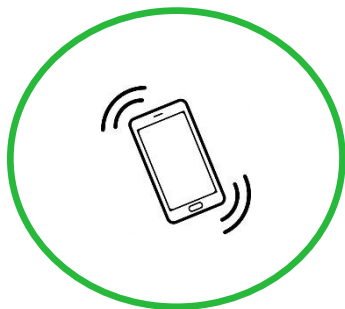
Что-то ещё скрывается в рекомендациях



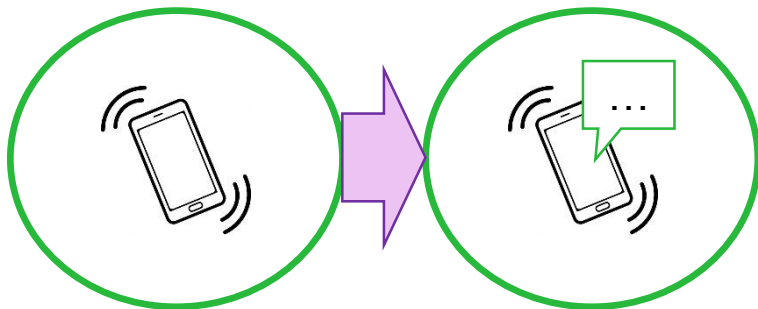
Селезнев Артем, Data Scientist



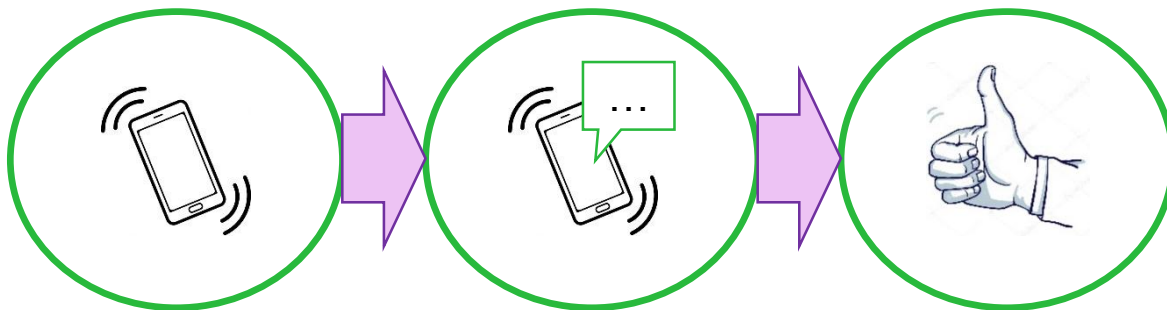
«Давайте будем трясти смартфон»



«Давайте будем трясти смартфон»



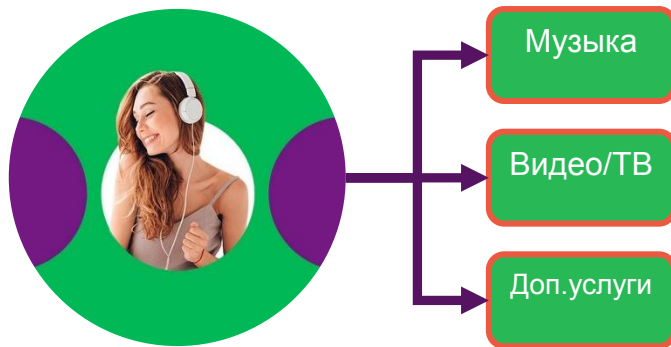
«Давайте будем трясти смартфон»



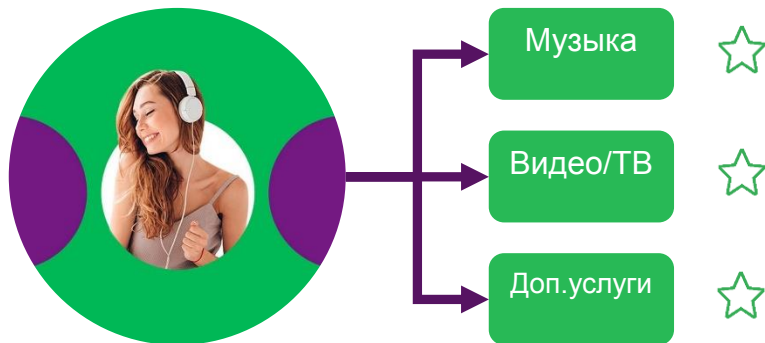
«Давайте будем трясти смартфон»



И другие продукты



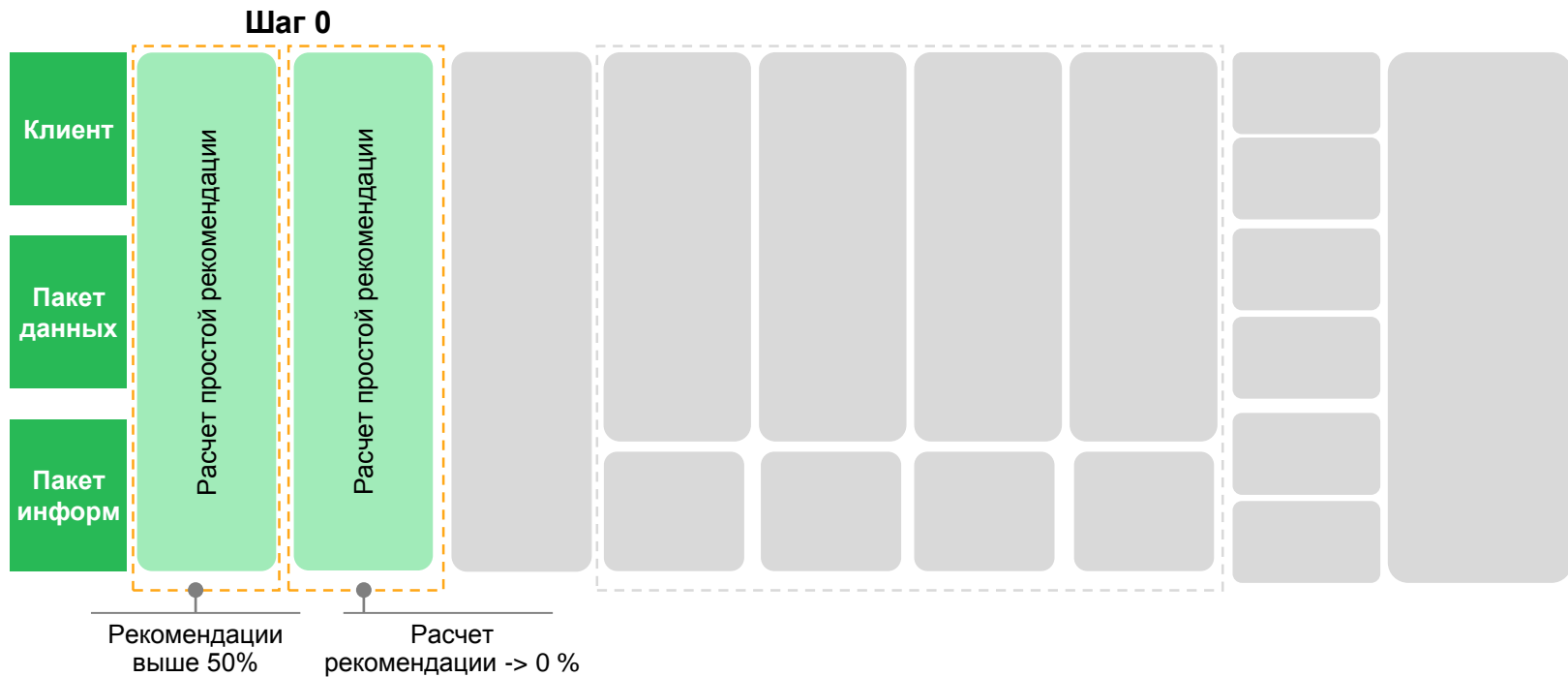
И другие продукты должны быть подходящие



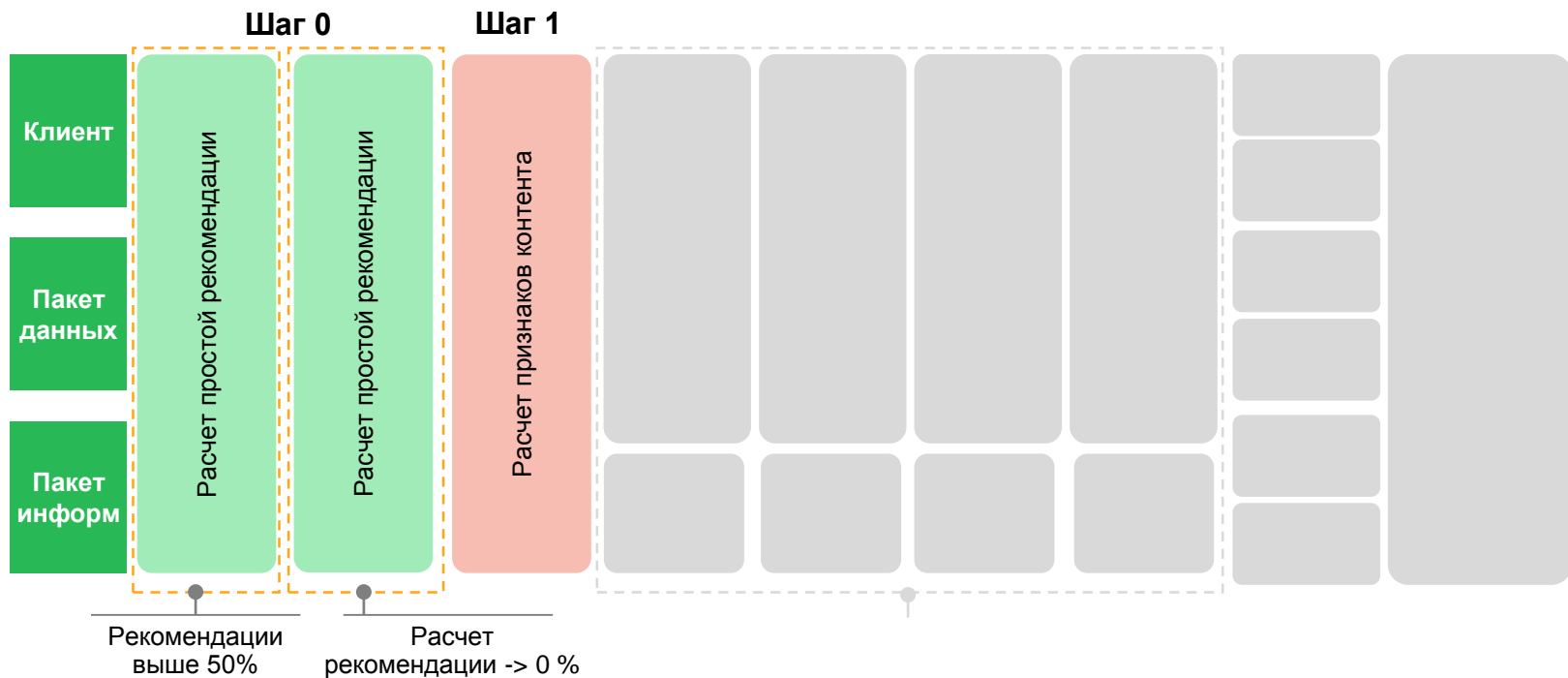
1

Что там «под капотом»?

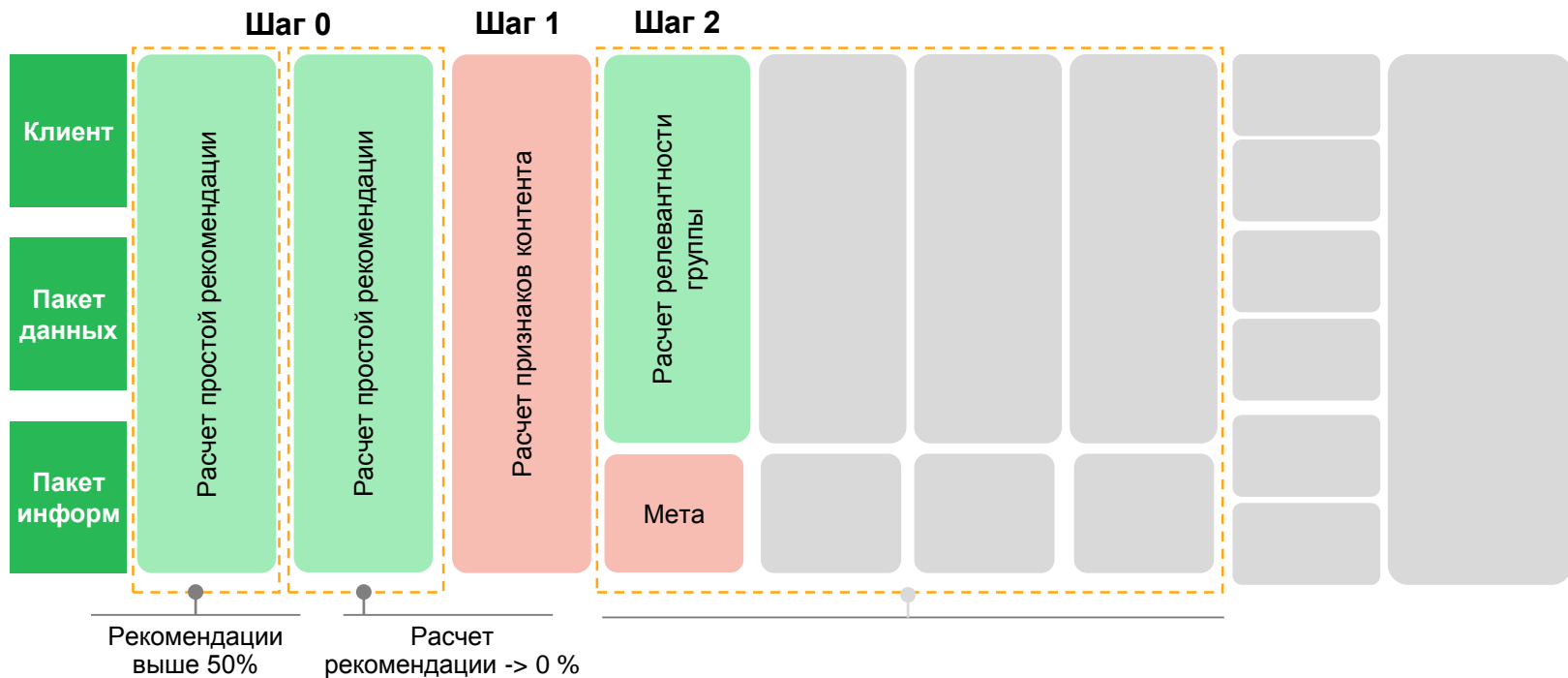
Данных действительно много



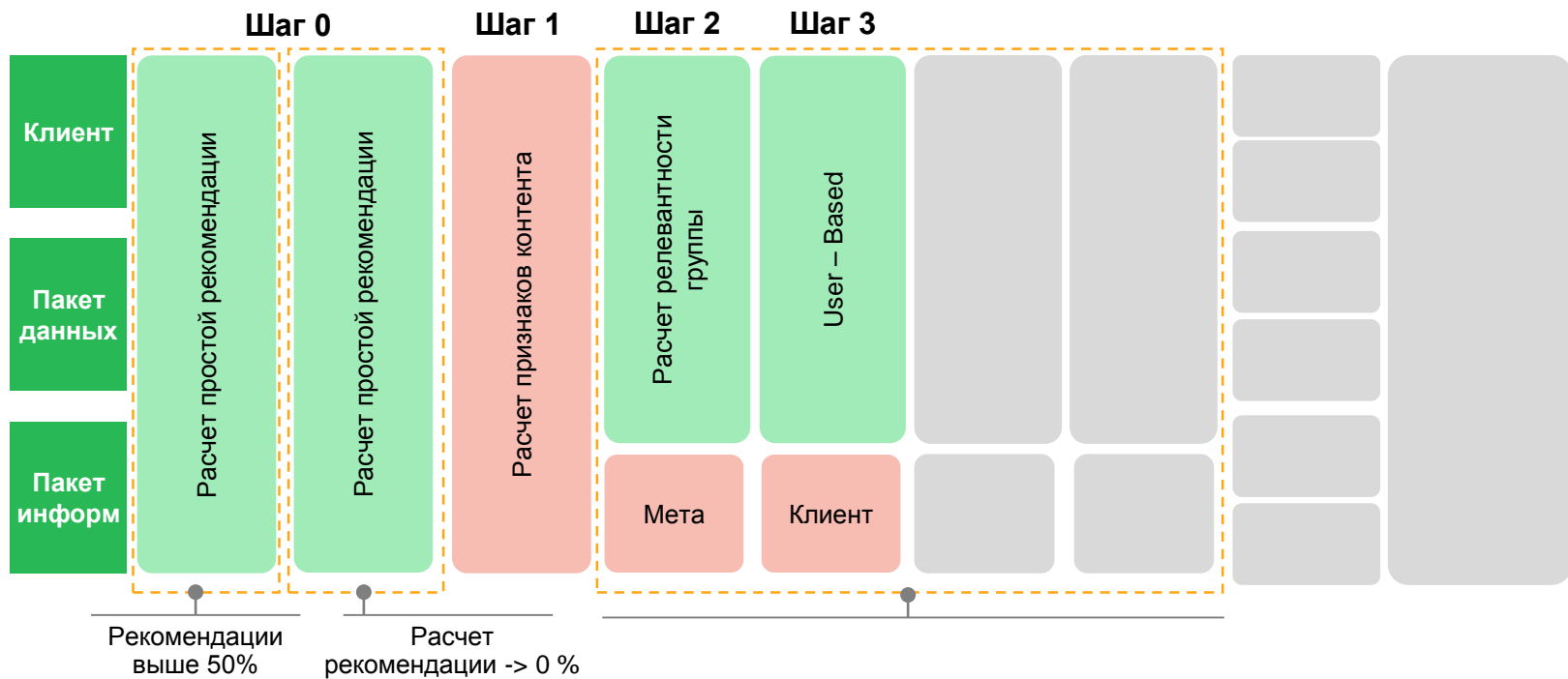
Данных действительно много



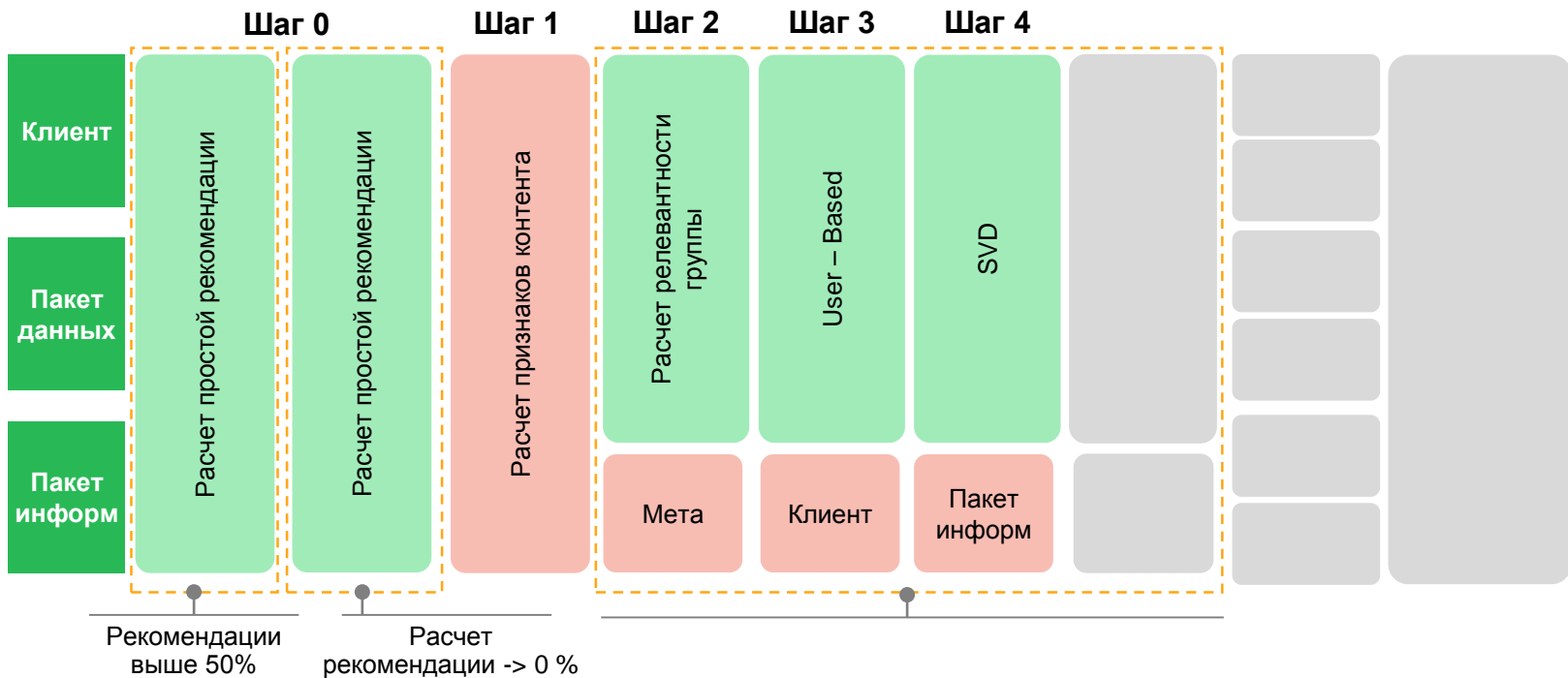
Данных действительно много



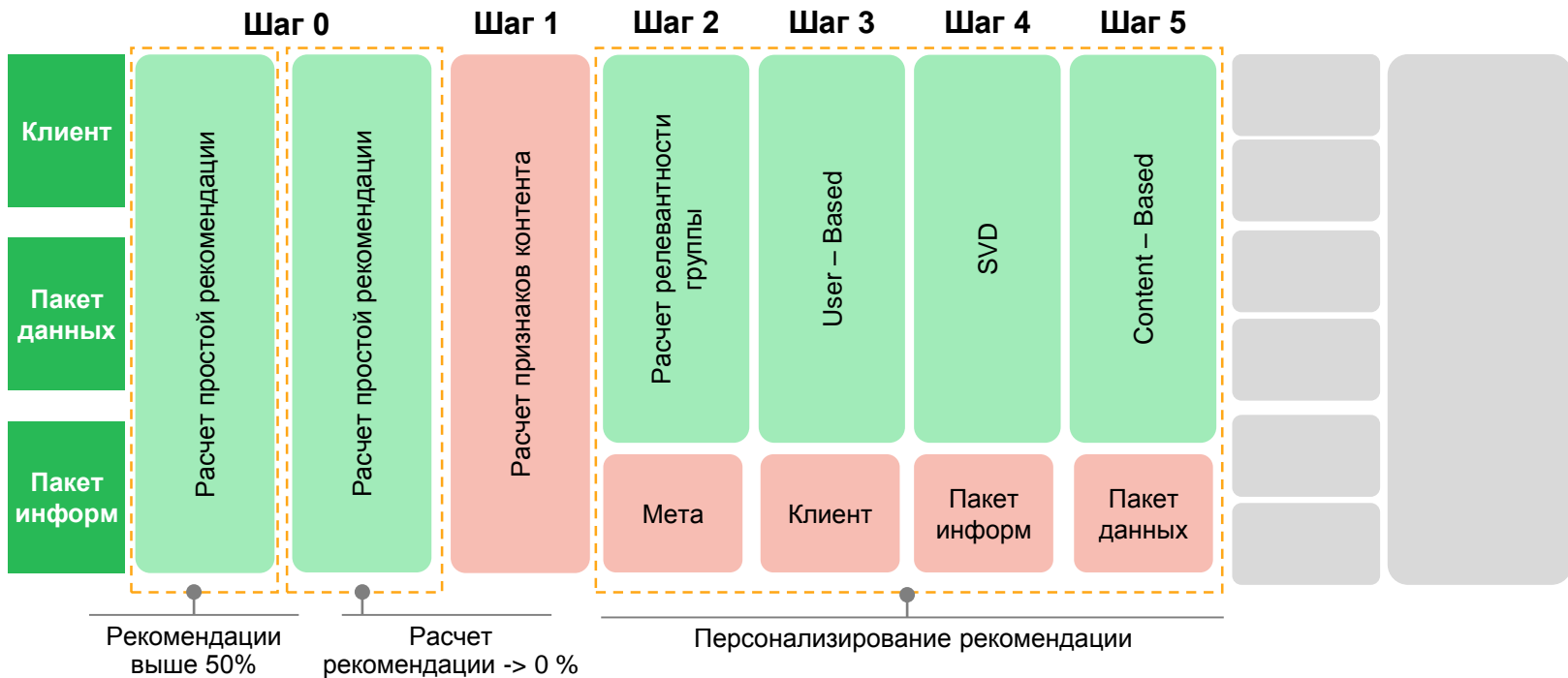
Данных действительно много



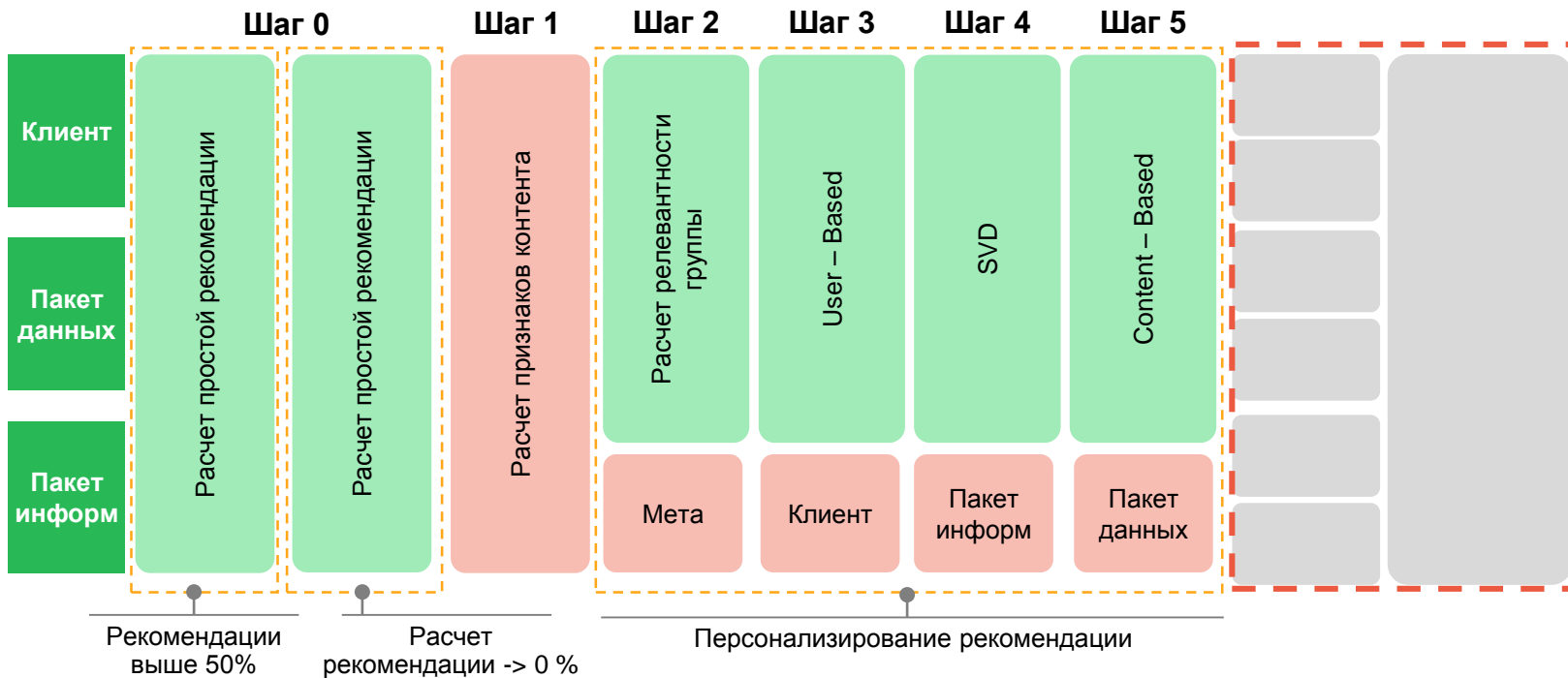
Данных действительно много



Данных действительно много



Данных действительно много. Но чего-то не хватало



2

Может быть не хватало...

Может быть дополнений?

1. Учет неявных предпочтений IALS

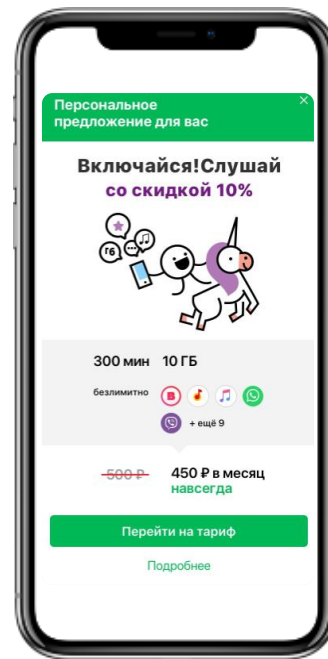
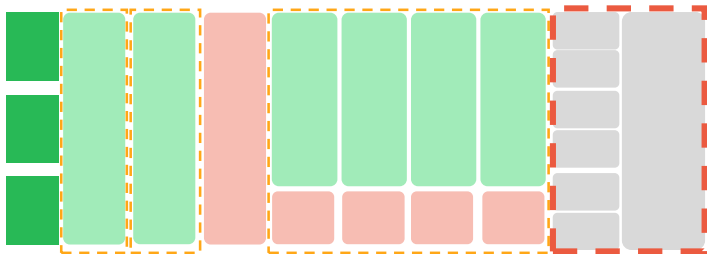
Может быть дополнений?

1. Учет неявных предпочтений IALS
2. Связь между объектами
(один плох, а два хороши)

Может быть дополнений?

1. Учет неявных предпочтений IALS
2. Связь между объектами
(один плох, а два хороши)
3. На вход алгоритму подавать агрегат
(CTR, доли данных, временные данные, отношение клиент/все)

Может быть дополнений. **Результат**



3

Метрики

Метрики

1. RMSE

Метрики

1. RMSE

2. NDCG@N

Метрики

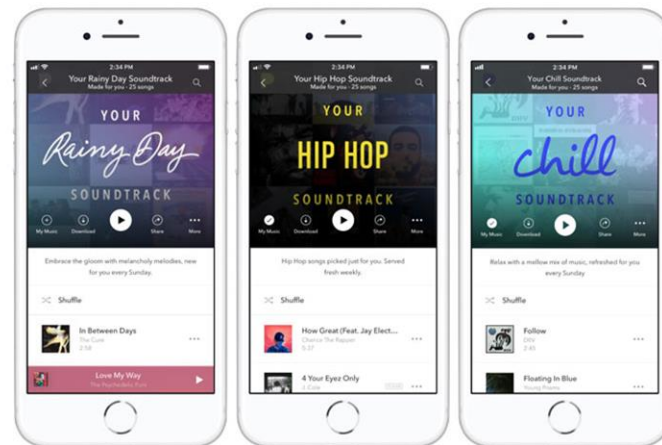
1. RMSE
2. NDCG@N
3. MAP@N
4. MAPE@N

Метрики не что – метадата все!

Метрики не что – метадата все!

1. Genome Project

2. Больше 400 атрибутов на трек

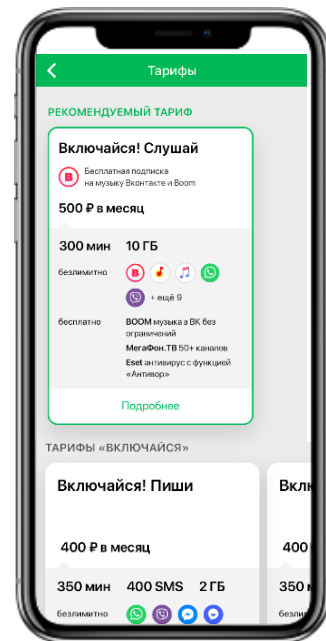
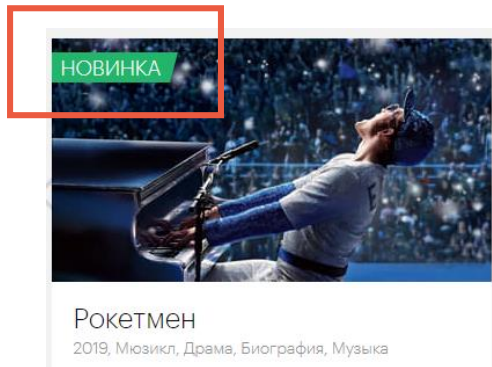


4

Метадата Контекстные данные

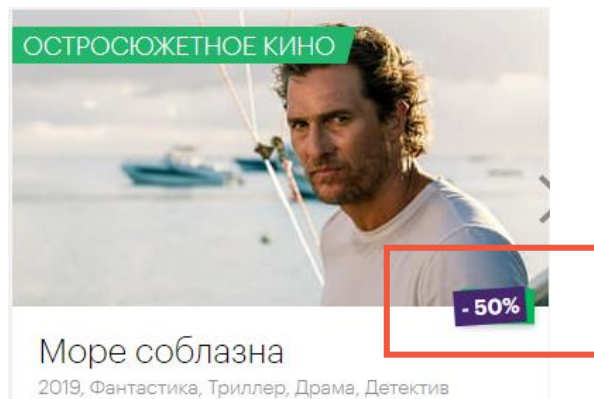
Контекст – основа рекомендаций

1. Новизна контента



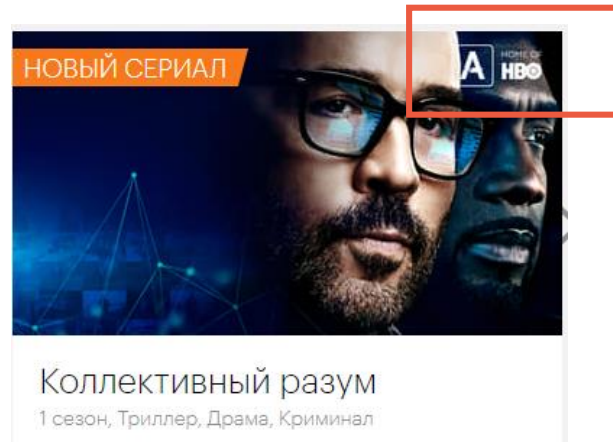
Контекст – основа рекомендаций

1. Новизна контента
2. Склонность



Контекст – основа рекомендаций

1. Новизна контента
2. Склонность
3. Качественные предпочтения



Контекст – основа рекомендаций

1. Новизна контента
2. Склонность
3. Качественные предпочтения
4. Тайминги



5

Новые проблемы контента

Проблемы контента

1. Голодание (специализированного) контента

Проблемы контента

1. Голодание (специализированного) контента
2. Каннибализация популярным контентом

Проблемы контента

1. Голодание (специализированного) контента
2. Каннибализация популярным контентом
3. Постоянство контента

Проблемы контента

1. Голодание (специализированного) контента
2. Канныбализация популярным контентом
3. Постоянство контента (не постоянство пользователя)

Проблемы контента

1. Голодание (специализированного) контента
2. Канныбализация популярным контентом
3. Постоянство контента (не постоянство пользователя)
4. Новые форматы и содержание предложений

Проблемы контента

1. Голодание (специализированного) контента
2. Канныбализация популярным контентом
3. Постоянство контента (не постоянство пользователя)
4. Новые форматы и содержание предложений
5. События по контенту

5

Новые проблемы контента и решения

Центричность дает решение

	Пользователь	Предложение
Основная задача	Что надо показать?	Кому показывать?
Метрика качества	По пользователю	По предложению
Зависимость данных	Исторические данные по пользователю	Около 1000 зависимостей

Центричность дает решение

	Пользователь	Предложение
Основная задача	Что надо показать?	Кому показывать?
Метрика качества	По пользователю	По предложению
Зависимость данных	Исторические данные по пользователю	Около 1000 зависимостей

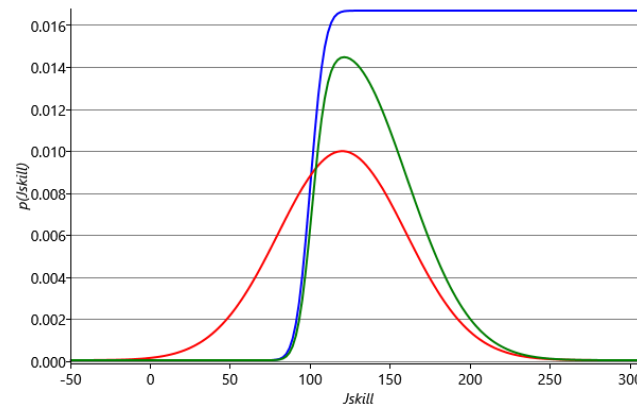
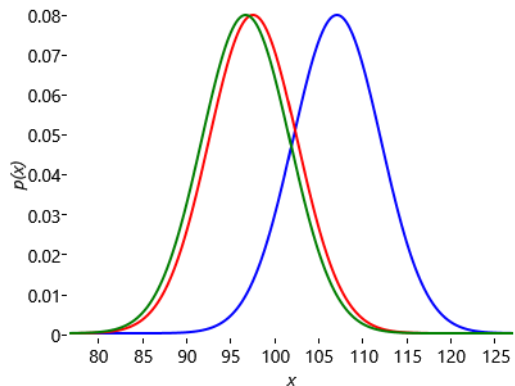
Центричность дает решение

	Пользователь	Предложение
Основная задача	Что надо показать?	Кому показывать?
Метрика качества	По пользователю	По предложению
Зависимость данных	Исторические данные по пользователю	Около 1000 зависимостей

Модель центрирования

Модель центрирования и True Skill

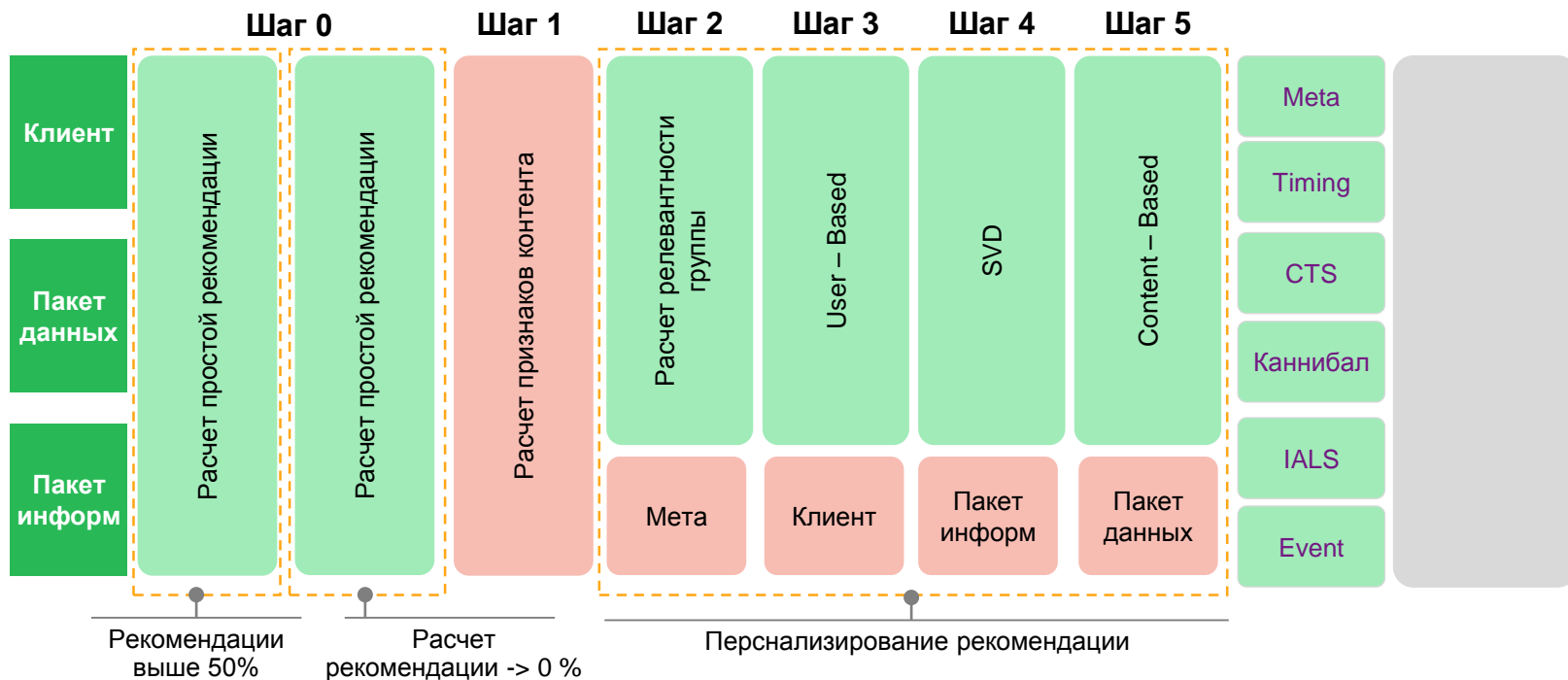
	Пользователь	Предложение
Основная задача	Что надо показать?	Кому показывать?
Метрика качества	По пользователю	По предложению
Зависимость данных	Исторические данные по пользователю	Около 1000 зависимостей



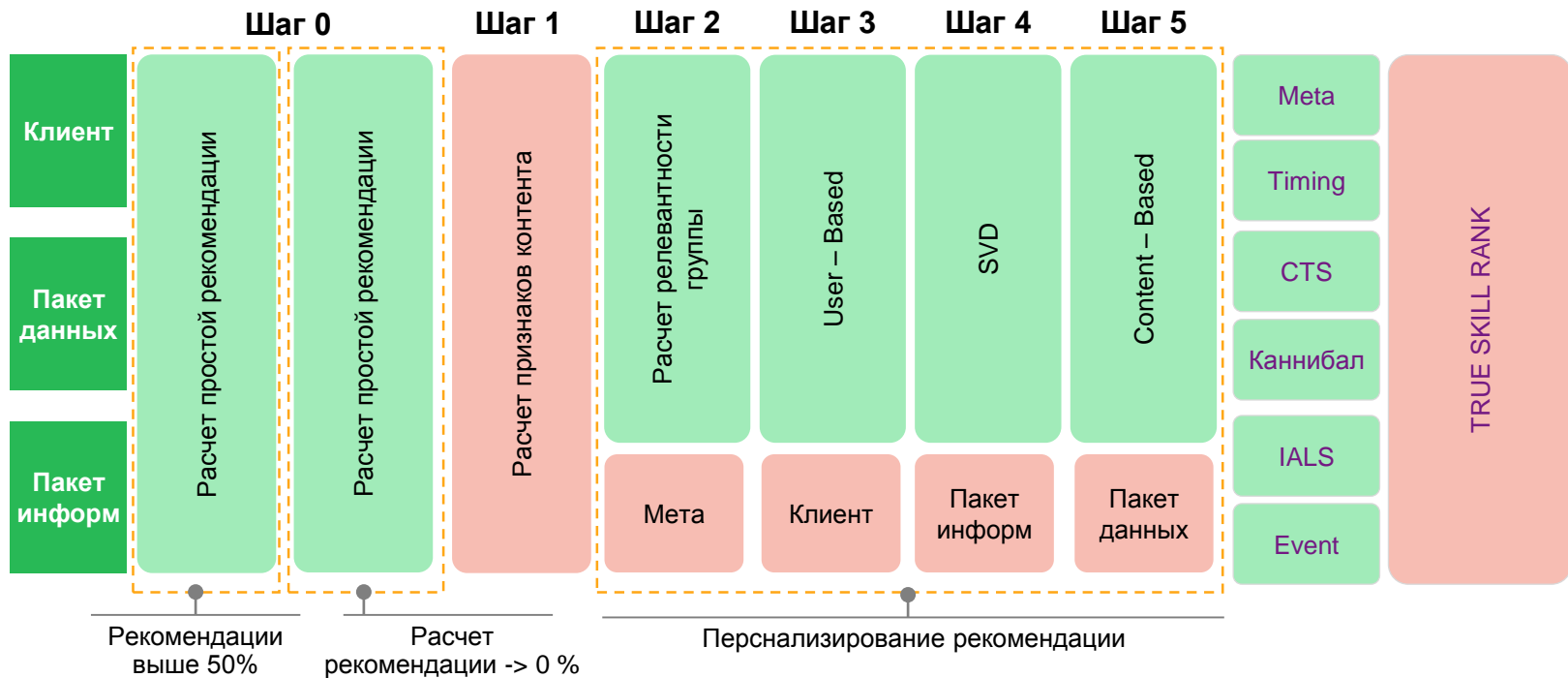
6

Добавим результат

Сложение рангов



Финальный ранг



Получаем нужный результат!



Рекомендации начинаются с тебя



Артем Селезнев, Data Scientist

@SeleznevArtem

