

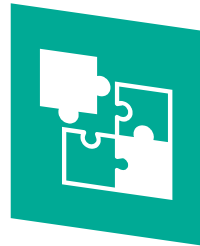
Mediascope Data Platform & Datalab

Михаил Цуприков
/ Директор по мобильным
технологиям и продуктам
Mediascope

Mediascope Data platform / Концепция



Уникальное технологическое решение для хранения и обработки в едином виде разнородных данных о медиапотреблении и потребительском поведении – как собственных данных Mediascope, так и сторонних данных, загружаемых клиентами



От анализа и поставки предрасчитанных данных по каждому медиа к работе со всем массивом «сырья» и агрегатами в одном «окне» (DataLab). Ускорение, углубление и персонализация аналитики



Открытая платформа с индустриальной моделью данных для ускорения трансформации отрасли и переходу к data driven моделям

Индустриальная модель данных

КТО?

Субъект, идентификация: опрос, куки, логи, устройства...

ЧТО?

Объект = элемент контента/ рекламы, с которым контактирует субъект

ГДЕ?

Обстоятельства контакта

КОГДА?

Точное время контакта



Под капотом платформы



Инструменты для анализа и агрегирования данных, хранения получаемых агрегатов, обеспечивает простой и надежный доступ к этим агрегатам



Платформа является высокопроизводительной системой, функционирующей в условиях большого числа параллельных процессов

Технологический стек

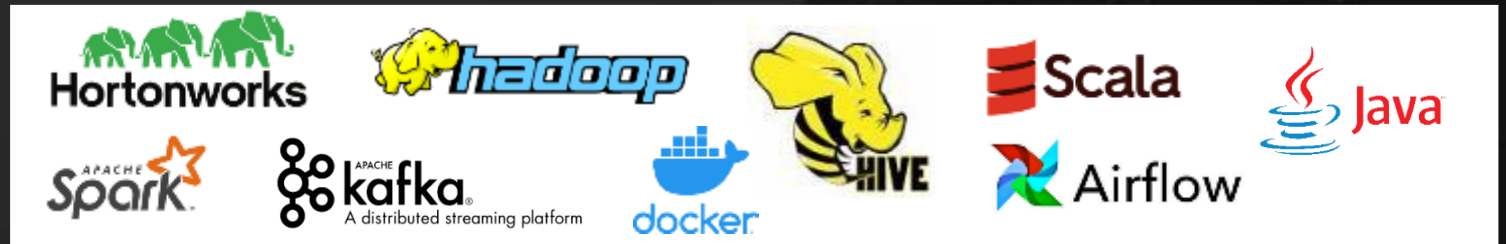
Инфраструктура

- Open source
- 2 data центра
- Облачное хранилище для больших данных Sbercloud в партнерстве с Huawei

Клиенты

- Data Platform используется в коммерческой поставке данных с 2020г.
- Клиенты уже работают с Datalab в тестовом исследовательском режиме

Платформа



Инструменты анализа



Документация и Metadata



Данные для анализа

01



Кластер

- 2 Пбайт в год / прирост 1-5 Тбайт в день

02



Набор загружаемых в платформу датасетов

- TV нац. рейтинг, посекундные логи (респонденты)*
- TV рег. Рейтинг, посекундные логи (респонденты)*
- Brand Pulse*
- Web Index**
- Web - просмотр видео**
- Mobile Apps / E-com (Android)**
- Мониторинги**

03



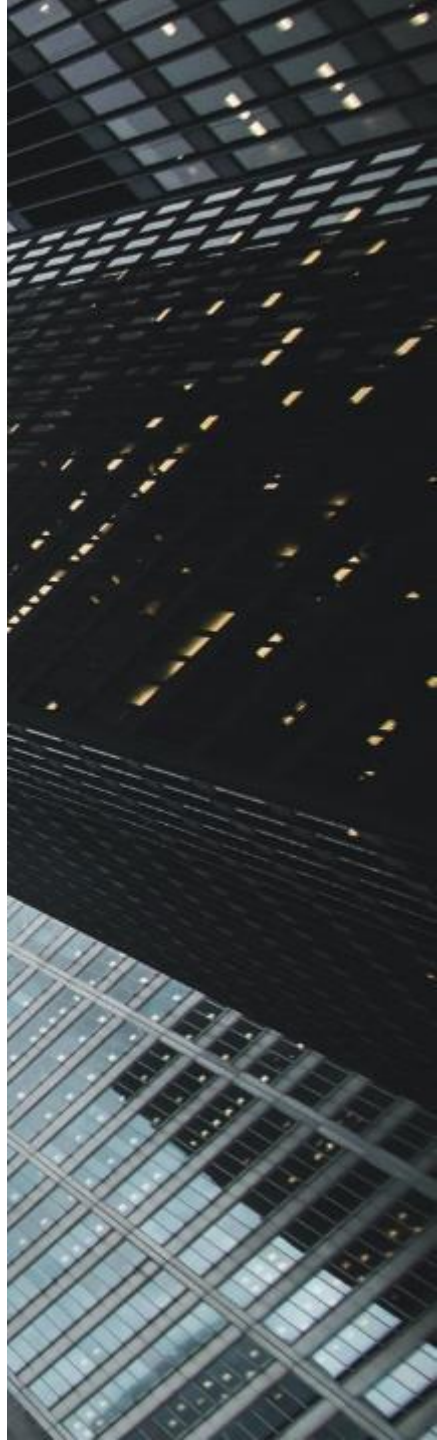
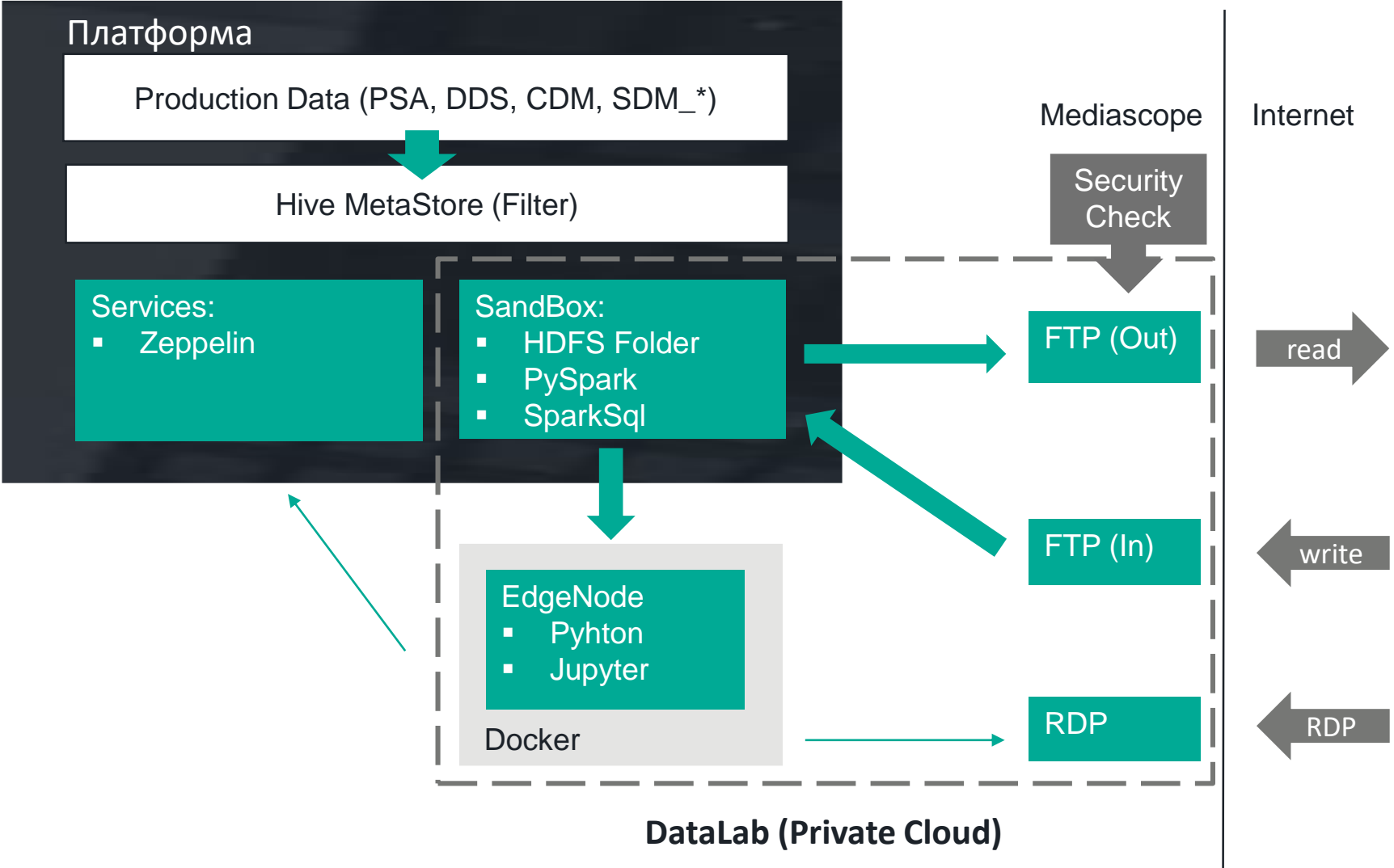
Сервисы

- Fusion as a service*

Процесс поставки данных потребителям



Архитектура платформы и Datalab



Результаты трансформации

Единая структура событий: «Event log»
Предметно-ориентированная аналитика

Упрощение доступа к данным

Разделение задач подключения новых источников
данных и создание аналитических продуктов

Снижение time-to-market
новых продуктов

Централизованный сервис контроля качества данных
и организация процесса Data Governance

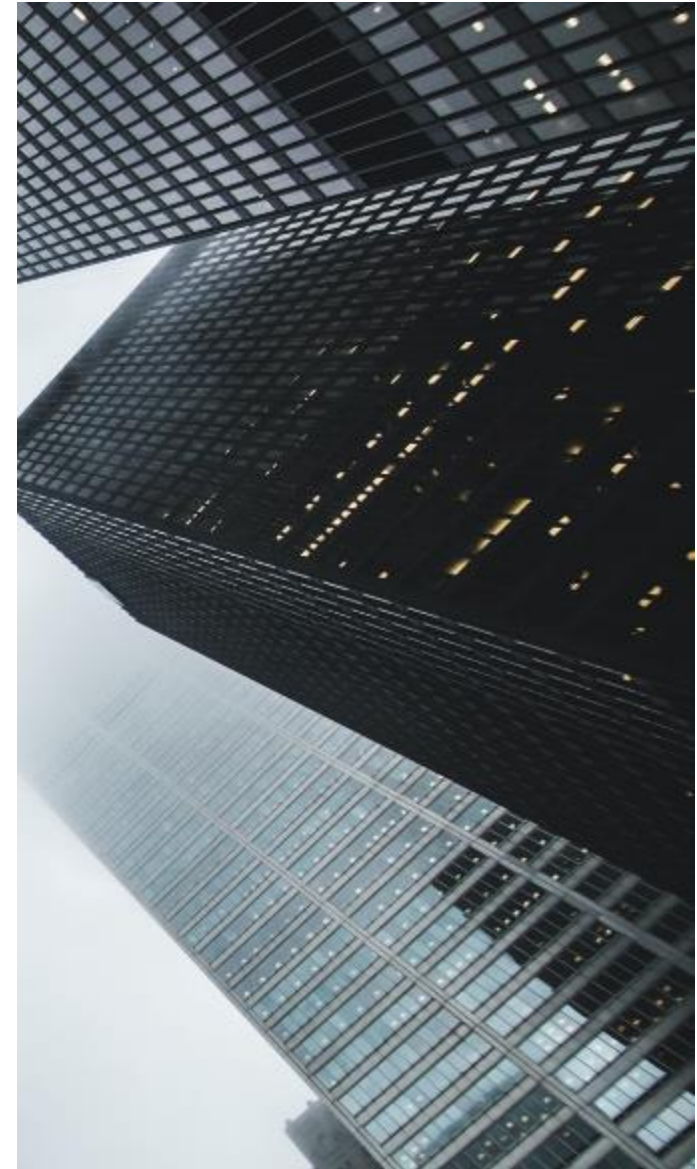
Управляемое качество
аналитики

Единый слой интегрированных данных платформы,
Индустриальная модель данных

Кросс-медийная аналитика

Все входящие транзакционные данные
обрабатываются в потоковом режиме

Near real-time аналитика



DataLab

Исследовательская лаборатория для дата-сайнтистов и аналитиков

Цели:

01

Повышение качества данных и вычислений в текущих клиентских моделях

02

Улучшение качества клиентских размещений (планирования, оптимизации и т.д.)

Задачи:

03

Перенос работы с данными в лабораторию для поиска новых инсайтов

04

Разработка принципиально новых алгоритмических big data продуктов

Ключевые преимущества для клиентов



Непротиворечивые данные по всем проектам в одной песочнице



Обогащение данных Mediascope данными из других источников

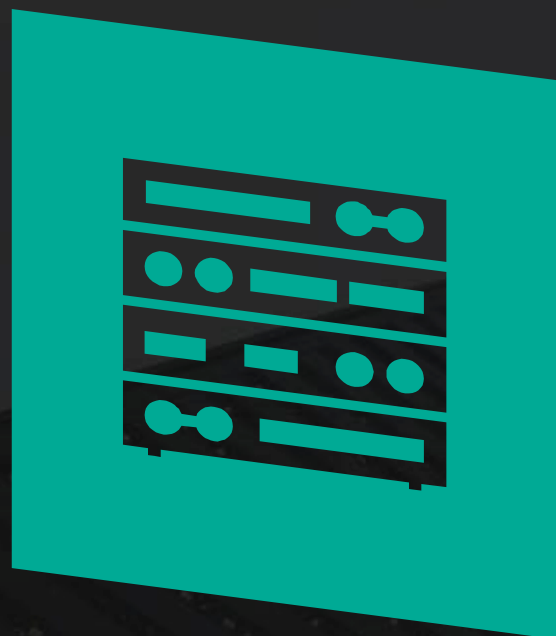


Инструменты быстрой аналитики для проверки новых гипотез и поиска закономерностей



Open API / Гибкость и персонализации в доставке данных

Система доступа и поддержки клиентов



Доступ:

- Через удаленную виртуальную машину клиент получает доступ к Private cloud и своему профилю* в рамках объема подписки на данные Mediascope

Поддержка:

- Запущен клиентский портал для поддержки и обучения клиентов работе в лаборатории
- Подготовлена и опубликована клиентская документация с описанием алгоритмов расчетов и преобразований

*Профиль – набор данных, приобретенных клиентом в рамках действующих контрактов с Mediascope (TV Index, WEB Index, Brand Pulse, мониторинги и т.д.)

Принципы использования DataLab



Загружаемая в лабораторию информация и расчеты не должны использоваться для идентификации панелистов и персональных данных пользователей



Результаты расчетов в лаборатории не могут использоваться для формирования сегментов для загрузки в рекламные системы



Отсутствие логов digital партнеров Mediascope



Отсутствие данных других клиентов Mediascope

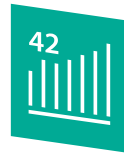


Варианты использования данных Mediascope



ПРОГНОЗЫ

- рейтинги
- оптимизаторы кросс-медиа охвата
- видео планировщики



ДЕТАЛИЗАЦИЯ

- посекундный лог тв + мониторинг
- доступ к предварительным данным расчетов
- просмотр промежуточных данных всех этапов производства (взвешивание, фьюжн, агрегация и т.п.)



АНАЛИТИКА

- ретроспективный анализ
- расчет любых собственных показателей (по своим целевым и т.д.)

Возможности пересечения данных Mediascope

+ **Client Data*** (1st /2nd party, CRM, Access panels)

- Media Mix Modelling
- Custom advertising results modelling
- Fusion CRM
- и т.д.



+ **3rd party data*** (RPD, ОФД, X5, etc)

- Эконометрика
- Планирование на потребительских сегментах
- Верификация точности попадания кампаний в целевую аудиторию (в т.ч. по поведенческим таргетингам)
- Атрибуция на продажи
- и т.д.

Datalab / Кейс

01

Эксперты dentsu Russia проанализировали посекундный телевизионный эфир

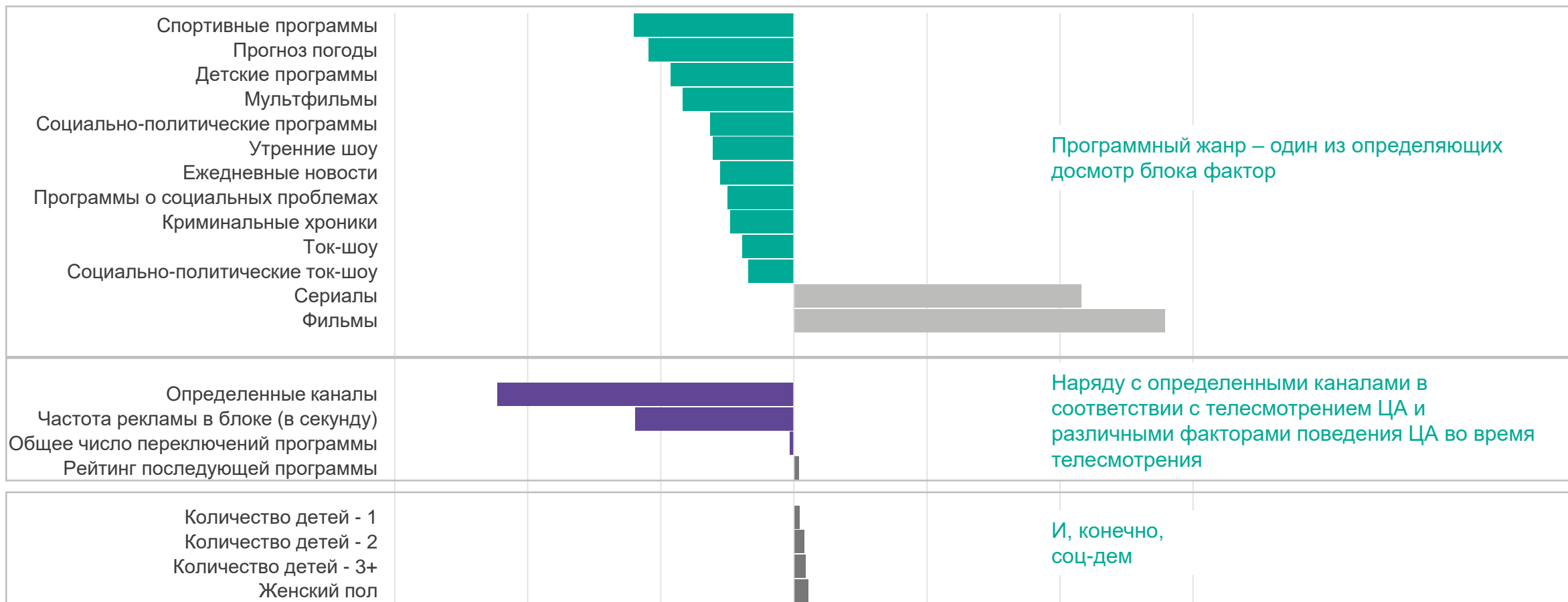
02

Построили модель, объясняющую, какие факторы телесмотрения влияют на досмотр рекламного блока



Кейс / Результаты работы модели*

Индексы влияния на базовый досмотр ТВ- блоков, все 100+



Резюме

- Происходит сращивание бизнес-консалтинга и исследовательской индустрии
- Рекламный рынок трансформируется и начинает развиваться за счет конкуренции data science специалистов, а не баинговых условий
- Выигрывает не тот, у кого больше данных (а их сейчас много у всех), а тот, кто может их правильно комбинировать и интерпретировать, разрабатывая новые коммерчески обоснованные алгоритмические продукты как для себя, так и для бизнеса своих клиентов

Спасибо!

