

Холистические
клиентские профили:
возвращаем смысл в
маркетинговые данные

Давид Вачадзе

INK Agency

Brand Mobile

Brand Mobile: краткая история большого пути



248 проектов

15 лет на рынке

17 CRM стратегий

Опыт работы с брендами различных студий

FMCG
CPG

ТЕЛЕКОМ
E-COMMERCE

КОСМЕТИКА
АВТОМОБИЛИ

Всестороннее управление CRM

Pre-campaign

Управление всеми стадиями проектов

Регулярная аналитика

Персональное ведение потребителя

Ad-hoc

Новые CRM форматы на рынке

Моделирование достижимости KPI

CRM action план

Собственная модель сегментирования

В 2006 году провели самый крупный CRM проект

накопительной лояльности в России при 4,5 млн активных участников

ЗНАЕМ, как достичь:

152% рост потребления

54% рост лояльных

50% улучшение ROI на медика

31% рост конвертации свитчеров

*Результаты одного из проектов

Неоднократно первыми в России

реализовывали уникальные проекты

в 80% проектов превышаем заявленные KPI

УМЕЕМ анализировать данные и повышать конверсии

Конверсия исторических баз на 30% и рост потребления на 70% за 1 год
*Результаты одного из проектов

Holistic Engineering CDP What ??

- Holistic
 - belief that the parts of something are intimately interconnected and explicable only by reference to the whole
- Engineering
 - creative application of science, mathematical methods, and empirical evidence to the innovation, design, construction, operation and maintenance
- Customer Data Platform
 - Marketer-managed system that creates a persistent, unified customer database that is accessible to other systems
 - System that centralizes customer data from all sources, unifies this data into customer profiles and then makes this data available to other systems for marketing campaigns, customer service and all customer experience initiatives.

С какой стороны подойти

1. Целеполагание, задачи и юз-кейсы
2. Источники данных
3. Набор метрик и способы их расчета, сегментирование на основе метрик, выделение KPI
4. Сбор и препроцессинг сырых данных, realtime, online, offline
5. Структура хранения профайлов потребителей
6. Инфраструктура использования данных профайлов потребителей

Пример целеполагания для потребительских рынков

- Создание и автоматическая актуализация 360⁰ профиля потребителя
 - базирующегося на фактических данных,
 - маркетингово-интерпретируемого и валидируемого
 - Интегрирующего соц-дем, транзакционную и/или поведенческую активность в различных онлайн и оффлайн каналах
- для прозрачной оценки и измеримого повышения эффективности бизнеса, включая
 - Продажи (Sales)
 - Вовлечение в коммуникацию с брендами (Awareness)
 - Лояльность (Loyalty)

Общая логическая структура данных

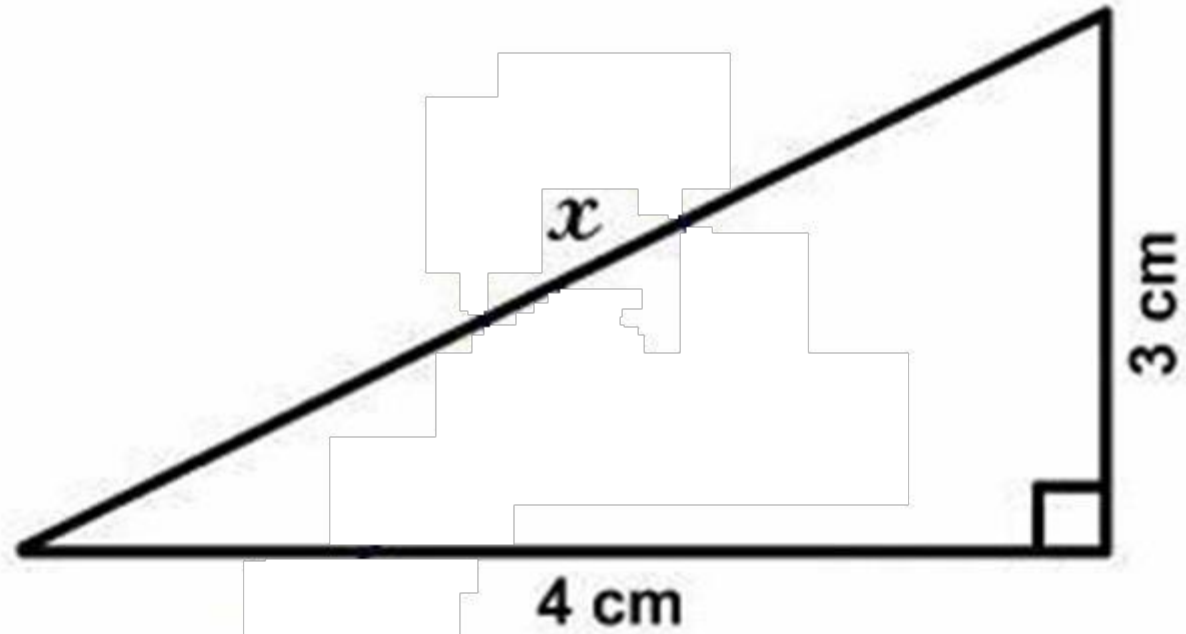
- Структурирование (справочники + параметры)
 - Проекты: вся активность потребителя структурирована по проектам
 - Продукты: иерархия категория-бренд-продукт-SKU, проекты привязаны к продуктам
- 1. Анкета потребителя
 - Базовая: введенная потребителем при регистрации – соц-дем, контакты, согласия
 - Расширенная: источники, качество контактов
- 2. Метрики персоны: кросс-проектные интегральные метрики и KPI-и
- 3. Метрики проектов: метрики, сегменты и KPI потребителей в рамках проектов
 - Общая информация: даты, активности, потребления
 - Блок потребления: продукты, объемы и частоты в покупках и в потреблении
 - Коммуникация: интегральный профиль и активность по pull и push
 - Остальное: Loyalty, DMP, Content, KPI-и
- 4. История (лог/динамика изменений)
 - проекты, метрики, коммуникации

Общая технологическая структура данных

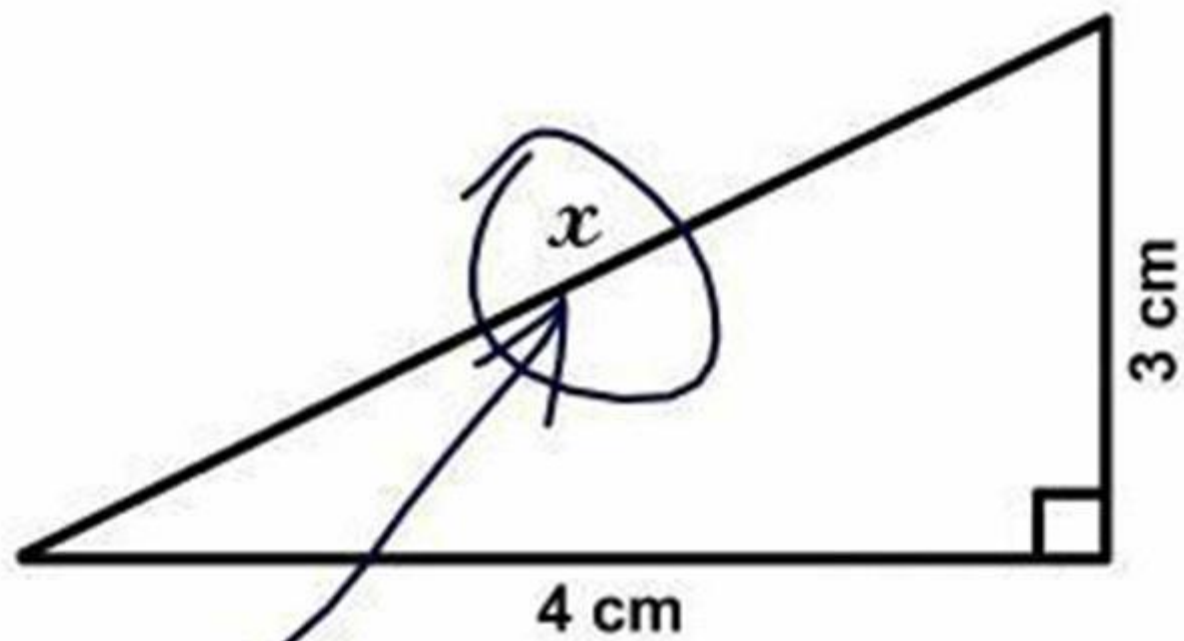
1. Исходные (сырые) данные (поведенческие и транзакционные), лежат рядом
 - Более-менее доступны, может потребоваться пересчет метрик или исследование
 - Например в GBQ
2. Инфраструктура (пере)расчета метрик
 - Алгоритмы могут быть непростые, нужна среда исполнения
3. Хранилище профайлов: персона + метрики + сегменты
 - Часто имеет не-реляционную структуру данных
 - Обычно работает в реальном времени
4. API
 - Сбор данных – профайлы и метрики
 - Отдача данных – профайлы и метрики
5. GUI

Примеры расчета нетривиальных метрик

3. Find x .



3. Find x .



Here it is

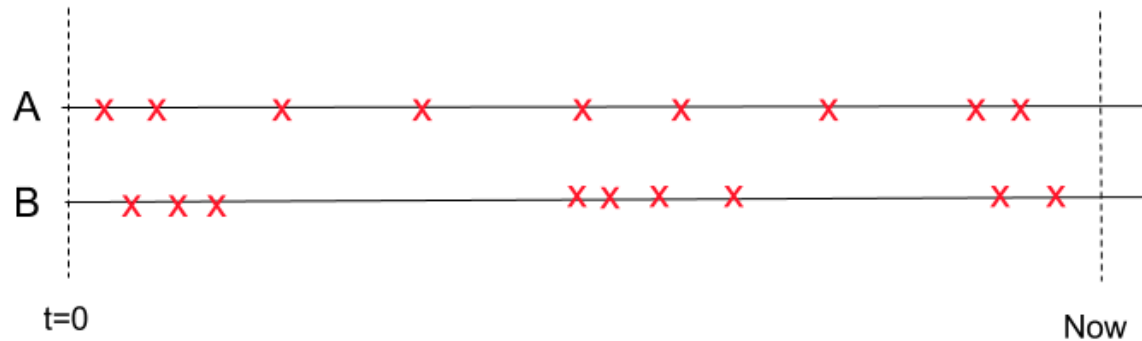
Распределения финансовых данных

- Метрики распределения денежных сумм (чеки, денежные потоки и пр.)
 - Нормализация через лог-трансформацию
 - Учитываем возможность нулевых чеков
 - Учитываем отрицательные величины
- Получаем

$$\text{sign}(x) * \log_{10}(1+\text{abs}(x))$$

- Использовать только при распределениях, часто важны абсолютные значения

Метрики «нерегулярности» (потребление и пр.)



$$1 + \frac{\sum_{i=1}^{n+1} \log\left(\frac{\Delta t_i}{N+1}\right) \frac{\Delta t_i}{N+1}}{\log(n+1)}$$

- Энтروпийная метрика, кастомизируется под конкретную задачу
- Например для товаров или услуг регулярного потребления параметры следующие:
 - n = общее количество покупок `total.consumption.n`, N – период между `reg.dt` и `last.consumption.dt`, выраженный в сутках, $\Delta t_i = t_i - t_{i-1}$ есть разница между таймстемпами последовательных регистраций покупок, выраженная в сутках, момент времени t_0 соответствует `reg.dt`

Частотность потребления для FMCG



Коммуникации: универсальные метрики

- Как сравнить email, sms, сайты, мобильные приложения и чат-боты?
- Например так – универсальная метрика коммуникационной сессии:
 - Тип коммуникации (pull, push)
 - Код канала (email, sms, web, mobile.web, mobile.push, web.push, chatbot, mobile.app)
 - Subject: тематика коммуникации
 - Таймстемп старт-стоп коммуникации
 - Количество активностей потребителя в коммуникации
- А теперь сформулировать способы расчета этой метрики для разных каналов и типов коммуникации
 - Это не всегда просто)

Consumer Journey: это состояние или переходы?

- «Какая аналитика? У нас есть Google Analytics for everything»
- Это оба, объединенные в цепочку по времени
- Как хранить?
 - использовать стандартные форматы SPS, TSE, SPELL и др.
- Как сегментировать по CJM, рассчитывать дискриминанты и пр.?
 - В CDP – никак, потом нужны внешние «инфраструктуры расчета метрик»
 - Есть метрики «расстояние» между CJ, «похожести» CJ
 - Дискриминирующие события или цепочки – шаг от «наблюдения» к «управлению» CJ

Интегральные метрики: Персона

- Sales ~ RR или Sales ~ доля NLHP * инкрементальный прирост NLHP,
 - во втором случае это какая доля аудитории постоянных потребителей на сколько увеличила свое потребление
- Awareness ~ $\sum N_{\text{участники}}(\text{push_сегмент} + 2 * \text{pull_сегмент}) * \text{дисконтированный_recency}$,
 - т.е. это сумма людей CLC: {New, Active, About to Lapse} по сегментам push (открывающим коммуникации) и pull (вовлекающим с активности бренда) с удвоенным весом, с учетом давности соответствующих действий (старые открытия входят с меньшим весом чем свежие открытия. Важно что тут люди, активно и регулярно участвующие в коммуникации, а не контакты, во втором случае были бы предпосылки к неконтролируемому росту частоты рассылок
- Loyalty ~ $\sum N_{\text{участники}}(\text{pull_сегмент}) * \text{дисконтированный_recency} + \sum N_{\text{участники}}(\text{NLHP})$
 - Лояльные потребители демонстрируют постоянное потребление + участие в вовлекающих активностях брендов в сейчас или в недавнем прошлом, кроме того не входили в сегмент А на каком-либо бренде

Примеры технологической реализации

Интеграция данных на платформе Marline

Сбор, обработка и накопление данных

индивидуальные потребители в не-агрегированном и не-сэмплированном виде

из разных источников, включая Web, Email, Backoffice, Call-center

на протяжении всего жизненного цикла

доступные для анализа в реальном времени

Идентификация и мэтчинг пользователей

в России по данным TNS люди выходят в интернет в среднем с 2-3 устройств

при этом используя до 10 онлайн личностей

идентифицируем анонимов

включаем email ремаркетинг

работа с аудиториями, а не куками

Расширенная кросс-канальная аналитика

какие люди или сегменты как реагируют на рассылки

когда читают

как и когда кликают

ходят на сайты по клику в рассылке или просто после прочтения

что делают на сайте

как долго длится эффект от рассылки

Marline: realtime data collection & processing

Google Cloud Platform | Marline | Multi-PID events demo

DASHBOARD | ACTIVITY

Project info

Project name: Marline
Project ID: marline-bm
Project number: 904070640889

Go to project settings

Resources

- App Engine: 0 versions
- Compute Engine: 3 instances
- Cloud Storage: 9 buckets
- BigQuery: 12 datasets

Trace

No trace data from the past 7 days

Compute Engine

CPU (%)

instance/cpu/utilization: 0.336

Go to the Compute Engine dashboard

RPI APIs

Requests (requests/sec)

Query history

Query editor: rawp

This is a partitioned table. [Learn more](#)

Schema: Details | Preview

Description: None | Labels: None

Table info

Table ID	marline-bm:default:rawp
Table size	222.08 GB
Long-term storage size	221.11 GB
Number of rows	552,852,688
Created	Sep 30, 2017, 10:03:05 PM
Table expiration	Never
Last modified	Sep 5, 2018, 1:47:39 PM
Data location	EU
Table type	Partitioned
Partitioned by	Day
Partitioned on field	_PARTITIONTIME
Partition expiration	2400 days

Streaming buffer statistics

Estimated size	697.2 KB
Estimated rows	1,828

Resources

Search for your tables...

- marline-bm
 - BM
 - default
 - raw
 - raw_backup
 - rawp**
 - DT
 - manual
 - moneymatika
 - results
 - SC
 - SD
 - stage_default
 - stage_testacc
 - stoloto
 - winline

Технологические требования к CDP

- Гибкость структуры данных
 - Никогда не угадаешь какие новые коллекции метрик возникнут (noSQL?)
 - Алгоритмы расчета – живые и эволюционирующие (версионирование метрик?)
- Multi-PID
 - Потребители с 5 имейлами или без имейлов – это нормально
 - Разные или множественные Social_ID – это нормально
 - (Не)идентифицированные анонимы - ? Да, и это нормально
- Доступ к сырым данным нужен часто
- Realtime нужен практически всегда
- Data integrity & consistency валидация и очистка - всегда

Marline Firebase CDP-demo

The screenshot displays the Firebase console interface for a project named "CDP-demo". The main header shows "Database" and "Cloud Firestore BETA". The left sidebar contains navigation options: Project Overview, Develop (Authentication, Database, Storage, Hosting, Functions, ML Kit), Quality (Crashlytics, Performance, Test Lab), Analytics (Dashboard, Events, Conversions, Au...), and Grow (Predictions, A/B Testing, Cloud Mes...).

The main content area shows the database structure for the "users" collection. The breadcrumb path is "users > jznNzRkV48Mx...". The tree view shows the following structure:

- cdp-demo-bfc1e
 - users
 - + Add collection
 - products
 - projects
 - users >
 - + Add document
 - jznNzRkV48MxoXAQeMu7 >
 - m4huhMcVTKTMGzSykEYp

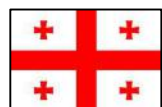
The selected document "jznNzRkV48MxoXAQeMu7" is shown with the following fields:

- + Add collection
 - Consumption
 - GEO
 - PID
- + Add field
 - CDP_ID: "123456789"
 - reg_date: July 5, 2018 at 3:03:12 AM UTC+3
 - source: "123"

Marketer-managed CDP?

Data Literacy как важное требование к «ЖИВЫМ» пользователям CDP

- Data-взгляд на повседневную деятельность маркетолога
- Уметь выделить измеримые метрики в решении бизнес-задачи
- Понимать как устроен маркетинговый эксперимент и проверка гипотез
- Тестирование и достоверность результата, как (не)поверить в цифры агентства
- Как читать графики
- Как ставить задачу аналитическим агентствам, и оценивать качество работы



Выездной воркшоп «Data Literacy for Marketers»

- 3+2 дня в Грузии: Тбилиси, Батуми, окрестные винодельни
- 4 академический пары Data Literacy ежедневно, экзамен, сертификат
- 2 эно-гастрономических пары ежедневно, экзамен, сертификат :)

