

# Мастер данные

Узнаем о том, что такое мастер-данные и как с ними работать:

- Зачем нужны мастер-данные.
- Какие бывают мастер-данные.
- Клиентский MDM (или CDI):
  - Почему CDI выделили в отдельное направление.
  - Ключевые характеристики.
  - Место в ландшафте компании.
  - Почему CDI не бывает без DataQuality
  - Что требуется от DataQuality.



H F Labs

**Никита Назаров**

*Архитектор проектов внедрения MDM*

Опыт работы в IT — 10 лет. С 2012 года в области MDM и DataQuality.

В роли архитектора успешно выполнил более 10 проектов в России по очистке и интеграции клиентских данных.

Запустил клиентский MDM в компаниях ВТБ Страхование, Метлайф, ВТБ Страхование жизни, Банк Открытие, Газпромбанк, Деловые линии.





H F Labs

**Павел Абдюшев**

*Директор по развитию продуктов*

С 2004 года работает в IT. Прошел путь от ручного тестировщика до руководителя группы автоматизации, потом заинтересовался смежными темами — аналитикой, UX и развитием продуктов.

С 2012 года занимается качеством данных, MDM и CDI-системами. Внедрил CDI в МТС, АльфаСтрахование, банк «ФК «Открытие»», Метлайф, Банк Союз.

С 2007 года разрабатывает и преподаёт тренинги по тестированию, автоматизации тестирования и SQL, анализу данных.

Автор докладов на конференциях SQADays, ЛАФ, TestLabs, ConfetQA.



## **RDM**

Фиксированные справочники, редко меняются.

## **Продуктовые (PDM)**

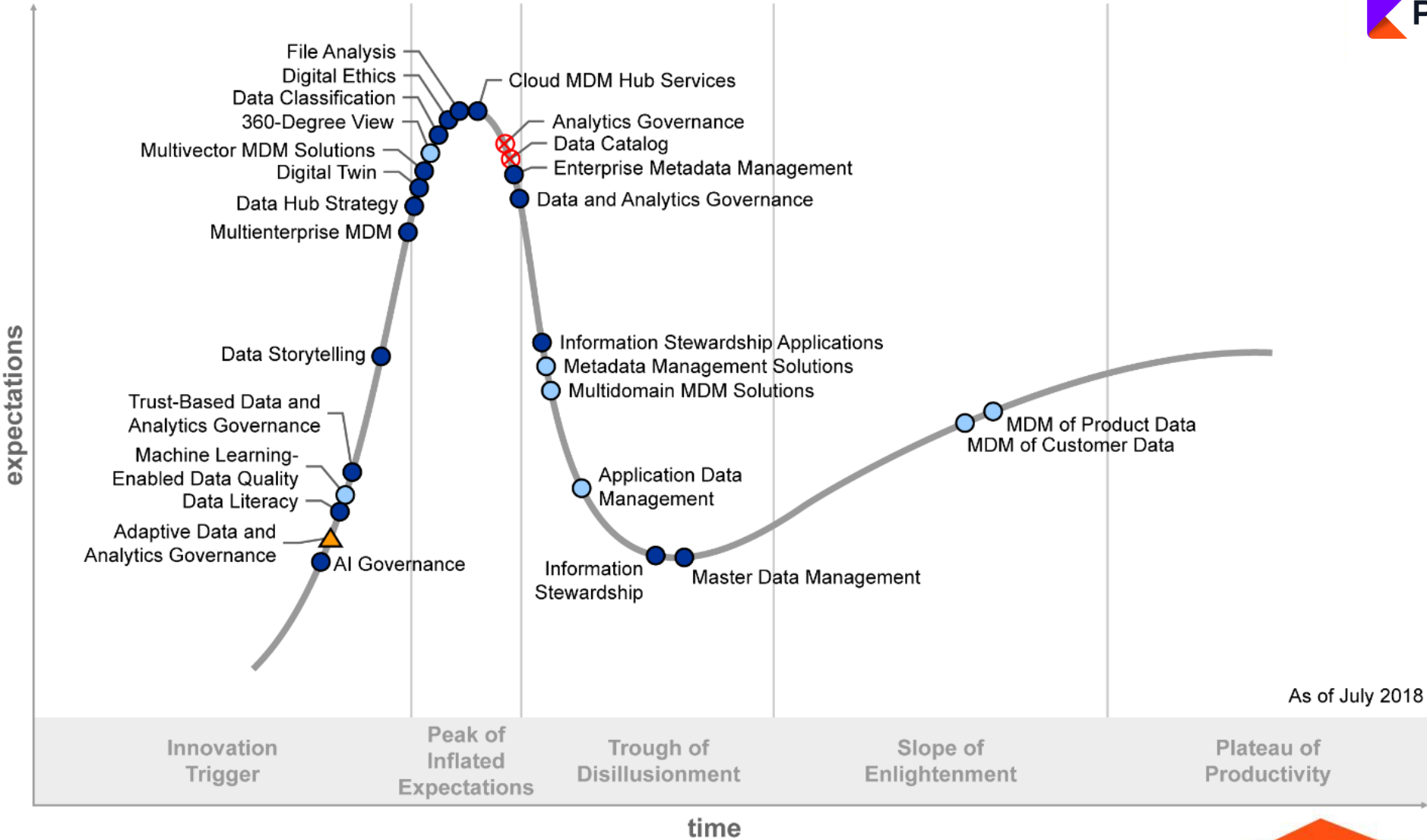
Продукты, товары, материалы, услуги.

## **Клиентские (CDI)**

Клиенты, контрагенты.

## **Мультидоменные**

Объединяют все домены.



Plateau will be reached:

- less than 2 years
- 2 to 5 years
- 5 to 10 years
- ▲ more than 10 years
- ⊗ obsolete before plateau



## Зачем CDI?



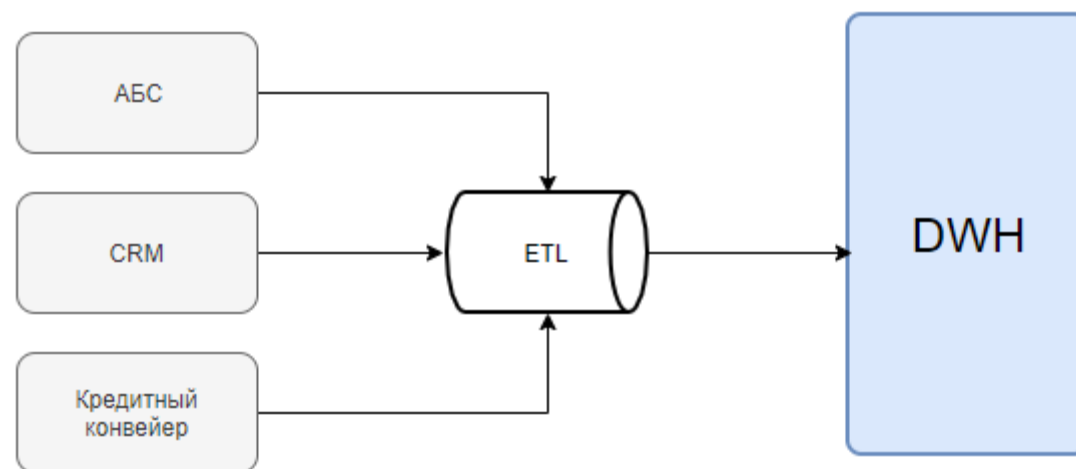
Сколько у компании клиентов?

Сколько клиентов, приносящих убытки?

Где продукты клиента?

Что вообще такое этот клиент?

Было





1

## ШАГ 1. РАЗРАБОТКА И СОГЛАСОВАНИЕ БИЗНЕС-ТРЕБОВАНИЙ

## ОПИСАНИЕ ШАГА

Бизнес-требования описывают решение задачи с точки зрения бизнес-заказчика.

Могут быть сформулированы в виде:

- Тезисов
- Пользовательских историй

В совокупности все требования к решению должны быть:

- Непротиворечивыми
- неповторяющимися
- Полными

Каждое требование должно обладать следующими свойствами\*:

- Неизбыточность
- Проверяемость
- Ясность
- Правдивость (отсутствие ложных сведений)
- Понятность формулировки
- Реализуемость
- Независимость
- Атомарность

## ВОВЛЕЧЁННЫЕ РОЛИ



Бизнес-эксперт  
Бизнес-аналитик  
Функциональный аналитик

## ОЖИДАЕМЫЙ РЕЗУЛЬТАТ



Формализованные бизнес-требования к решению (BRD)



Бизнес-кейс на реализацию решения



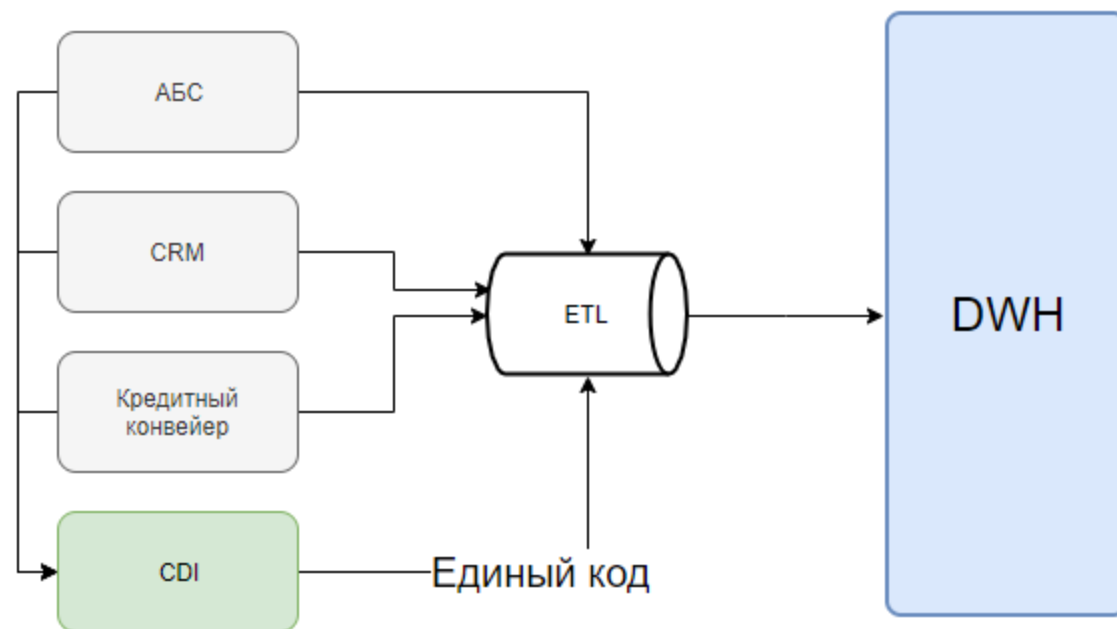
- Удобная витрина данных для построения, калибровки, пересчета моделей.
- Загрузка из других систем в витрину данных.
- Реализация алгоритма расчета всех показателей Basel III



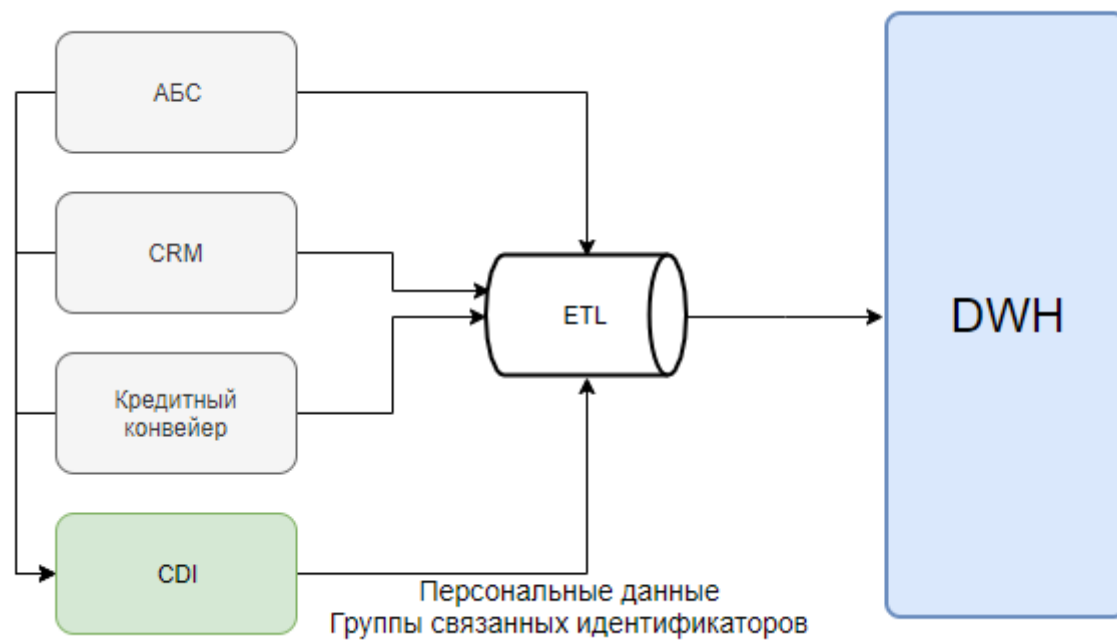
- Необходимо на ежедневной основе формировать единую клиентскую витрину на основе информации из систем Equation, Credit Fabric, CRM, MDM по следующим атрибутам...
- Сопоставление клиентов из систем осуществляется по атрибуту «Единый код клиента MDM»

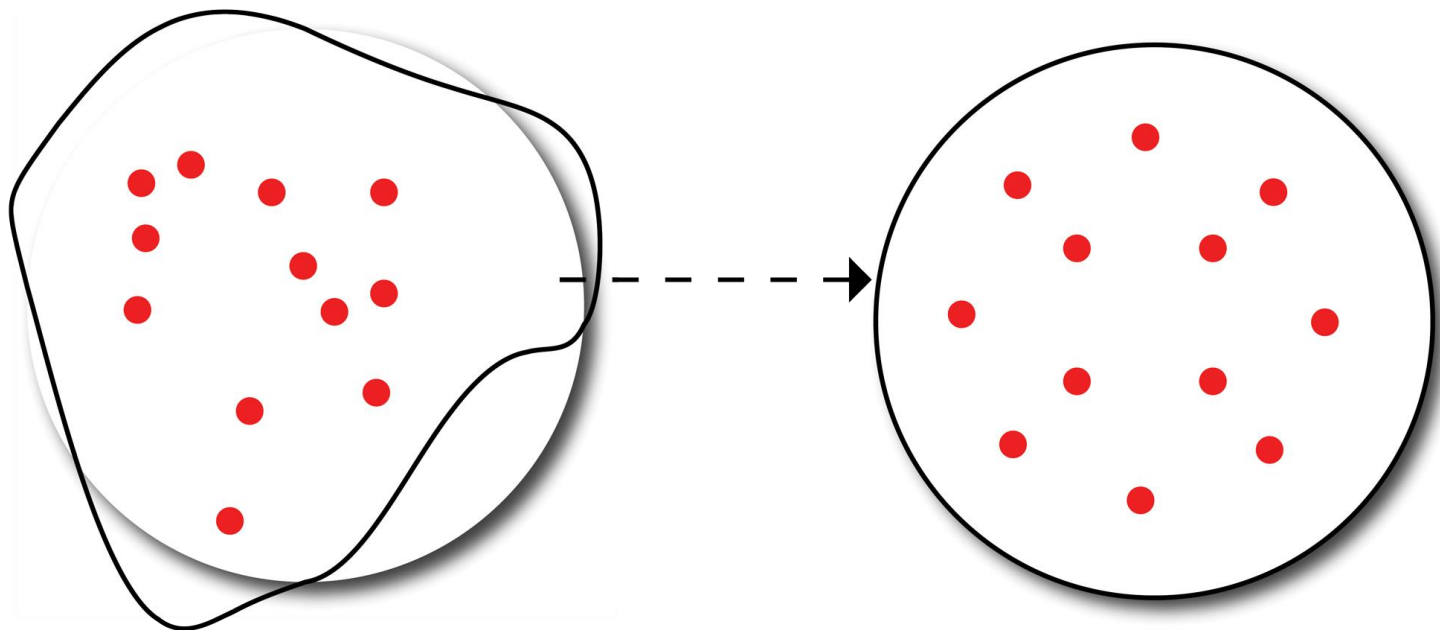


Стало?



Стало!







### **ПРОДАВАТЬ БОЛЬШЕ**

Клиентский MDM предоставляет данные для клиентской аналитики, сегментации и телемаркетинга.

### **СОБИРАТЬ ЗАДОЛЖЕННОСТЬ**

Клиентский MDM собирает все контакты клиента, помогает оценить полный размер задолженности.

### **ОЦЕНИВАТЬ РИСКИ И БОРОТЬСЯ С МОШЕННИКАМИ**

Клиентский MDM находит подозрительные изменения.

Проверяет клиентскую базу в одном месте по «чёрным» и «белым» спискам.



### ГОТОВИТЬ ОТЧЁТНОСТЬ

Клиентский MDM — источник клиентских данных для управленческой и регуляторной отчетности по клиентам.

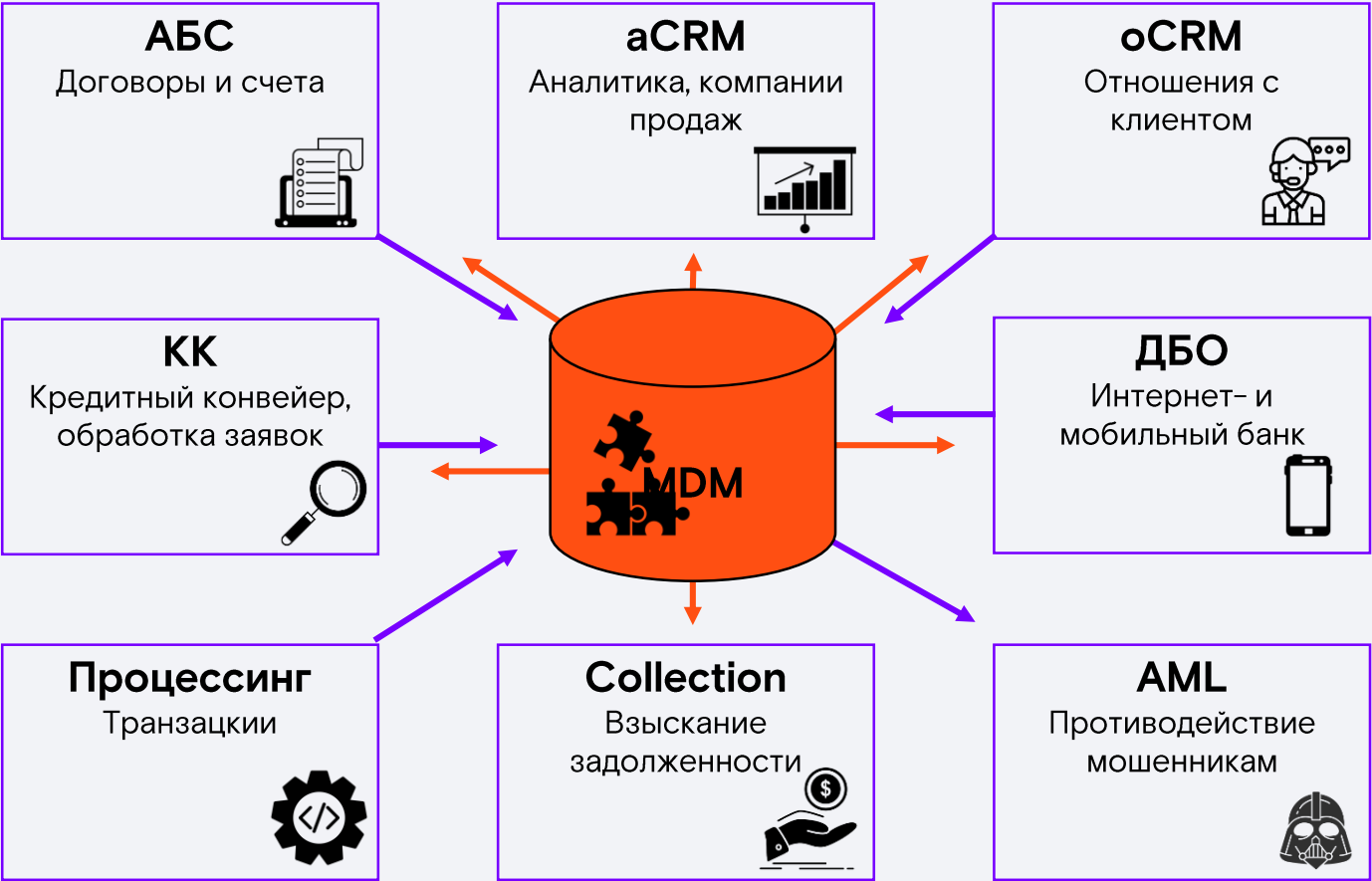
### ПОВЫШАТЬ ЛОЯЛЬНОСТЬ

Клиентский MDM помогает узнавать клиента в режиме онлайн в личном кабинете, приложении или при звонке в Call-Center.

### ПОЛУЧАТЬ НОВЫЕ ЗНАНИЯ

Клиентский MDM выявляет домохозяйства. Соединяет разные направление бизнеса: b2c и b2b.

# Клиентский MDM в банке



Знаковые проекты





## МНОГО СИСТЕМ С КЛИЕНТСКИМИ ДАННЫМИ

В больших банках часто для каждого продукта своя АБС. В России тренд на укрупнение банков. Большие банки покупают маленькие и интегрируют их в свой ландшафт.

## ФИЗИКИ И ЮРИКИ

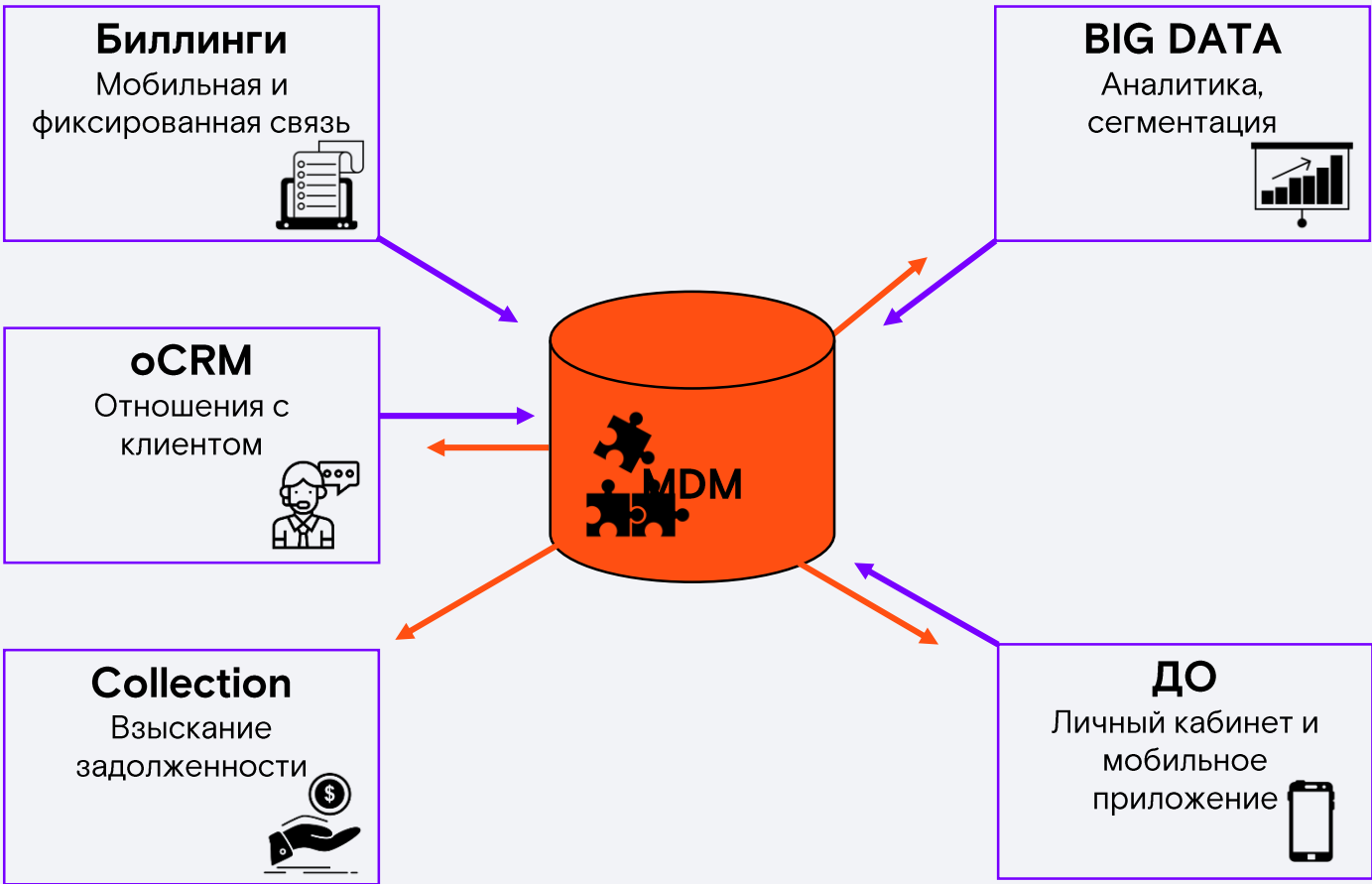
Клиент с маленьким депозитом может оказаться генеральным директором большой компании.

## РЕГУЛЯТОРЫ

Банковская деятельность очень зарегулирована. Например, Агентство по страхованию вкладов требует корректные данные вкладчиков.



# Клиентский MDM в телекоме



Знаковые проекты



Ростелеком





## ПЕРЕЙТИ ОТ SIM-КАРТЫ К КЛИЕНТУ

Оператор знает о вас всё. Но без MDM только о вашей SIM-карте.

## ХОРОШИЕ КОНТАКТЫ, НО ПЛОХИЕ ДОКУМЕНТЫ

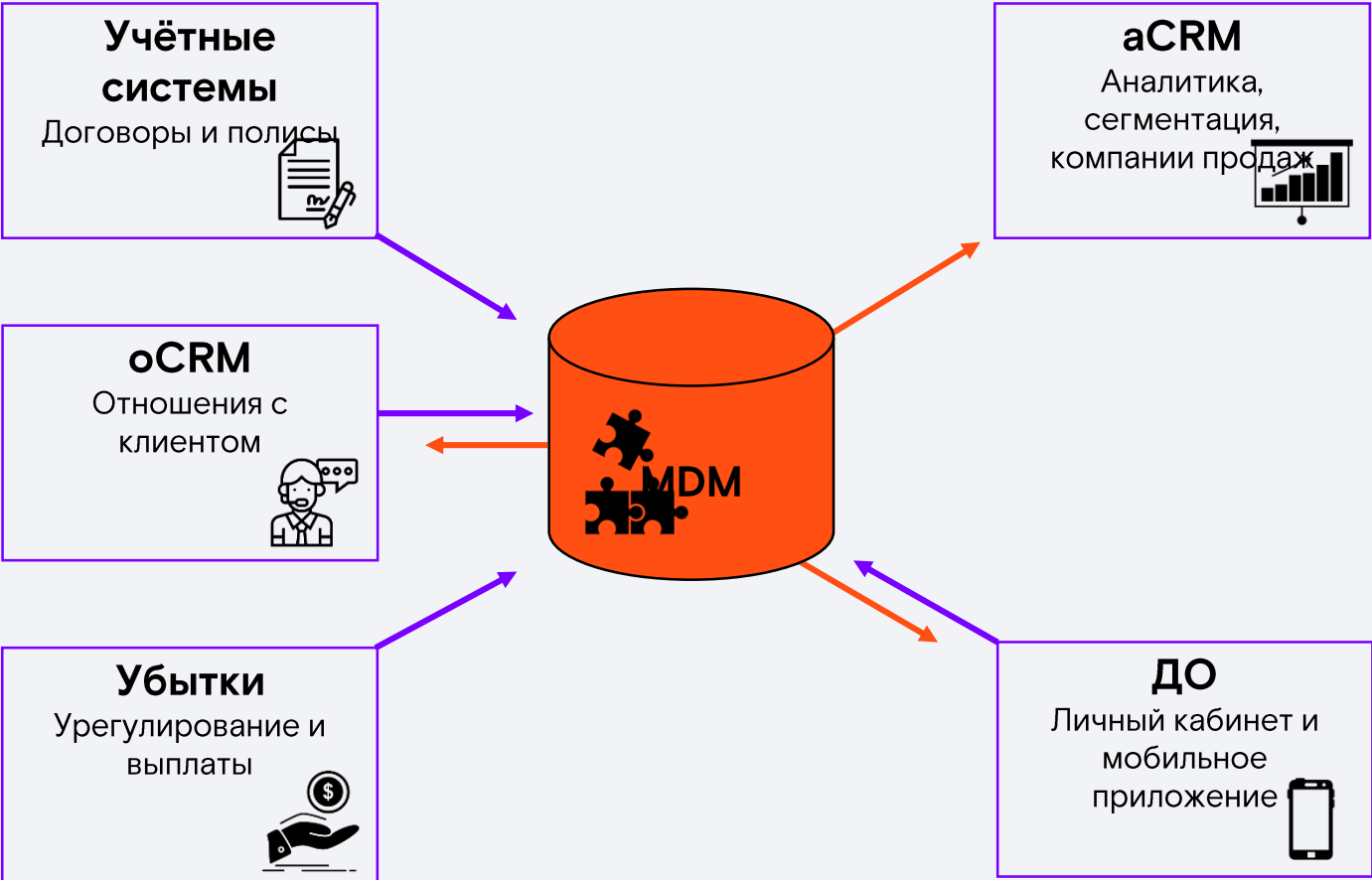
Оператор знает ваши контакты, но не знает актуальных документов.

## МНОГО БИЛЛИНГОВ

Для разных продуктов часто используются разные биллинговые системы. Например, для мобильной и фиксированной связи.

Биллинги для мобильной связи обычно геораспределённые.

# Клиентский MDM в страховой компании



Знаковые проекты



Агент прячет от страховой компании реальные контакты клиентов.

H

## Задание про разметку дублей

# КЛИЕНТСКИЕ ДАННЫЕ — ЭТО ПАЗЛ



Сергей В. Иванов
31.07.1983
Дубна 2-й Театральный
8(916)151-07-23
3-56-03

ИВАНОВ СЕРГЕЙ
1-I-2016
096213-56-03
ivanov_s@mycorp.ru
3-56-03

sregey vlodimirovich
ivanov
Дубна, Татральный 2, 6 17
0079161510723-моб
4957371293доб101



Мечта

Сергей  
Владимирович  
Иванов

31.07.1983

141983, Московская обл.  
г. Дубна  
проезд Театральный 2-й  
д. 6, кв. 17

+7 (49621) 3-56-03

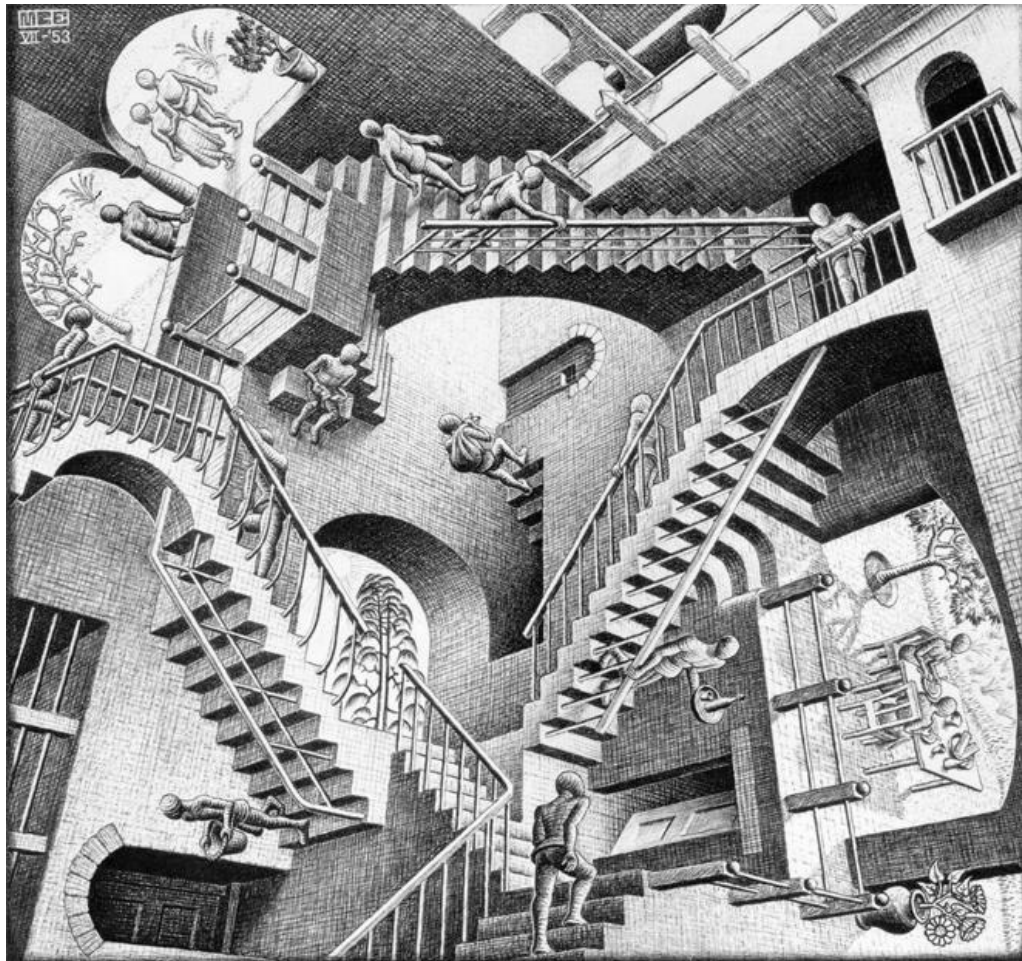
+7 (916) 151-07-23

+7 (495) 737-12-93\*101

ivanov\_s@mycorp.ru



## Путь к мечте

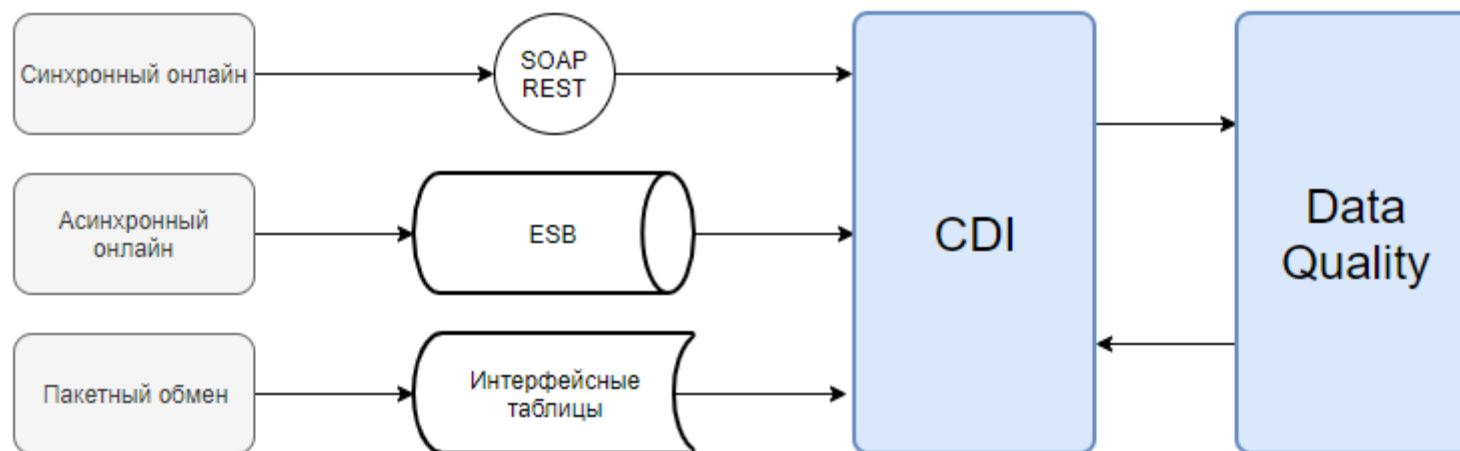


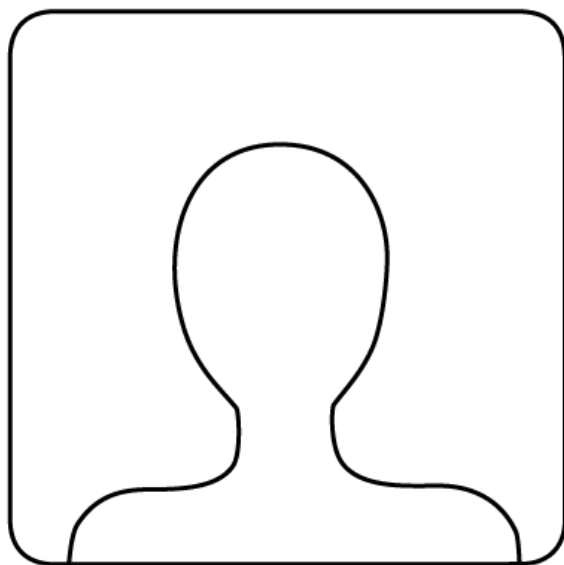
Донести

Стандартизировать

Разметить дубли

Слить дубли

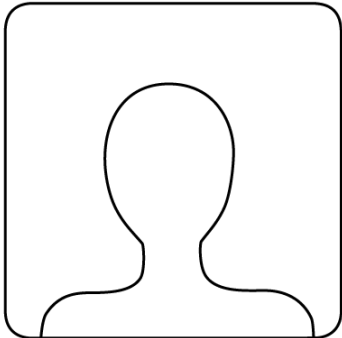




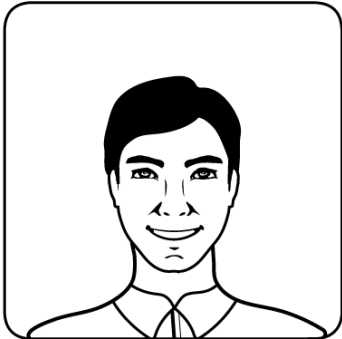
ПРИВЕСТИ В ПОРЯДОК



Приводим



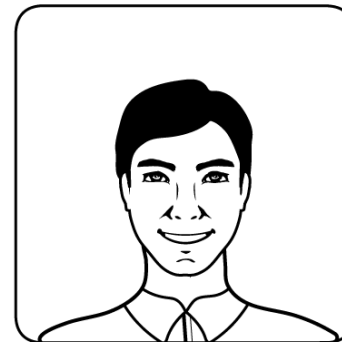
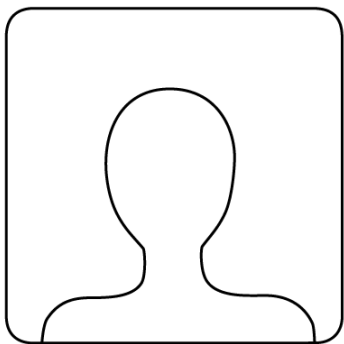
@



sregey vlodimirovich ivanov

фамилия	Иванов
имя	Сергей
отчество	Владимирович
пол	мужской

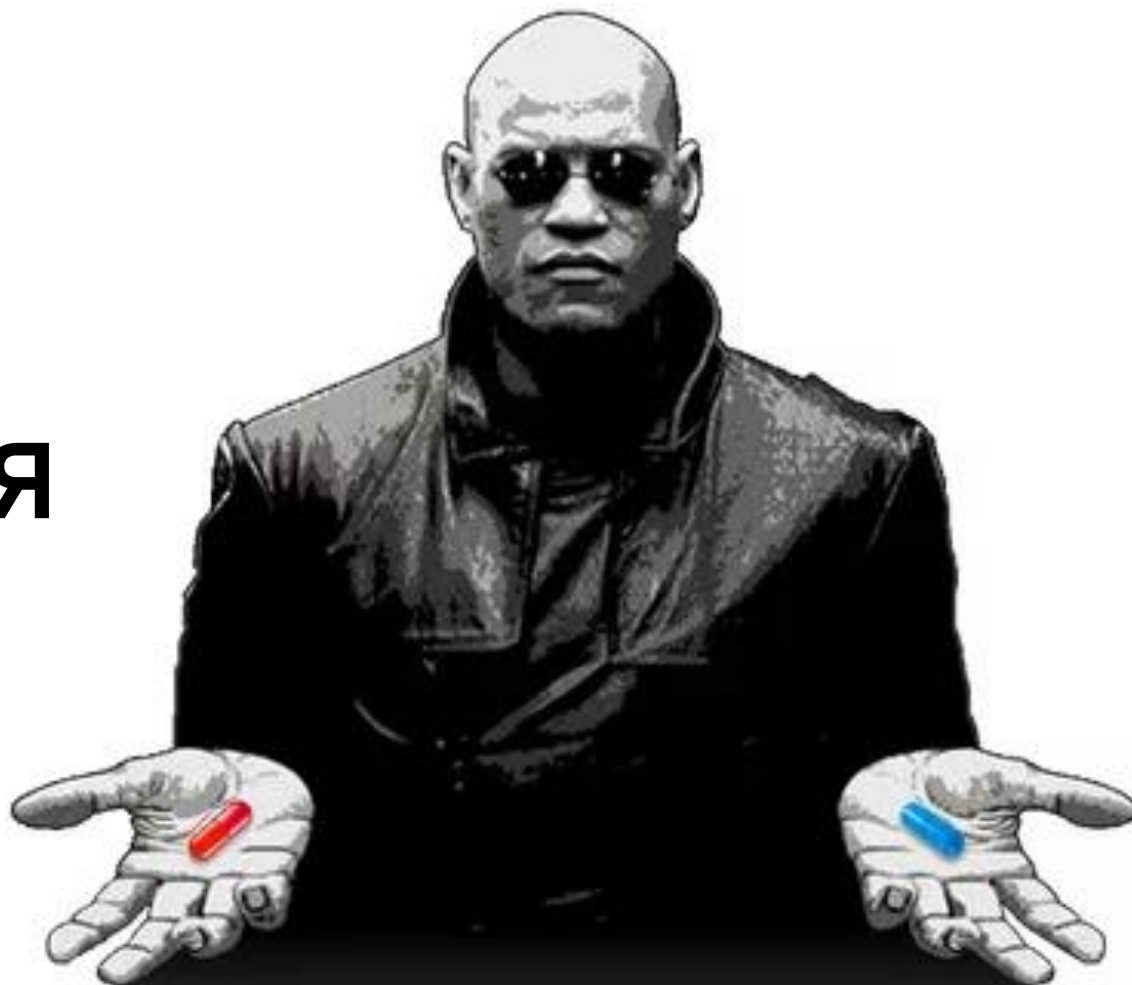
Что если?



Савченко Саша  
Николай Иванович  
КУЗЬМИН ЛА  
Сереза Корнеев  
Марина Марина

Александра?  
Бранислав Иванович  
КУЗЬМИН ЯН  
Юра Мовсисян  
Елена Елена

НАТАЛЬЯ



НАТАЛИЯ



Даты



01.01.1900

01.01.1970

01.01.2018

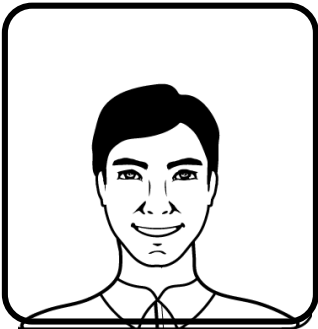
01.01.00

21.03.2019

21.03.2025



Дубна, Татральный 2, 6 17



индекс		
141983		
тип региона	регион	тип города
область	Московская	город
тип улицы	улица	
проезд	Театральный 2-й	
дом	квартира	площадь
6	17	43.4 м²



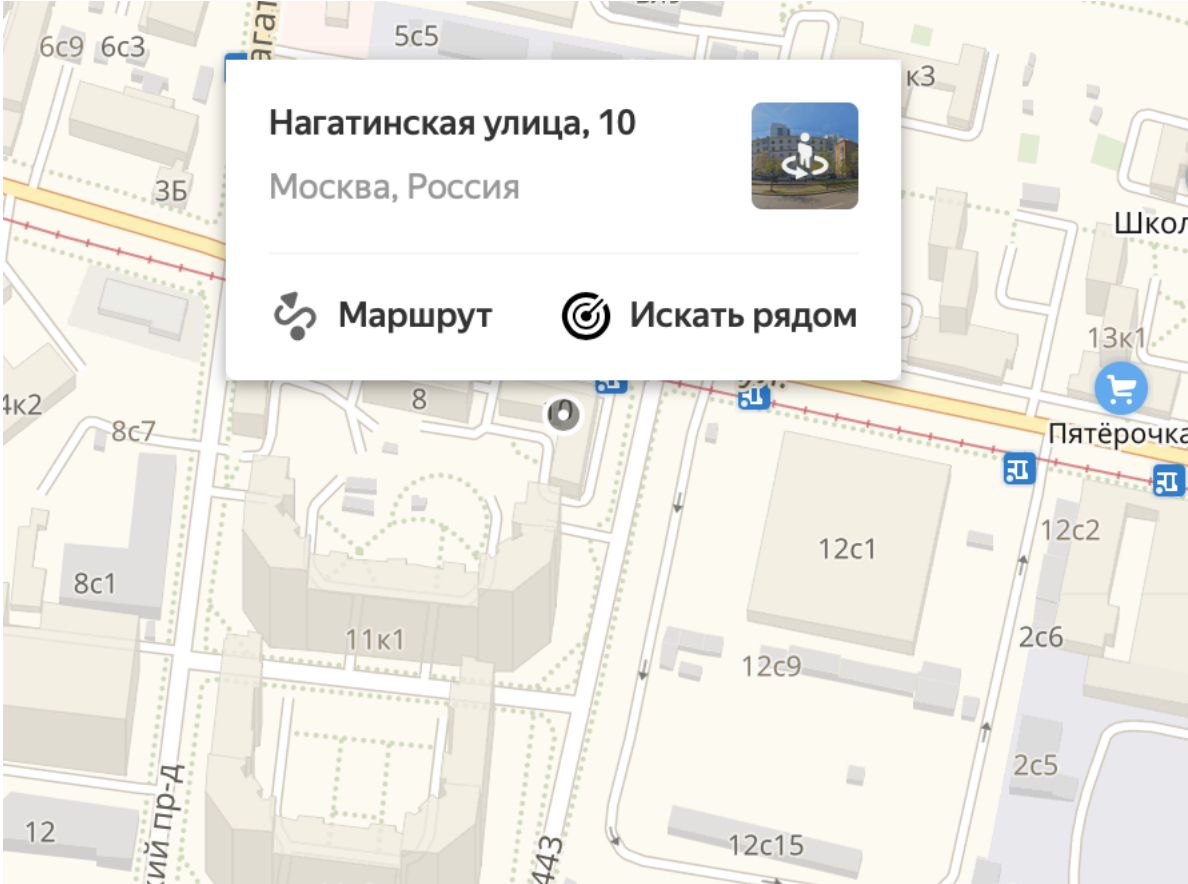
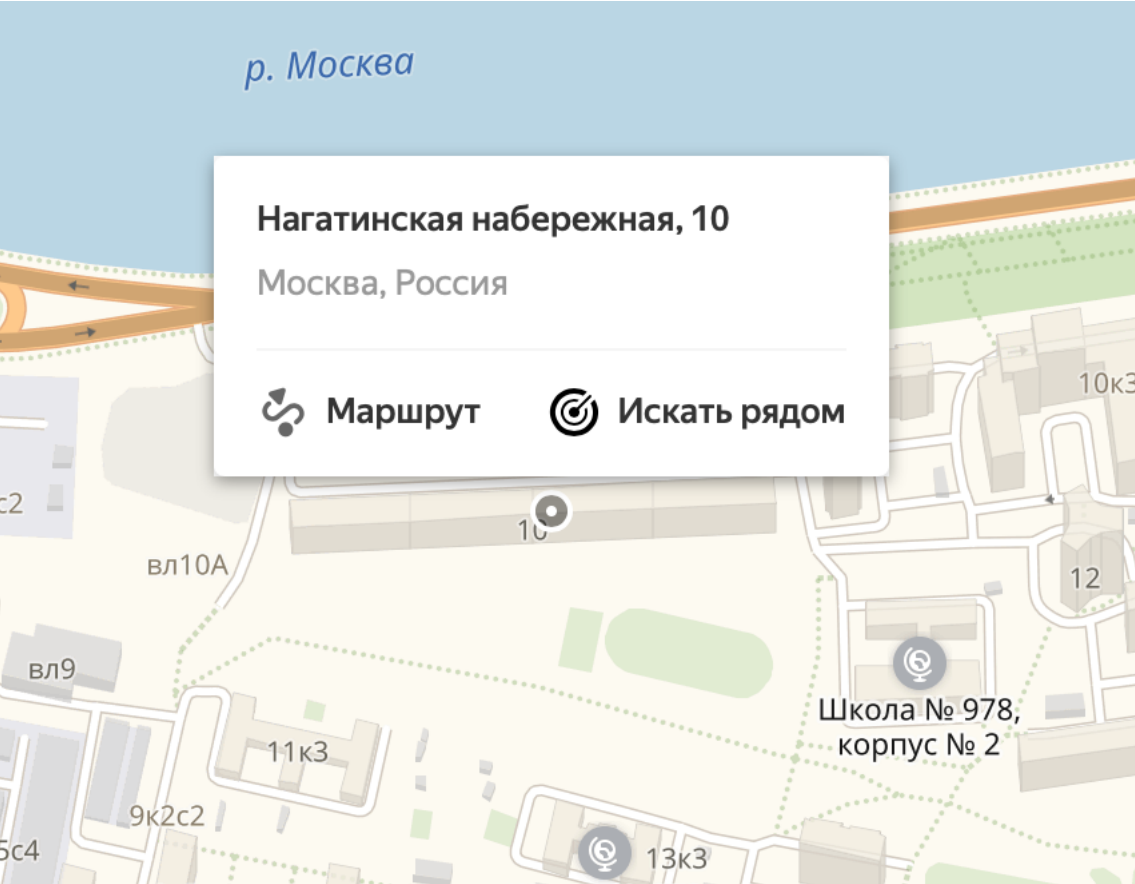
Разберём адрес



Москва Нагатинская 10



# Разберём адрес



И целого Левенштейна мало



Космодемьянская

Краснодомианская

Космодамианская  
набережная

Козьмодомьянская

Космодоминская



# Телефон



8903111111

955 7752550

8(999)999-99-99

+7/095/5325353

20551

# Имейл



n.k.valuev@gmail/com

info@mysite.ru

vasiliy@gmail.ru

gaucomole@10minutesmail.com

1@ya.ru

# Документ



1111 000000



Какой же это клиент?



Номер полиса

VIN автомобиля

Лицевой счет



Данные пригодны



Нужна ручная проверка

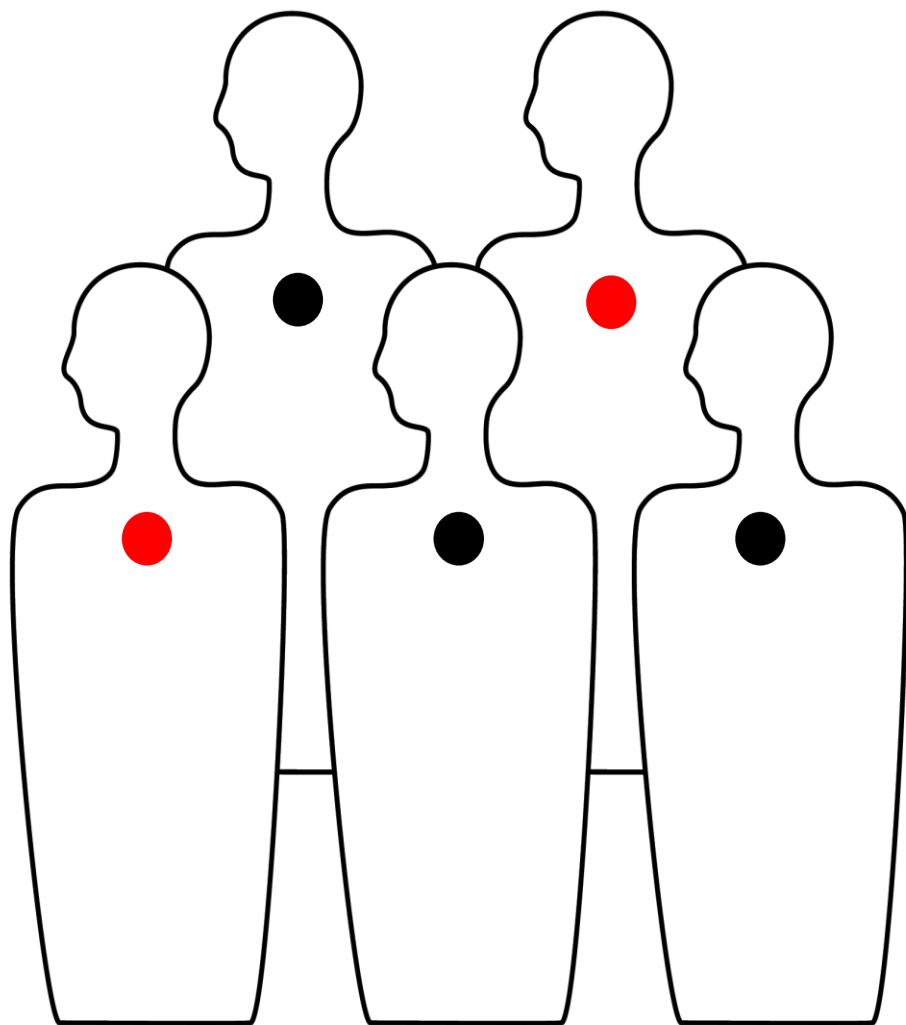


Непригодные данные

И регулярно актуализировать!

# Стандартизированные карточки

Дубли



Разметить



Иван  
Травин

960-42-42



Иван  
Травин

960-42-42



Иван  
Травин

960-42-42

Иван Травин 960-42-42

=

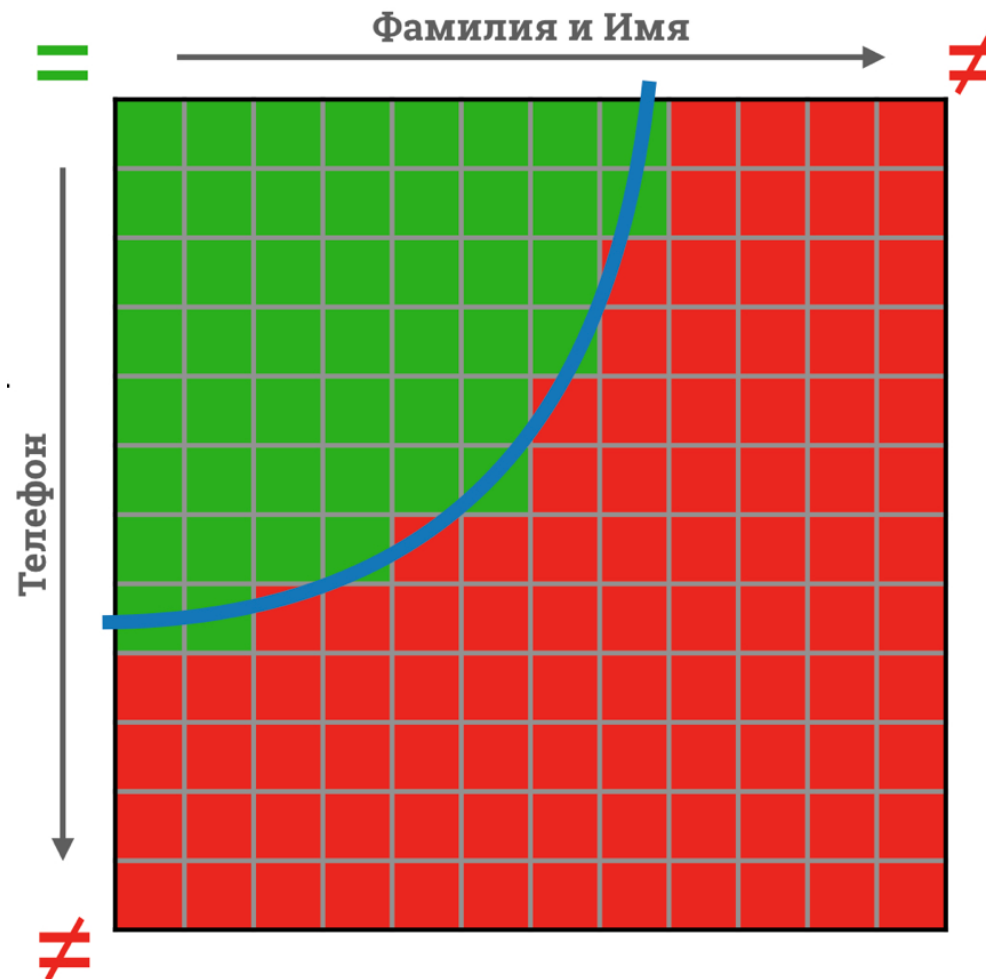
Иван Травин 960-42-42

Иван Травин 960-42-42

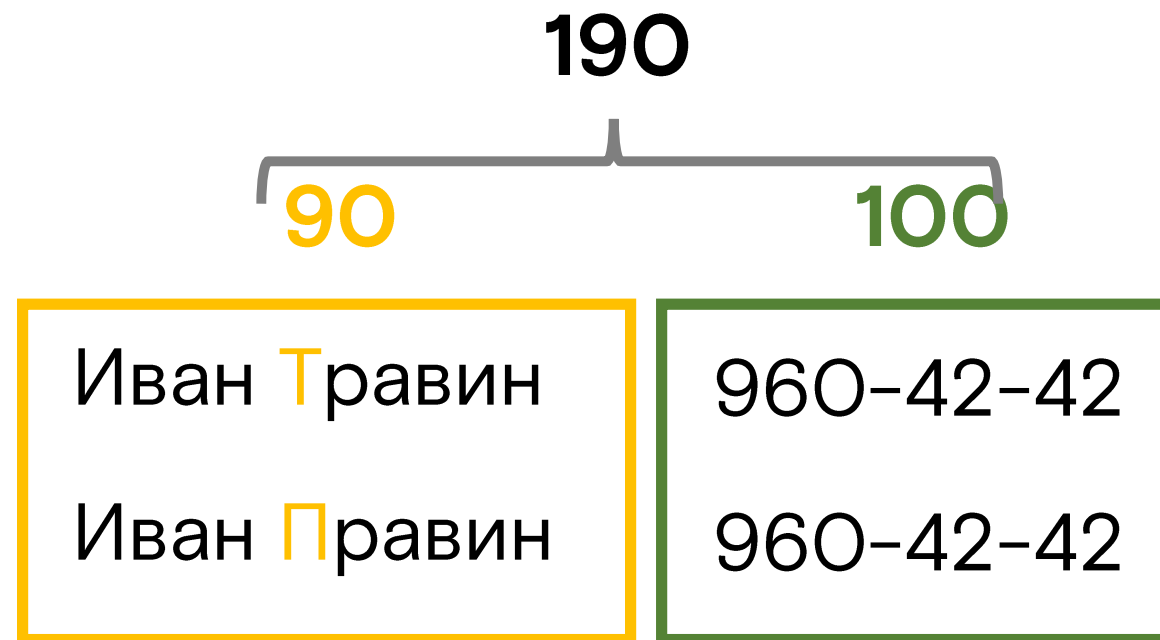
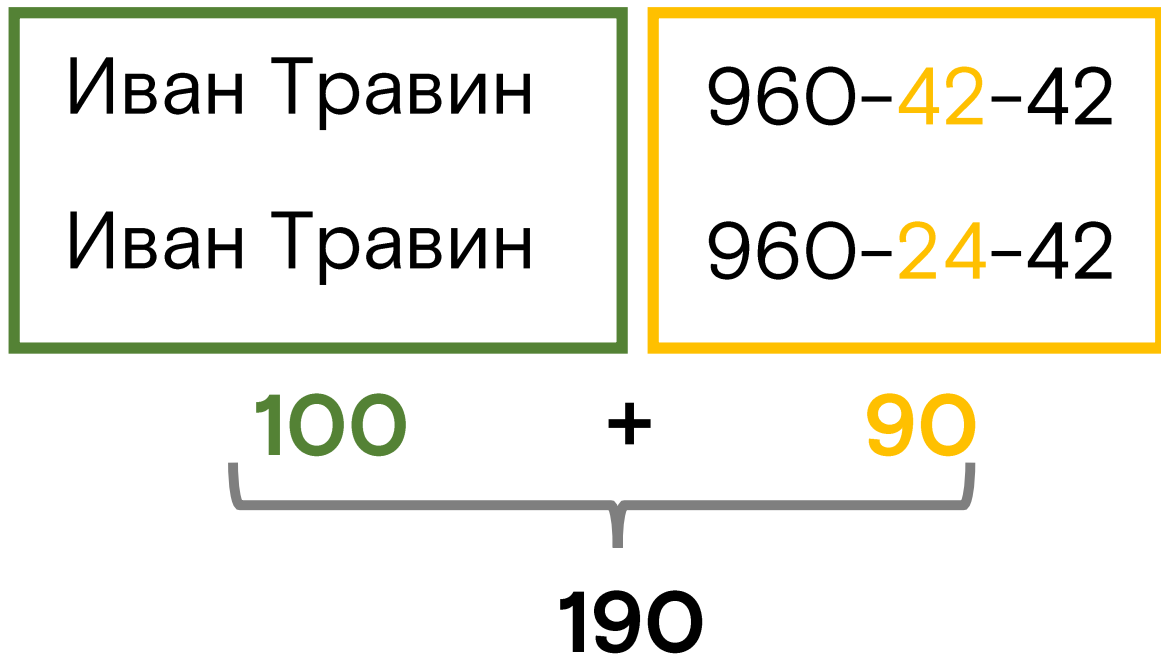
≠

Иван Травин 960-24-42

Скоринговый коэффициент  
«похожести»







90% ФИ + 100% телефон — 180 единиц

100% ФИ + 90% телефон — 190 единиц

Иван Травин 960-42-42

Иван Травин 960-24-42

=

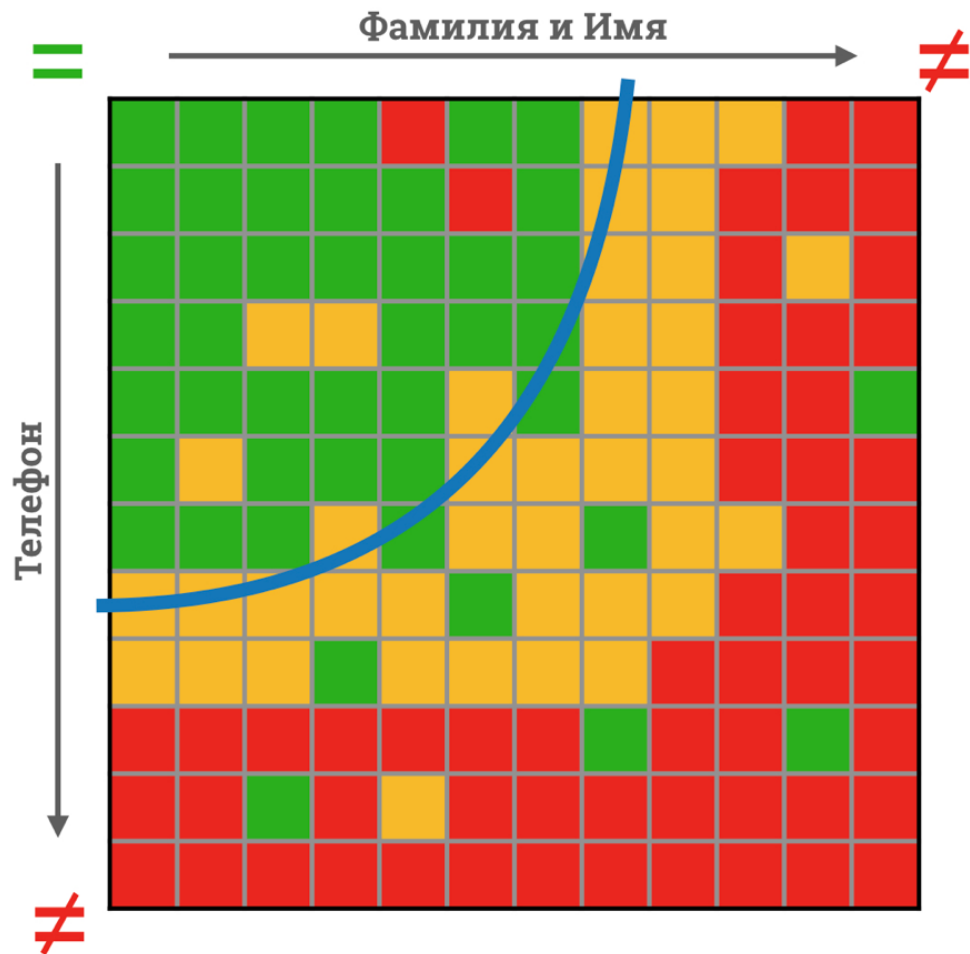
Дубликат

Иван Травин 960-42-42

Иван Правин 960-42-42

=

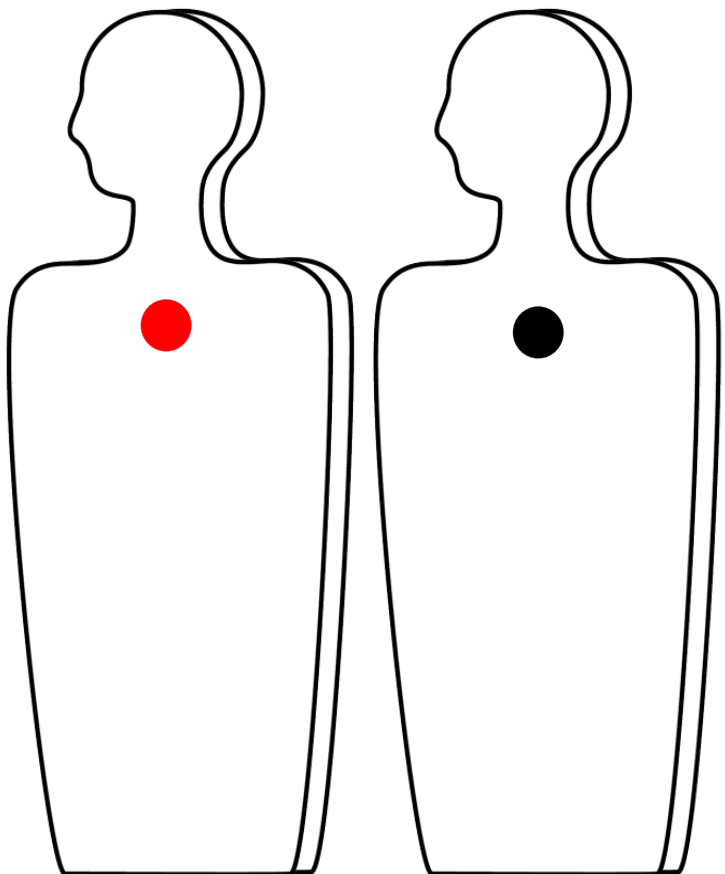
Проверка



Кто совпадает



Дубли



Слить

# Что делать с дубликатами



## Карточка АБС ID=10

## Карточка CRM ID=20

ФИО

Травин Иван

Травин Иван Сергеевич

Телефон

(495) 960-42-42

960-42-42

Дата рождения

31.03.1984

Документ

7701 359743

7701 359743





# Две стратегии эталонизации

Карточка АБС ID=10

~~Карточка CRM ID=20~~

ФИО

Травин Иван Сергеевич

~~Травин Иван Сергеевич~~

Телефон

(495) 960-42-42

~~960-42-42~~

Дата рождения

31.03.1984

~~-~~

Документ

7701 359743

~~7701 359743~~

Карточка АБС ID=10

Карточка CRM ID=20

Золотая (хид + идентификаторы)

ФИО

Травин Иван

Травин Иван Сергеевич

Травин Иван Сергеевич

Телефон

(495) 960-42-42

960-42-42

(495) 960-42-42

Дата рождения

31.03.1984

31.03.1984

Документ

7701 359743

7701 359743

7701 359743

## Задание про слияние

А теперь пообновляемся



Вдруг что-то изменится



Карточка АБС ID=10

~~Карточка CRM ID=20~~

Карточка CRM ID=20

ФИО

Травин Иван

~~Травин Иван Сергеевич~~

Травин Иван Сергеевич

Телефон

(495) 960-42-42

~~960-42-42~~

960-42-42

Дата рождения

31.03.1984

-

Документ

7701 359743

~~7701 359743~~

7712 346424

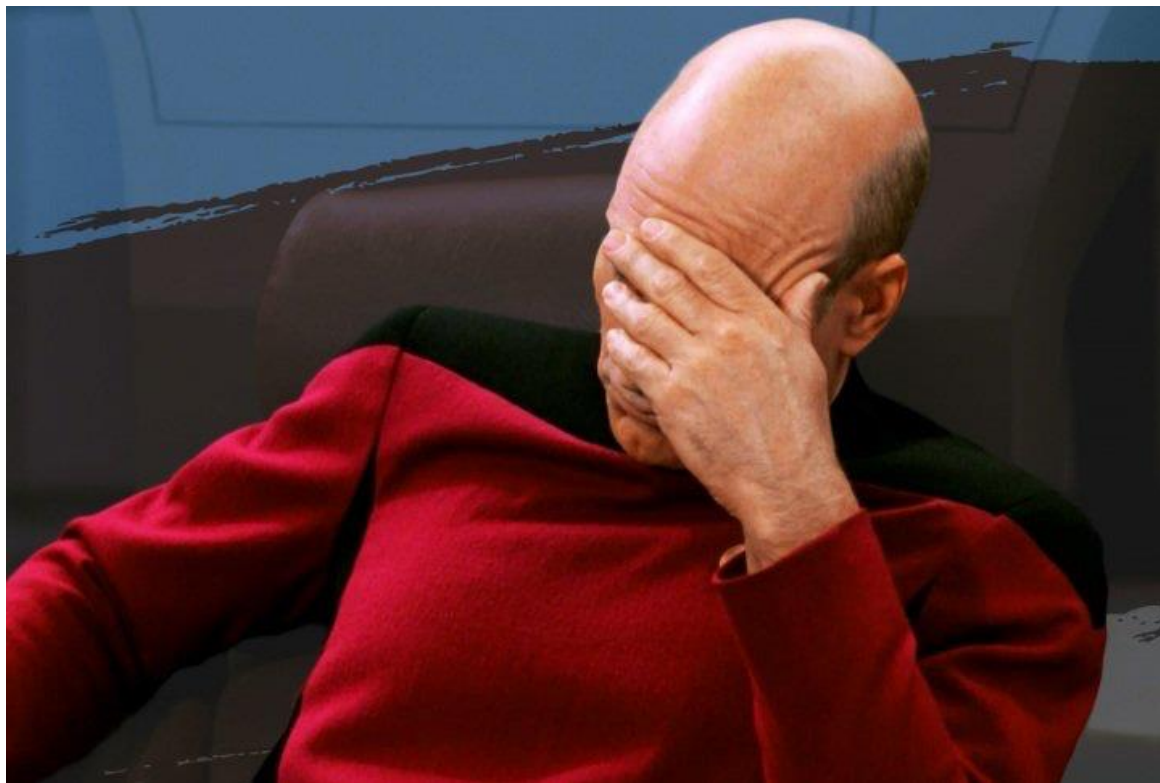


## Плюсы и минусы





Вдруг что-то изменится



При каждом обновлении:

1. Создать техническую карточку.
2. Найти дубли.
3. Найти лучшие данные в дублированных карточках.
4. Объединить лучшие данные в одной карточке.
5. Удалить техническую карточку.



Единая точка ввода (зарплатники, поглощения)

CRM — тот же MDM (синергия с задачами самого CRM)

Постоянный идентификатор (+картинки)

Распространим чистые данные в источники (обратные потоки)

Один раз почистим, а потом заживем новой жизнью (данные живые)

Давайте загрузим в CDI всё (сроки, отсутствие владельца, бесполезность, никогда не спрашивать что хотят — хотят всё)

Статья в блоге HFLabs — <https://tinyurl.com/yyab6d44>



Multienterprise MDM —  
федеративный

152-ФЗ, GDPR

Облака

СМЭВ

Идентификация по чему угодно

Геоданные и цифровой отпечаток

Подписывайтесь, ставьте лайки



# H F Labs

Канал в Telegram — [https://t.me/hflabs\\_official](https://t.me/hflabs_official)

Блог — <https://blog.hflabs.ru/>

Контакты — [nikitan@hflabs.ru](mailto:nikitan@hflabs.ru) и [pavela@hflabs.ru](mailto:pavela@hflabs.ru)

