

H

F

Labs

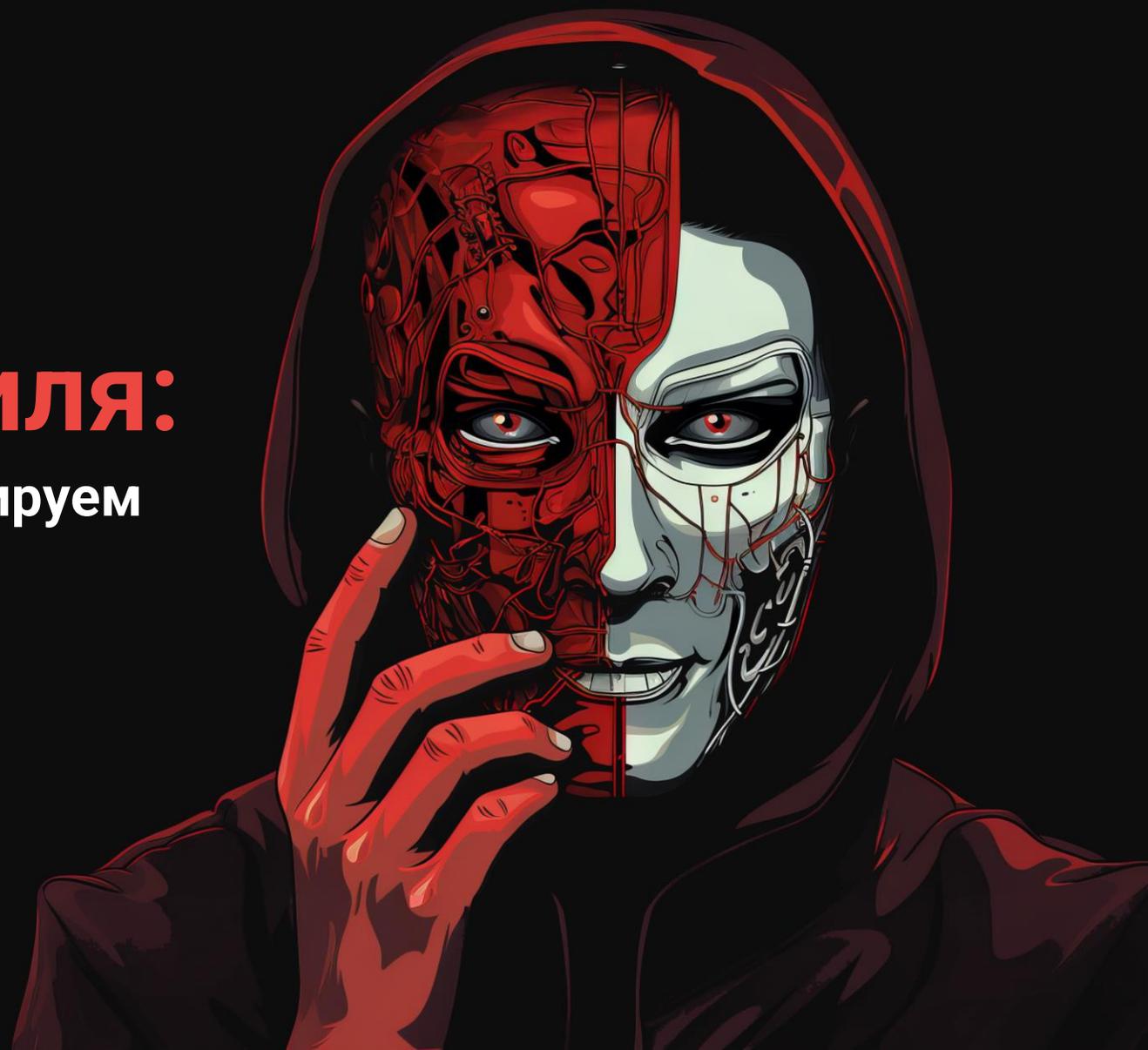
Централизация клиентского профиля:

как мы собираем, защищаем и маскируем
данные и согласия



Никита Назаров

Технический директор



С 2023 г.

чистим и обогащаем клиентские данные

Строим эталонные клиентские
базы с сотнями миллионов записей

1

Приводим данные в порядок

- собираем
- систематизируем
- помогаем монетизировать

2

Помогаем единообразно работать с данными

даем целостный взгляд на все составляющие бизнеса, в том числе на источники данных, авторство, качество, полноту и на потенциальное использование данных

Зачем все это?



ХОРОШАЯ ЕДИНАЯ КЛИЕНТСКАЯ БАЗА

1

Дает бизнесу новые знания о клиенте

2

Работает с минимальным участием дата-стюардов

3

Быстро адаптируется под бизнес-процессы

5

Прощает ошибки и помогает их исправить

4

Не тормозит на десятках миллионов записей

Создание единой эталонной базы клиентов позволяет:

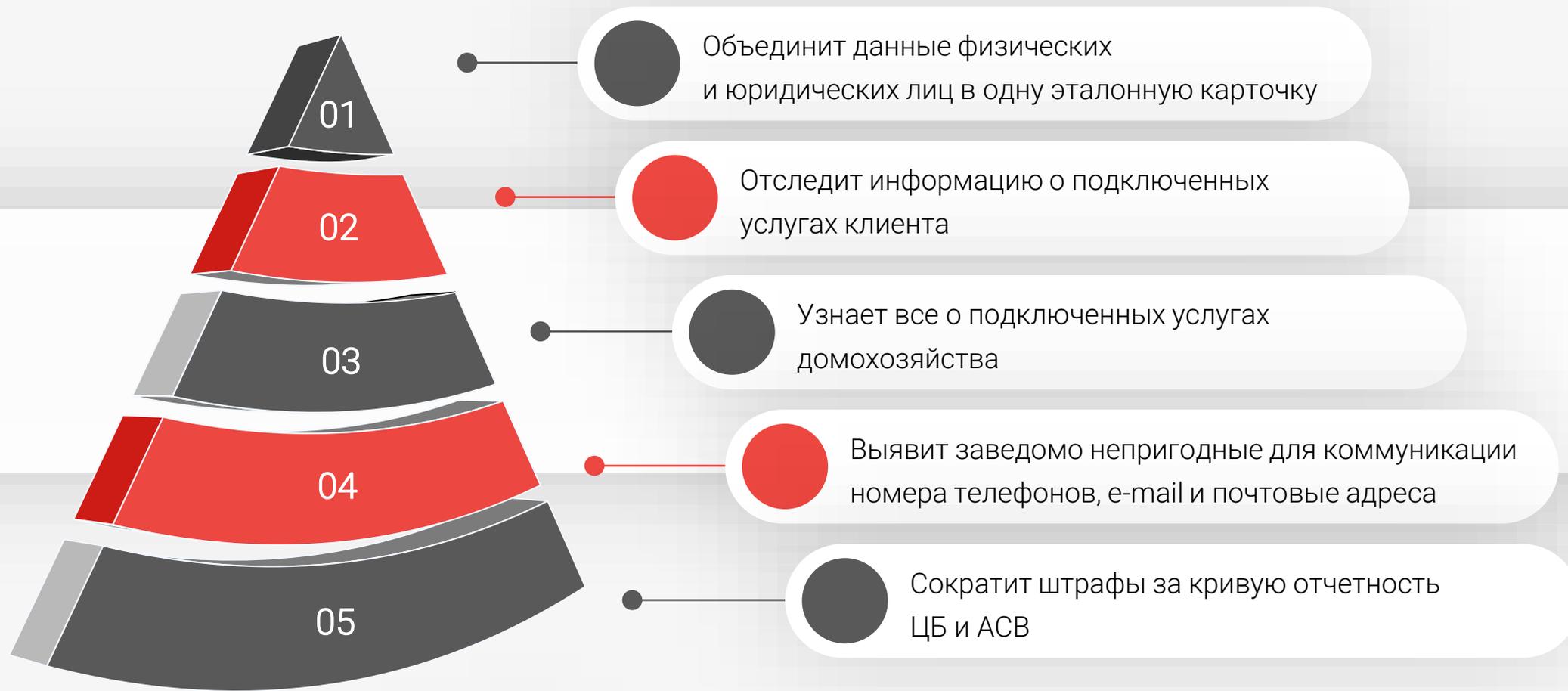
1

Сделать все виды продуктов банка технически доступными всем клиентам

2

Повысить качество информации о клиенте и узнать его предпочтения

Чем «Единый клиент» порадует CIO?



Чем «Единый клиент» порадует CISO?



Обработка ПД возможна ТОЛЬКО с согласия клиентов, **НО**

H F Labs

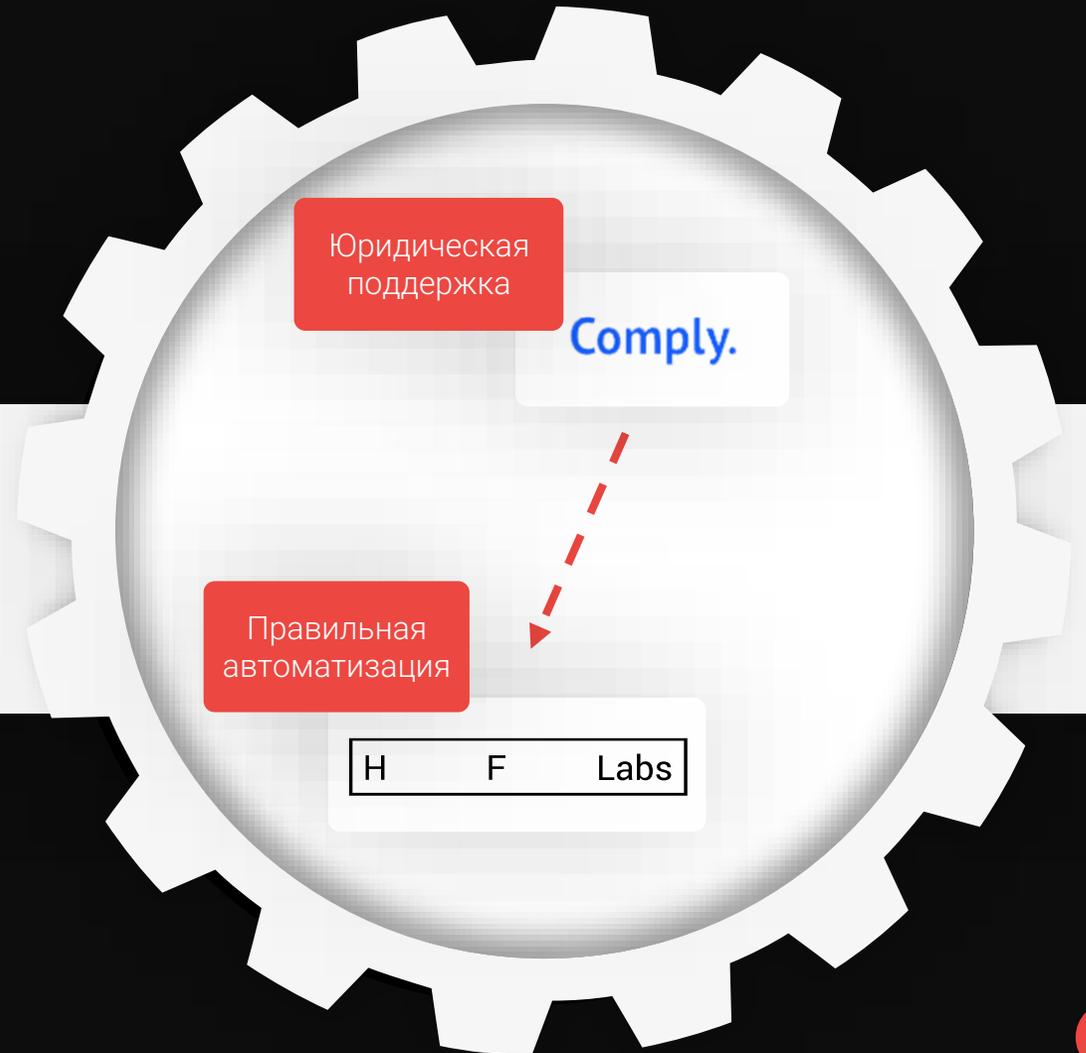
1 Какие согласия собираются и где?

2 Что в них написано?

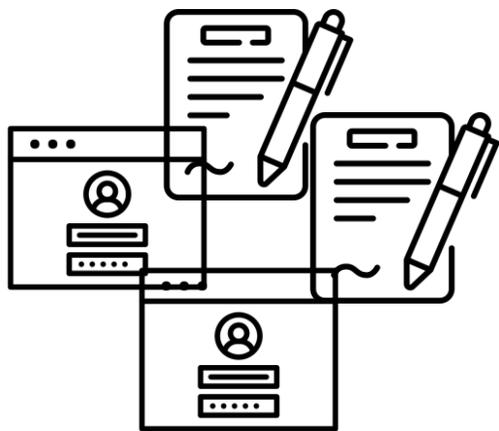
3 Кто за все это отвечает?

4 Кто знает «как надо»?

5 Что делать, если?...



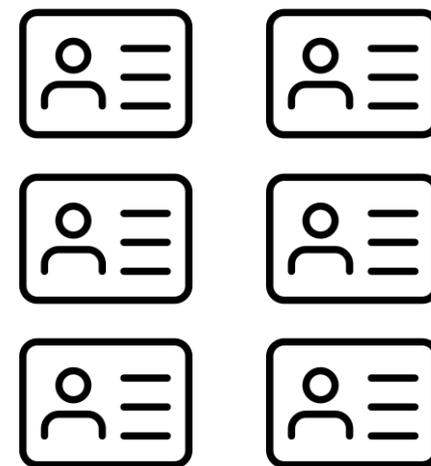
Центр управления согласиями



Десятки разных согласий собираются бизнесом в реальном времени



ЦУС собирает согласия в единую базу



Бизнес получает информацию по согласиям в разрезе своих клиентов

Статистика нарушения ФЗ «О рекламе»

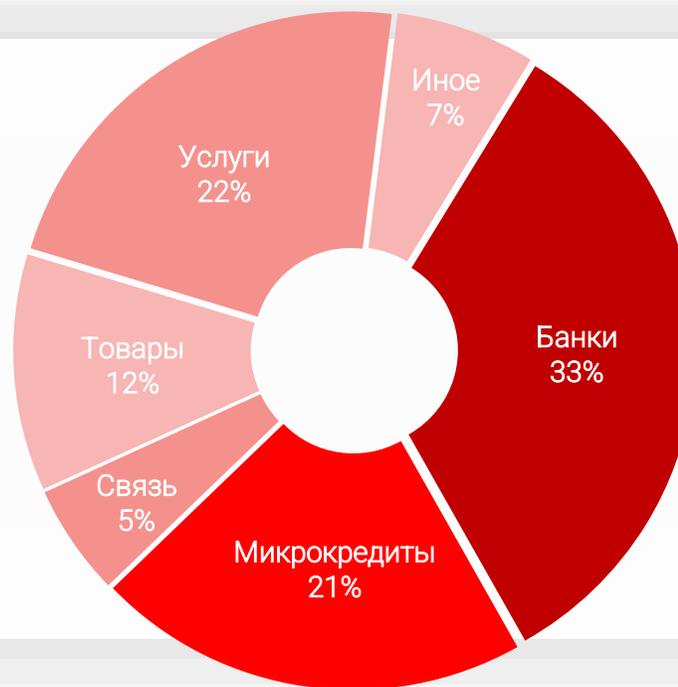
H F Labs

75%

Жалоб в ФАС на рекламу без согласий
привели к штрафам

на **100-150** тыс.
руб

Может быть оштрафована компания
за каждый звонок, смс, e-mail



**Какие отрасли чаще
всего попадают в поле
зрения ФАС**

источник: br.fas.gov.ru

Общение с регуляторами

1

Чем активнее маркетинг, тем больше шанс, что попадется клиент без согласия

2

Бремя доказательства наличия согласия ложится на компанию. Клиенту регулятор «верит на слово»

КАК ПОДАТЬ ЖАЛОБУ НА НЕЗАКОННУЮ СМС-РЕКЛАМУ ИЛИ РЕКЛАМНЫЙ ЗВОНОК

Распространение рекламы по сетям электросвязи, в том числе телефонной, в отсутствие предварительного согласия абонента на получение рекламных сообщений не допускается. Ответственность за данное нарушение несет распространитель рекламы (часть 1 статьи 18 Федерального закона от 13.03.2006 №38-ФЗ «О рекламе»)



До обращения в антимонопольную службу с заявлением о рассылке смс-рекламы или поступлении звонков рекламного характера получателям такой информации рекомендуется следующее:

Читать подробнее:



Центр управления согласиями – это:

1

Централизованное ведение всех форм согласий

2

Хранение ссылок на сканы согласий и всех проверочных кодов по каждому клиенту

3

Отслеживание, кому клиент дал согласие из партнеров

4

Отслеживание, когда срок согласия вышел и нужно перестать обрабатывать данные

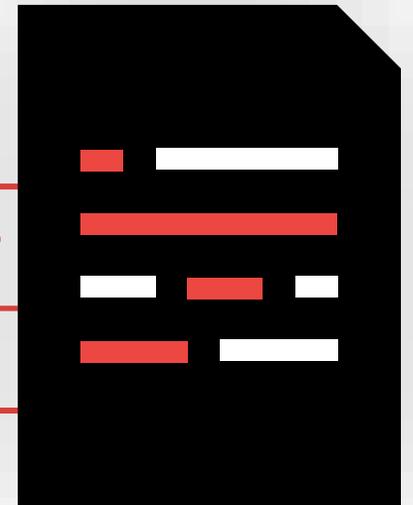
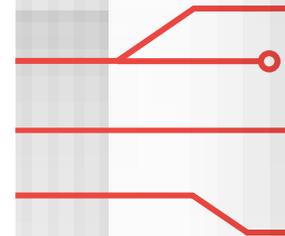
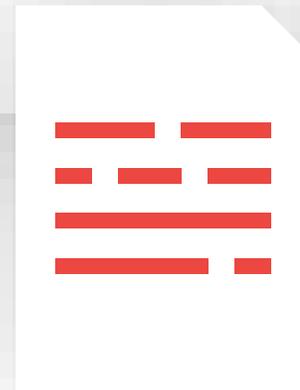
Помогает собрать доказательства по инциденту и успеть в регламентированный 10-дневный срок

Как еще защититить персональные данные?

Маскировать по-умному!



Компании хотят избежать утечек персональных данных клиентов. Поэтому ограничивают доступ сотрудников к персданным



Подход 1

Замена на «звездочки»

H F Labs

Меняет тип данных

Поэтому алгоритм не годится для обезличивания чисел и дат

Слабо защищает популярные имена и адреса

Даже если сократить число «звездочек» до одной, популярные имена, отчества и города легко расшифровывают

Убивает смысл данных

Обезличенные данные теряют семантику, валидность, социально-демографические характеристики и связи по домохозяйствам

Фамилия, Имя, Отчество

Абдюшев Павел
Рашитович



А*****вП***л
Р*****ч

Дата рождения

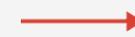
21.01.1979



2*.*11**9

Паспорт

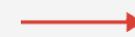
6806 108711



6**61****1

Телефон

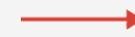
8 926 118 1212



89*****12

Адрес

Москва,
Турчанинов пер. 6,
стр.2



М****а,
Т*****в пер. 6,
стр.*

Подход 2

Замена букв на буквы, цифр на цифры

H F Labs

Поддается расшифровке

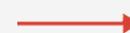
Если алгоритм замены простой, то обезличенные данные можно восстановить

Портит качество и полноту данных

Заменяя случайными буквами и цифрами осмысленные, мы теряем семантику, валидность, сождем характеристики, связи домохозяйствам

Фамилия, Имя, Отчество

Еременко Наталья
Сергеевна



Нонингел Гьпдбч
Мношнагп

Дата рождения

21.07.1961



11.02.1973

Паспорт

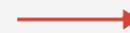
6806 108711



7187 315818

Телефон

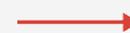
8 926 118 1212



9 037 229 2323

Адрес

Тамбов,
ул. Советская, 11



Тамбов

Как работает «Маскировщик»

**Еременко Петр
Сергеевич**



Сохранит родственные связи.
Если Петра превратил в Сергея,
то Петровичей – в Сергеевичей

21 июля 1960

Ванина 1, Тамбов

6806 108771

8 926 118-12-12

mario@gmail.com

**Антонов Сергей
Андреевич**



Сохранит гендерный
баланс

11 февраля 1961

Сохранит социально-
демографические группы

Дорожная 5, Тамбов

В адресе сохранит страну,
город и район

6807 203771 8 926 311-89-84 elf@mail.ru

Серию паспорта привяжет
к году рождения, чтобы
она прошла форматно-
логический контроль

Оставит оператора
и страну для номера
телефона

Понимает домены емейлов:
рабочие, личные, одноразовые
такими и останутся

Как работает «Маскировщик»

Иванов Петр Ильясович

21.07.1961

Тамбов,
Советская 11

Петров Андрей Фатихович

11.02.1961

Тамбов, Астраханская 5

Иванов Ильяс Петрович

05.03.1988

Тамбов,
Советская 11

Петров Фатих Андреевич

11.08.1987

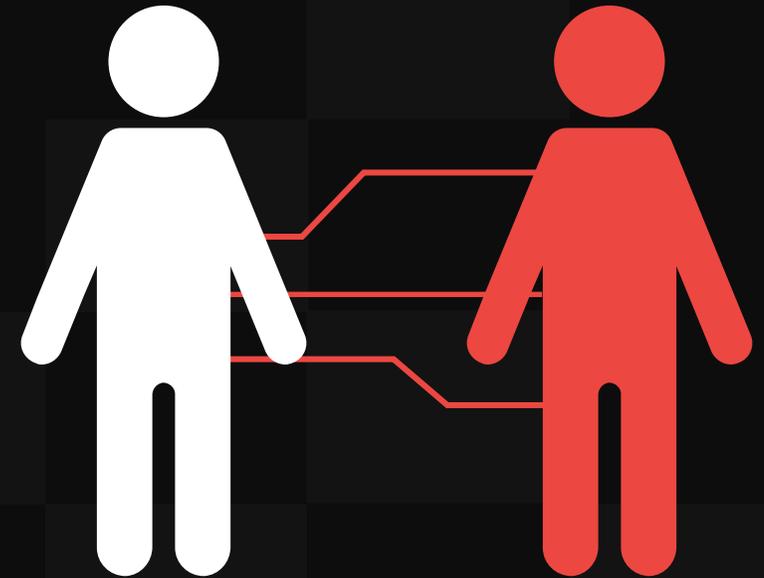
Тамбов, Астраханская 5

«Маскировщик» понимает, что Ивановы Петр Ильясович и Ильяс Петрович, живущие по адресу город Тамбов, улица Советская, 11 — родственники. И учитывает эту информацию, когда маскирует их ФИО и адрес

И что в итоге?

Из базы данных, содержащих записи о миллионах **реальных** людей, «Маскировщик» делает ровно такую же, неотличимую для человека базу данных людей, только **нереальных**

Сохраняет родственные связи, качество данных, соцдем и географическое распределение



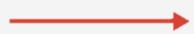
Маскировщик безопасен?

Да, абсолютно!

H F Labs

11 НОЯБРЯ

Петр



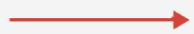
Степан

Ирина



Татьяна

Наталья

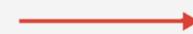


Марина



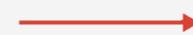
15 ДЕКАБРЯ

Петр



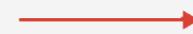
Иван

Ирина



Марина

Наталья



Татьяна

В каждой сессии «Маскировщик» подбирает замены случайным образом.

Данные в файле замен хранятся в виде хеша, набора цифр и букв. Не зная алгоритма, восстановить исходные персональные данные невозможно. По умолчанию кэш сессии удаляется по ее завершению

Какие типы данных понимает «Маскировщик»?

Фамилии, имена
и отчества

01.

Даты
рождения

02.

Документы
физлиц

03.

Адреса

04.

Номера
телефонов

05.

Email
адреса

06.

ИНН

07.

СНИЛС

08.

Номера банковских
карт

09.

Номера
банковских
счетов

10.

Отличает типы данных и понимает их структуру. Качество данных не портится

HFLabs — полностью российская компания

2005 — год основания

1

Среди
собственников
нет иностранных
граждан

3

Имеем
аккредитацию
Минцифры

2

В структуре
доходов
нет поступлений
из-за рубежа

4

Продукты входят
в Реестр
Отечественного
ПО и совместимы
с АльтЛинукс

39%

сотрудников работают
дольше трех лет

130

человек в штате

10

собственных
IT-решений

Наши клиенты:



РОСГОССТРАХ



ГОСУСЛУГИ

OZON



МЕГАФОН

росбанк



открытие | Банк



X5Group



Yota

лента

Ренессанс.
СТРАХОВАНИЕ



АЛЬФА
СТРАХОВАНИЕ

ВТБ

М.ВидеоЭльдорадо

H F Labs

Спасибо
за внимание!

Telegram



Сайт



Хабр

