

H F Labs

«Маскировщик»

Общее описание продукта

«Маскировщик» – программный продукт для быстрого автоматического обезличивания персональных данных.

Для обезличивания используются умные алгоритмы, которые распознают специфику данных и изменяют их, сохраняя базовые характеристики и качество данных. Восстановить конкретного человека – невозможно. Данные остаются консистентными и полезными для тестирования.

Такое решение позволяет повысить качество тестирования и упростить работу с внешними вендорами, ускорить и удешевить вывод новых решений в продакшен.

Программа заменяет персональные клиентские данные на похожие. При этом «Маскировщик» сохраняет их смысл: семантику, валидность, социально-демографические характеристики и связи по домохозяйствам.

Благодаря справочникам «Маскировщик» сохранит недействительный документ недействительным, популярное имя – популярным. Редкое ФИО заменит на редкое, подберет похожий адрес и телефон.

Данные в файле замен хранятся в виде хеша, набора цифр и букв. Не зная алгоритма, восстановить исходные персональные данные невозможно. «Маскировщик» умеет сохранять консистентность при обезличивании нескольких стэндов.

Таким образом персональные данные хранятся в безопасности, а тестирование, аналитики или подрядчики получают данные, похожие на настоящие. «Маскировщик» помогает в интеграциях, тестировании, аналитике и построении моделей машинного обучения.

«Маскировщик» – превращает данные реальных людей в ваших данных в данные нереальных людей.

	Исходные данные	Любой другой продукт	«Маскировщик» HFLabs	Комментарий
ФИО	Иванов Николай Петрович	Авфы Пвфуке Авфпм	Семенов Петр Иванович	ФИО преобразуется в настоящее ФИО, для одинаковых ФИО преобразование одинаково в рамках одной сессии
Дата рождения	12.01.1975	10.04.2007	05.06.1978	Даты рождения меняются случайно, но сохраняется общее распределение. Люди старше 18 останутся такими.
Телефон	+79265764264	+74536642434	+79263253254	Сохраняем регион и оператора
Емейл	mail@mail.ru	fdaeg@fease.ge	graw@gmail.com	Тип домена сохранится, корпоративные, одноразовые и публичные домены останутся такими же
Адреса	Москва, Тверская, 10	Южно-Сахалинск, Ленина, 10	Москва, Ленинградский, 25	У адреса сохранится регион или город (точность можно выбрать), люди массового не сменят регион
Паспорт	7703 425465	5412 535143	7705 235212	Паспорт останется из того же региона, а также с консистентной датой и местом выдачи

Маскирует красиво

Данные надежно замаскированы, но при этом ими можно будет продолжать пользоваться:

- для маркетинга и аналитики сохранится соцдем распределение,
- для анализа рисков сохраняются семейные взаимосвязи,
- тестирование будет работать с неотличимыми от реальных данными, а не с набором букв и цифр.

Маскирует быстро

«Маскировщик» читает данные в несколько потоков пачками и обрабатывает их параллельно. За счет этого эффективно используются все доступные ресурсы сервера.

Маскирует надежно.

Консистентность данных при маскировании сохраняется в рамках одной сессии. Метаданные сессии никому недоступны, их можно удалить сразу после завершения сессии.