

Дмитрий Маслов

CPO, Guardora

Олег Фатюхин

TL, Guardora

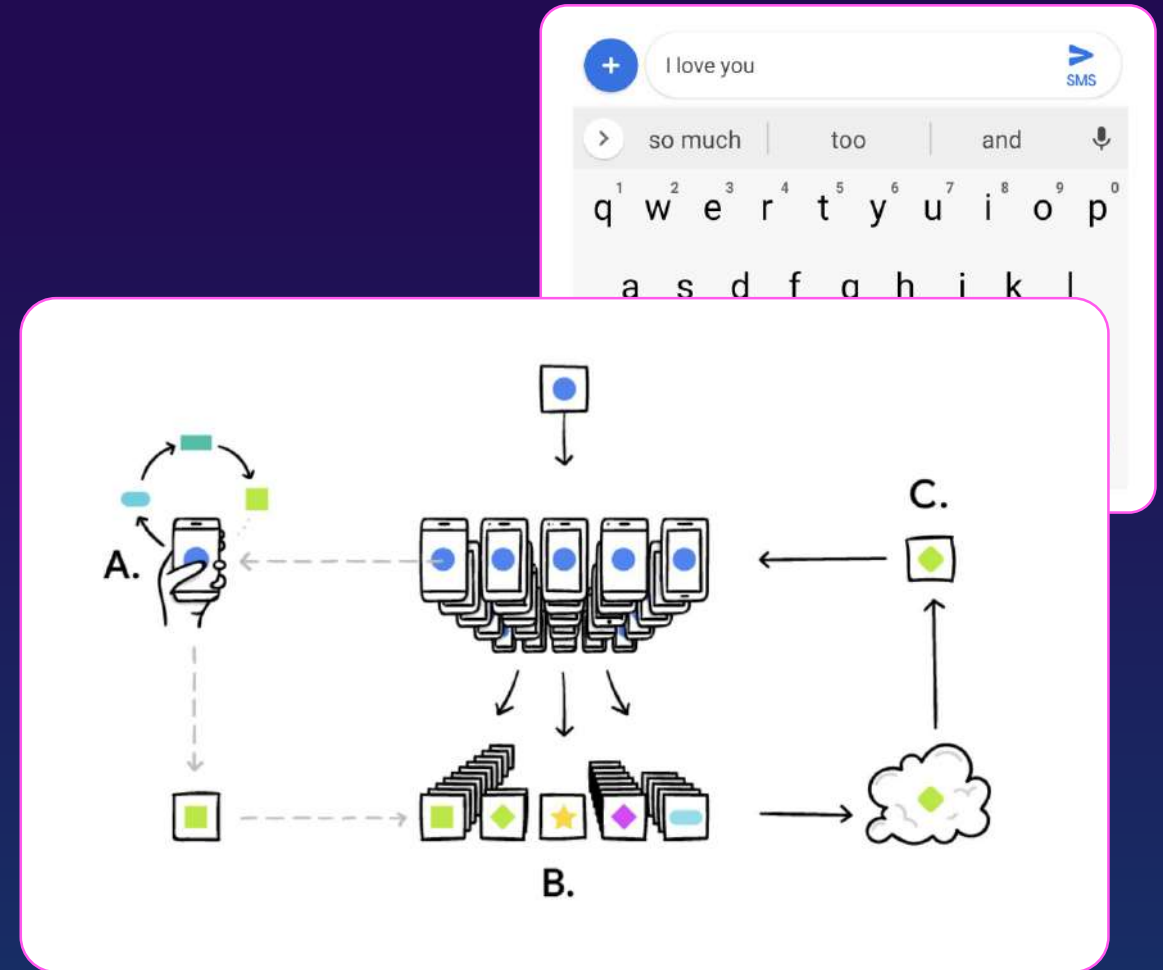
Обеспечение конфиденциальности
данных при разработке ML-моделей с
использованием *технологии
федеративного обучения и
криптографических методов*



FedML в GBoard



- A. Мобильные устройства вычисляют обновления модели на локальных данных
- B. Сервер объединяет обновления и строит глобальную модель
- C. Новая модель отправляется клиентам, и процесс повторяется

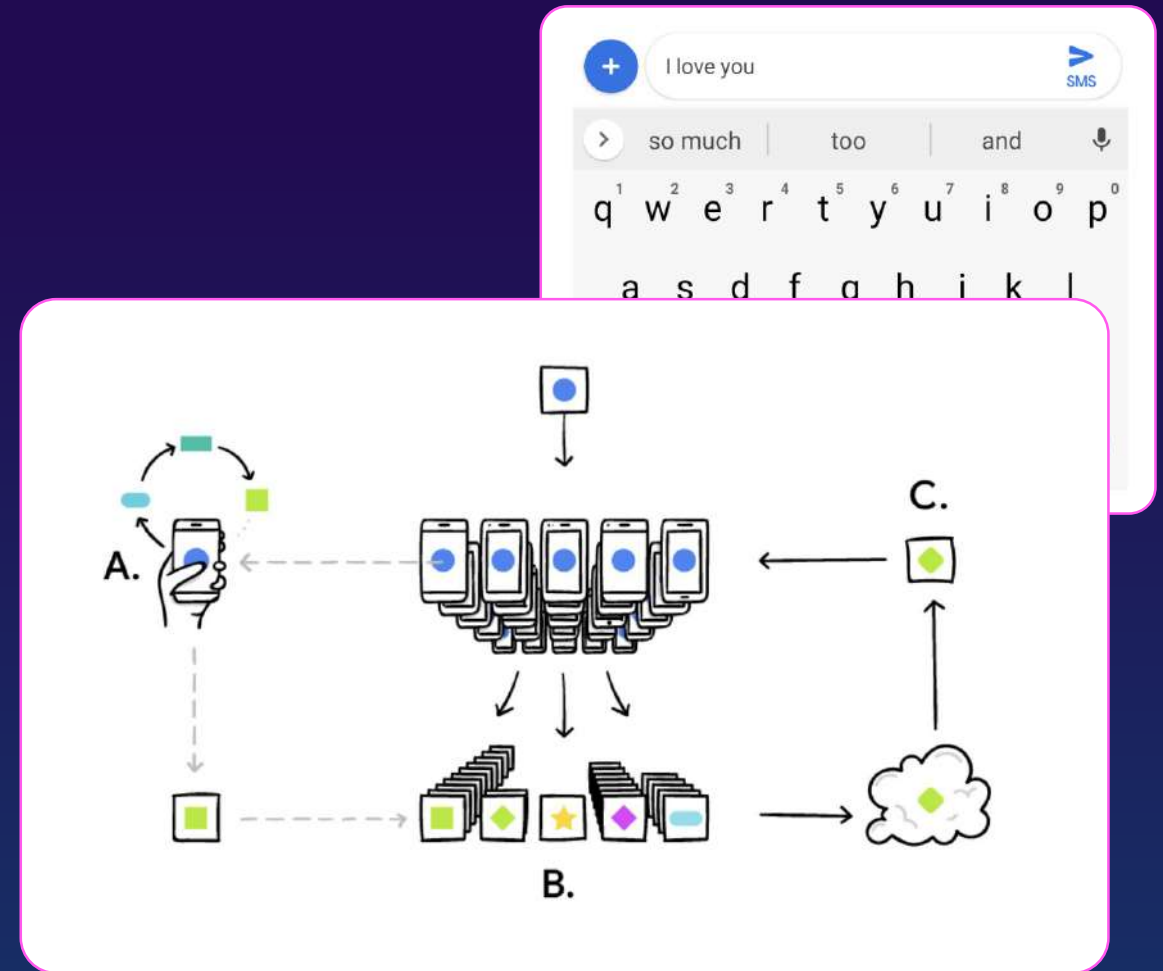


FedML в GBoard

- A. Мобильные устройства вычисляют обновления модели на локальных данных
- B. Сервер объединяет обновления и строит глобальную модель
- C. Новая модель отправляется клиентам, и процесс повторяется

Google таким образом...

- > Улучшает точность модели предсказания вводимого текста
- > Не передает конфиденциальную информацию на сервер
- > Перераспределяет вычислительную нагрузку с серверов на устройства



01



О технологиях FedML и HE

Технологии конфиденциальных вычислений (Confidential Computing)

Машинное обучение с защитой приватности данных (Privacy Preserving Machine Learning)

Многосторонние
безопасные вычисления
(Secure Multi Party
Computation)

Дифференциальная
приватность
(Differential Privacy)

Синтетические данные
(Synthetic Data)

Функциональное
шифрование
(Functional Encryption)

Федеративное обучение
(Federated Machine
Learning - FedML)

Гомоморфное
шифрование
(Homomorphic Encryption)

Технологии конфиденциальных вычислений (Confidential Computing)

Машинное обучение с защитой приватности данных (Privacy Preserving Machine Learning)

Многосторонние
безопасные вычисления
(Secure Multi Party
Computation)

Дифференциальная
приватность
(Differential Privacy)

Синтетические данные
(Synthetic Data)

Функциональное
шифрование
(Functional Encryption)

Федеративное обучение
(Federated Machine
Learning - FedML)

Гомоморфное
шифрование
(Homomorphic Encryption)

Как работает FedML

1

Управляющий сервер генерирует начальные значения параметров глобальной модели



Как работает FedML

1

Управляющий сервер генерирует начальные значения параметров глобальной модели



2

Стороны инициализируют локальные подмодели и проводят раунд обучения



Как работает FedML

1

Управляющий сервер генерирует начальные значения параметров глобальной модели



2

Стороны инициализируют локальные подмодели и проводят раунд обучения



3

Стороны отправляют свои параметры на управляющий сервер



Как работает FedML

1

Управляющий сервер генерирует начальные значения параметров глобальной модели



2

Стороны инициализируют локальные подмодели и проводят раунд обучения



3

Стороны отправляют свои параметры на управляющий сервер



4

На сервере выполняется агрегирование локальных параметров



Как работает FedML

1

Управляющий сервер генерирует начальные значения параметров глобальной модели



2

Стороны инициализируют локальные подмодели и проводят раунд обучения



3

Стороны отправляют свои параметры на управляющий сервер



4

На сервере выполняется агрегирование локальных параметров



5

Формируются новые параметры глобальной модели и направляются сторонам



Как работает FedML

1

Управляющий сервер генерирует начальные значения параметров глобальной модели



2

Стороны инициализируют локальные подмодели и проводят раунд обучения



3

Стороны отправляют свои параметры на управляющий сервер



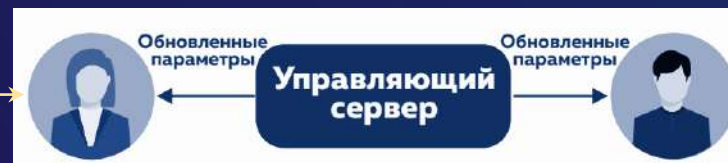
4

На сервере выполняется агрегирование локальных параметров



5

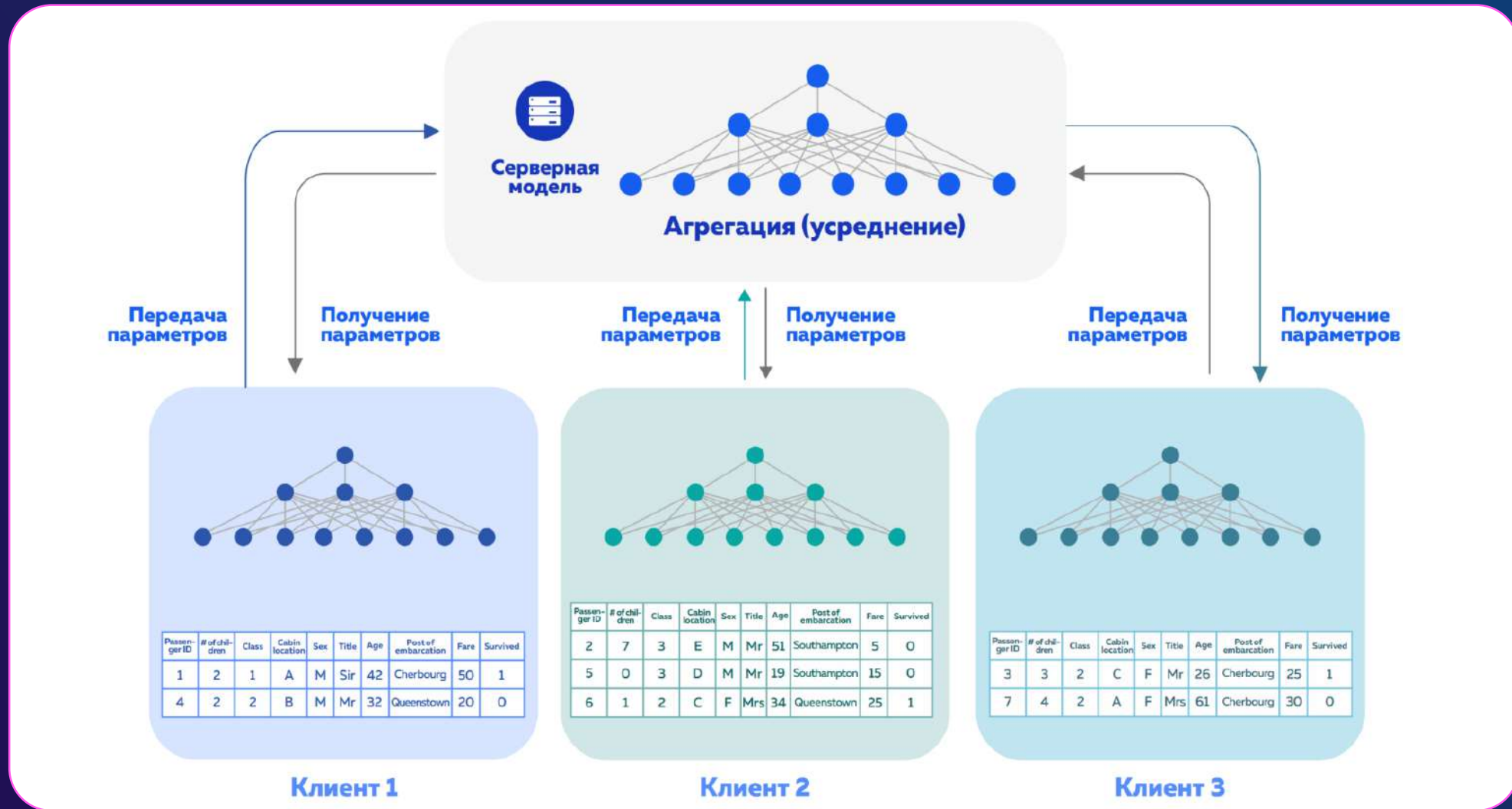
Формируются новые параметры глобальной модели и направляются сторонам



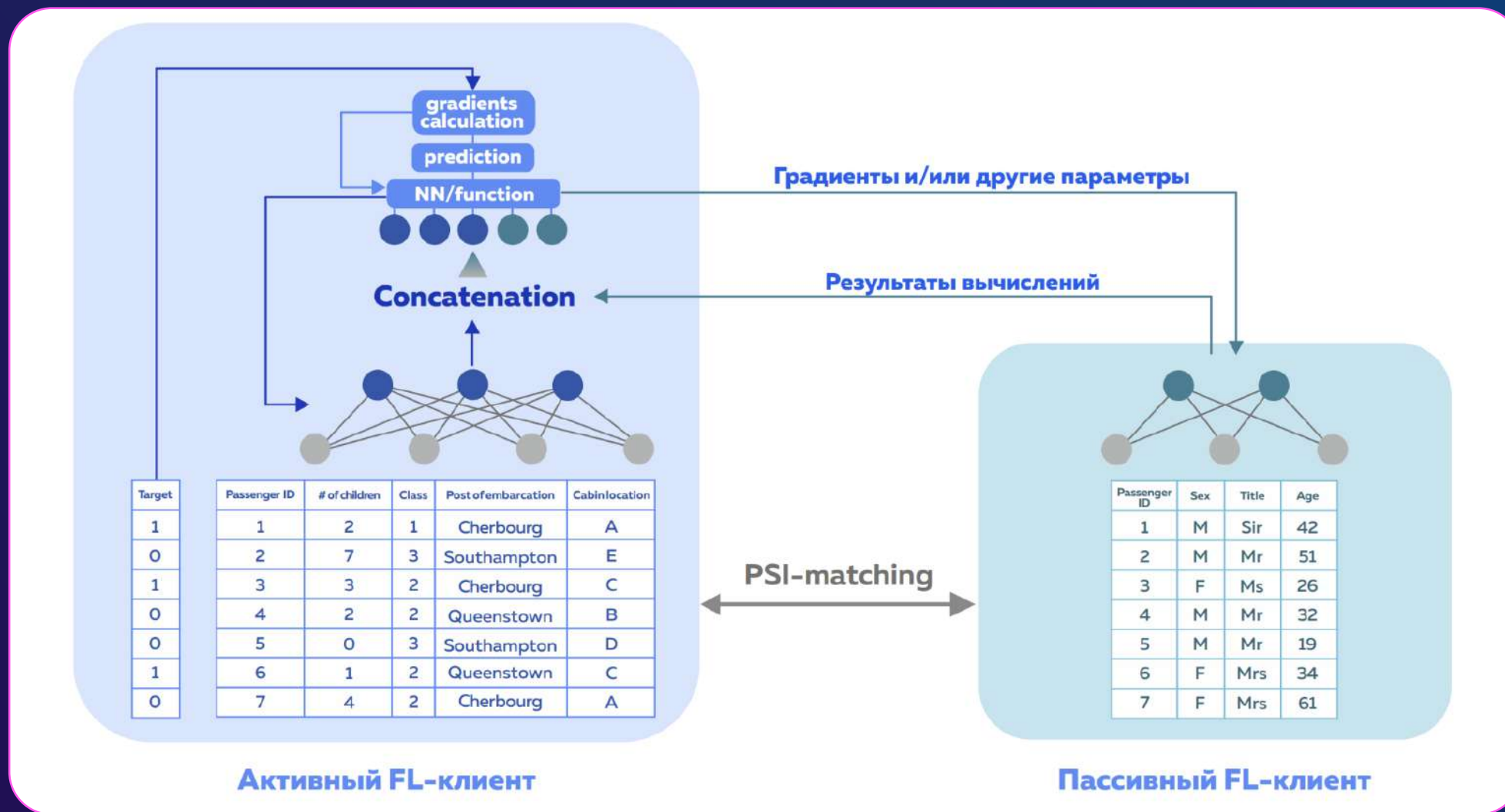
6

Действия шагов 2–5 выполняются, пока не будет выполнен критерий остановки процедуры обучения

Горизонтальный FedML (HFL)



Вертикальный FedML (VFL)

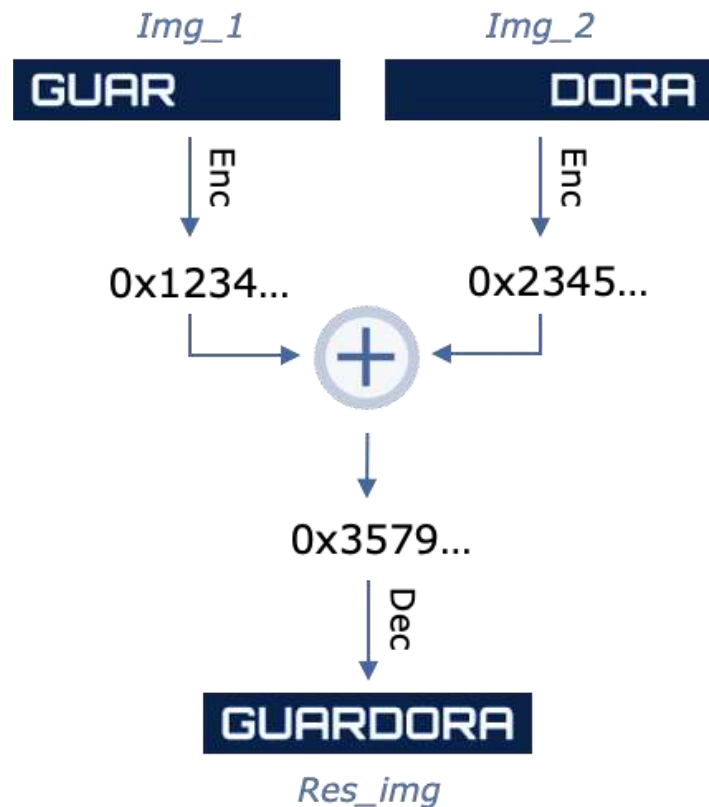


HFL vs VFL

	HFL	VFL
Разделение данных	Пространство сущностей	Пространство признаков
Сценарий обучения	Cross-device, cross-silo	Cross-silo
Циркулирующие данные	Параметры модели	Промежуточные результаты вычислений
Тип локальной информации	Данные	Модели клиентов, данные
Сложность адаптации алгоритма обучения	Средняя	Возможно
Участник получает в итоге	Глобальную модель	Локальную часть модели
Самостоятельное использование модели	Возможно	Невозможно

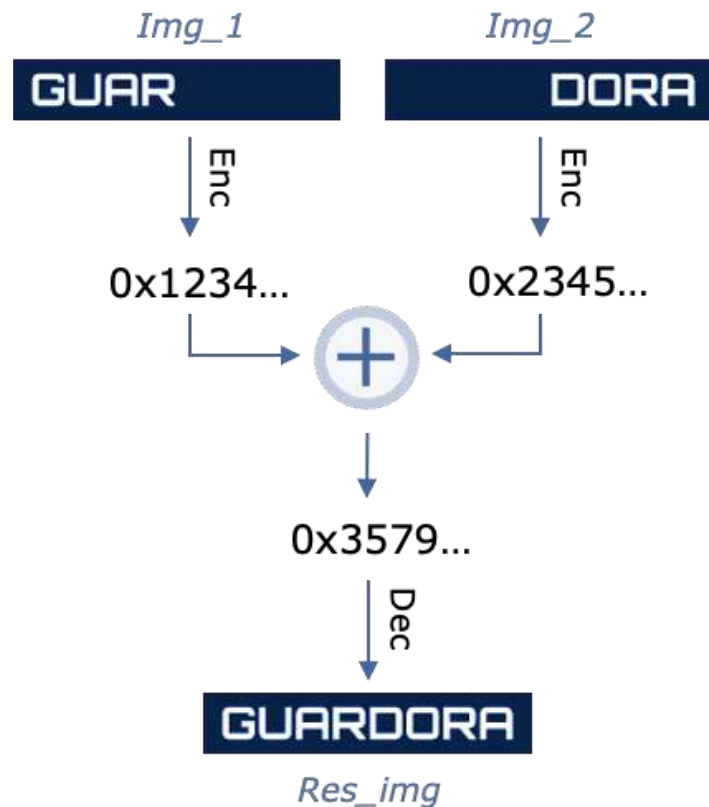
Гомоморфное шифрование

Гомоморфность по сложению

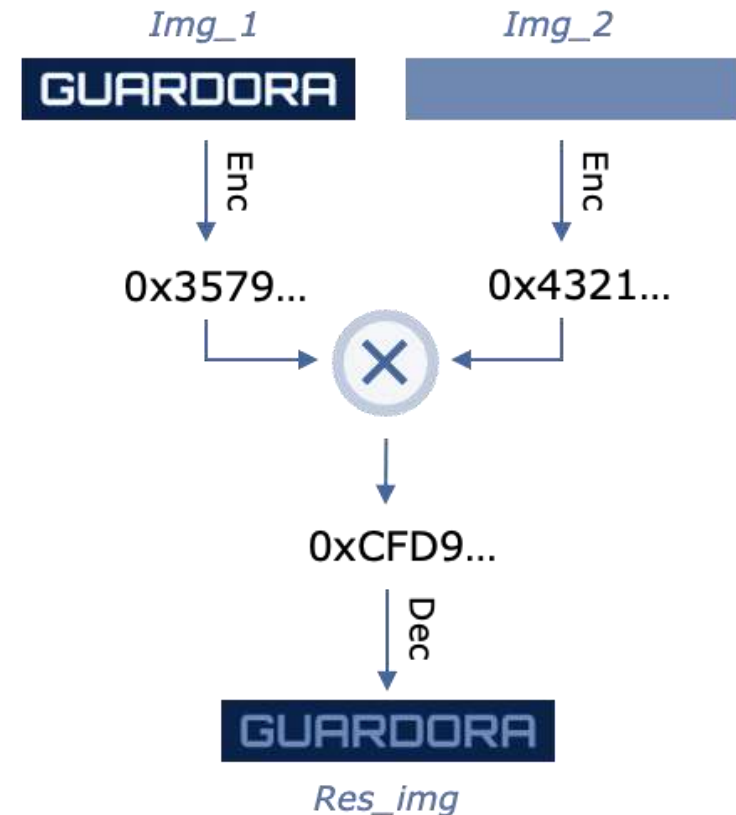


Гомоморфное шифрование

Гомоморфность по сложению



Гомоморфность по умножению



02

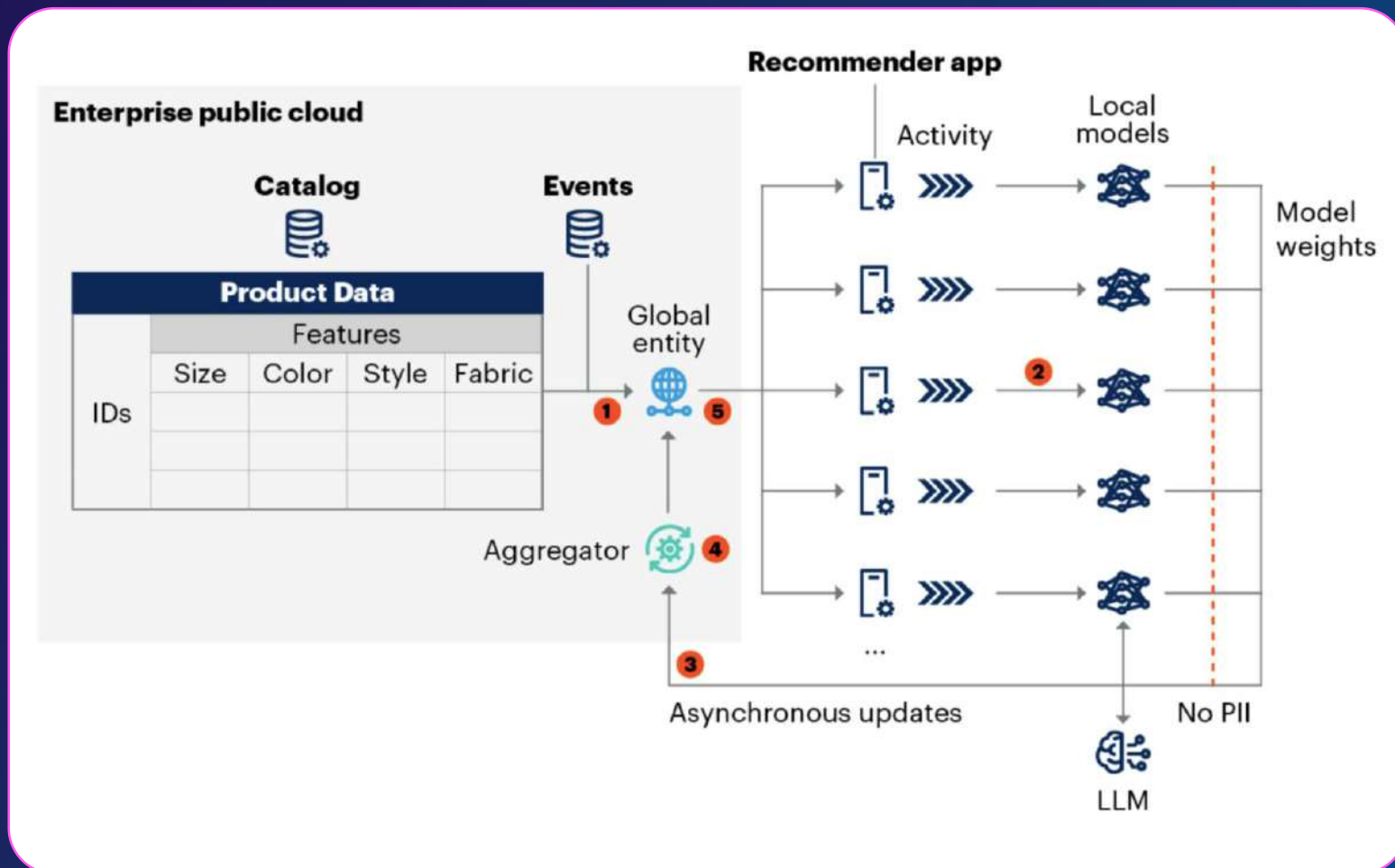


Бизнес-кейсы использования FedML

Пример из маркетинга

Рекомендательные системы

- 1 Разработана базовая модель и встроена в приложение
- 2 Действия пользователей уточняют модель на их устройствах
- 3 Приложение отправляет веса и другие параметры на сервер асинхронно
- 4 Агрегатор консолидирует обновления и уточняет глобальную модель
- 5 Вместе с обновлением приложения устанавливается новая модель



Применение FedML



Маркетинг

Рекомендательные системы (On-Device)

Платформы монетизации данных (DMP)

Применение FedML



Маркетинг

Рекомендательные системы (On-Device)

Платформы монетизации данных (DMP)

Финтех

Антифрод скоринг

Кредитный скоринг

Применение FedML



Маркетинг

Рекомендательные системы (On-Device)
Платформы монетизации данных (DMP)

Финтех

Антифрод скоринг
Кредитный скоринг

Медицина

Диагностика и прогнозирование
заболеваний
Мониторинг (IoT)

Применение FedML



Маркетинг

Рекомендательные системы (On-Device)
Платформы монетизации данных (DMP)

Финтех

Антифрод скоринг
Кредитный скоринг

Медицина

Диагностика и прогнозирование
заболеваний
Мониторинг (IoT)

IoT

Обеспечение конфиденциальности данных
Эффективное использование ресурсов,
снижение сетевой нагрузки
Совершенствование продуктов
(компьютерное зрение, распознавание речи,
умный дом, автономные автомобили ...)

Применение FedML



Маркетинг

Рекомендательные системы (On-Device)
Платформы монетизации данных (DMP)

Финтех

Антифрод скоринг
Кредитный скоринг

Медицина

Диагностика и прогнозирование заболеваний
Мониторинг (IoT)

IoT

Обеспечение конфиденциальности данных
Эффективное использование ресурсов,
снижение сетевой нагрузки
Совершенствование продуктов
(компьютерное зрение, распознавание речи,
умный дом, автономные автомобили ...)

Кибербез

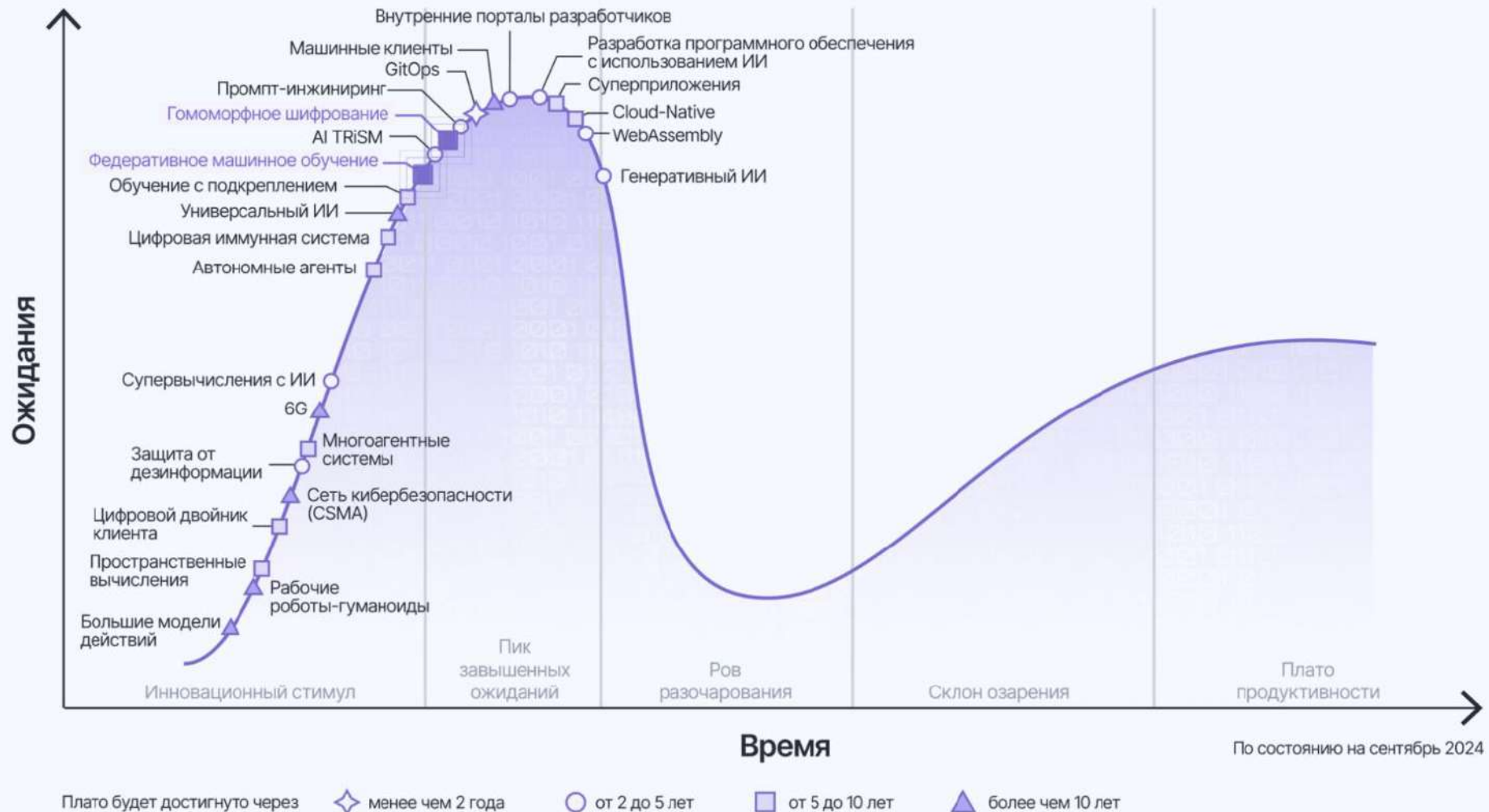
Улучшение ML моделей на данных клиентов
Обнаружение взлома устройств,
вредоносных приложений, дипфейков (On-Device)

03



Перспективы и трудности внедрения FedML

Hype Cycle 2024, Gartner



Факторы развития FedML



Драйверы

Конфиденциальность

Масштабируемость

Эффективное использование ресурсов

Снижение эффекта дрейфа данных

Факторы развития FedML



Драйверы

Конфиденциальность

Масштабируемость

Эффективное использование ресурсов

Снижение эффекта дрейфа данных

Трудности

Необходимость кооперации

Необходимость согласования данных

Новизна технологии

Вычислительные ограничения

ТЕХНОЛОГИИ!
В ТВОИХ
РУКАХ



dmitry.maslov@guardora.ai



guardora.ru



**ТЕХНОЛОГИИ
В ТВОИХ
РУКАХ**



Спасибо