

приоритет2030[^]

лидерами становятся



СПБГЭТУ «ЛЭТИ»
ПЕРВЫЙ ЭЛЕКТРОТЕХНИЧЕСКИЙ



SMARTILIZER RUS

Пилотные проекты СПБГЭТУ «ЛЭТИ» & «Смартилайзер Рус»

Холод Иван Иванович

д.т.н., декан ФКТИ СПБГЭТУ «ЛЭТИ»

Филиппов Евгений Васильевич

к.т.н., исполнительный директор
«Смартилайзер Рус»



Коротко о главном



**Мы ищем крупных промышленных партнеров для
пилотирования новых технологических решений:**

Проект. Платформа анализа распределенных данных (финансовый сектор)

ПРОЕКТ. ПЛАТФОРМА АНАЛИЗА РАСПРЕДЕЛЕННЫХ ДАННЫХ



Направление: Новые подходы к обработке клиентских данных, обогащение данных.

Краткое описание: Реализован прототип платформы машинного обучения распределенных данных без их передачи 3й стороне и без централизованной обработки с использованием технологии федеративного обучения.

Что это дает:

- использовать в обучении конфиденциальные данные;
- обогащать данные с разных источников;
- снизить стоимость владения BigData.

ПЛАТФОРМА АНАЛИЗА РАСПРЕДЕЛЕННЫХ ДАННЫХ



СПбГЭТУ «ЛЭТИ»
ПЕРВЫЙ ЭЛЕКТРОТЕХНИЧЕСКИЙ

SMARTILIZER RUS

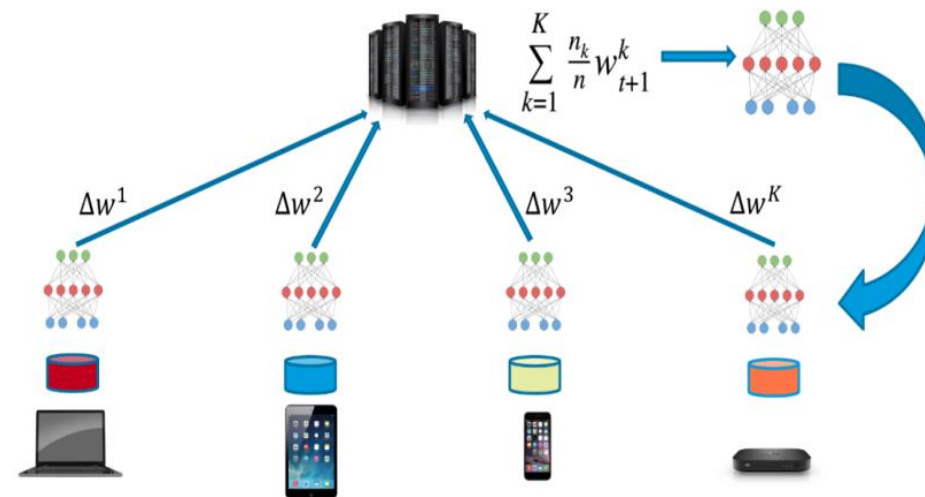
приоритет2030^

лидерами становятся

Платформа анализа информации (в т.ч. конфиденциальной) из разнородных, территориально распределенных источников, без ее передачи третьим лицам с использованием технологии федеративного обучения.

Федеративное обучение - предполагает анализ данных и построения различных аналитических моделей (регрессионных функций, прогнозных моделей и т.п.) на стороне владельца данных без их передачи на сервер для обучения.

Предложена компанией Google Inc в 2017 г.



Преимущества технологии:

- снижаются риски «утечки» информации, т.к. данные остаются у своих владельцев и не передаются;
- происходит обогащение данных, за счет использования в анализе нескольких источников;
- могут использоваться каналы связи с низкой пропускной способностью, т.к. передаются только результаты анализа;
- не требуются большие вычислительные мощности, т.к. анализ выполняется распределено.

ПЛАТФОРМА АНАЛИЗА РАСПРЕДЕЛЕННЫХ ДАННЫХ



СПБГЭТУ «ЛЭТИ»
ПЕРВЫЙ ЭЛЕКТРОТЕХНИЧЕСКИЙ

SMARTILIZER RUS

приоритет2030[^]

лидерами становятся

Области применения для Банков:

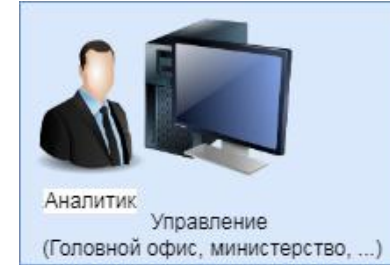
- оценка риска и кредитоспособности клиента;
- формирование персональных предложений;
- подбор кадров и др.;

Источники данных:

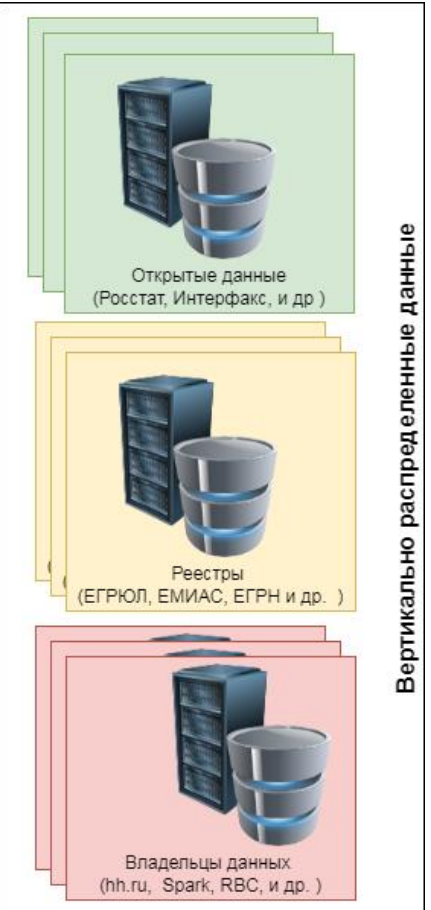
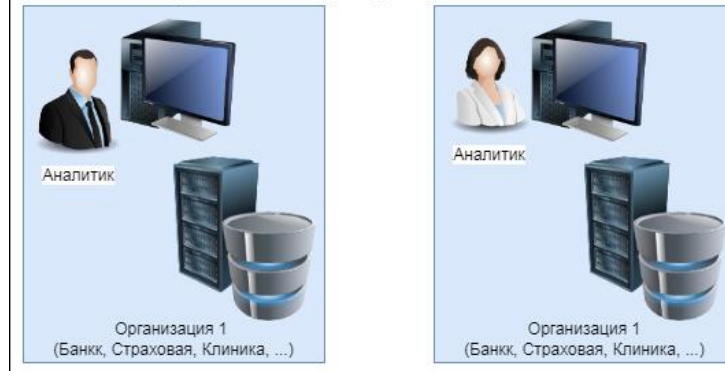
- данные других банков и смежных организаций;
- данные клиентов (приложения смартфонов, датчиков и др.);
- открытые данные (Росстат, Интерфакс, метеоданные, геоданные и др.);
- реестры с частично платным доступом (ЕГРЮЛ, ЕМИАС, ЕГРН и др.);
- данные других компаний (Head Hunter, Spark, Bloomberg, RBC и др.).

Экономический эффект:

- увеличение доходов за счет более персонализированного подхода к клиентам;
- снижение потерь за счет более точной оценки рисков.



Горизонтально распределенные данные



Вертикально распределенные данные

ПЛАТФОРМА АНАЛИЗА РАСПРЕДЕЛЕННЫХ ДАННЫХ



приоритет2030[^]
лидерами становятся

Заделы:

- исследованы существующие фреймворки федеративного обучения:

<https://www.mdpi.com/1424-8220/21/1/167>

<https://www.mdpi.com/1424-8220/22/8/2983>

- разработана библиотека алгоритмов федеративного обучения - FL4J:

<https://ieeexplore.ieee.org/document/9396508>

- создан стенд для апробации применения технологии для разных прикладных задач;
- разрабатываются компоненты для быстрого прототипирования систем федеративного обучения.

ПЛАТФОРМА АНАЛИЗА РАСПРЕДЕЛЕННЫХ ДАННЫХ



СПБГЭТУ «ЛЭТИ»
ПЕРВЫЙ ЭЛЕКТРОТЕХНИЧЕСКИЙ
SMARTILIZER RUS

приоритет2030[^]
лидерами становятся

Текущий уровень УТГ 2-3 - завершается разработка компонентов для быстрого прототипирования и создания систем анализа конфиденциальной информации:

- модуль владельца данных для настройки доступа к данным;
- реестр данных для хранения информации о доступных данных;
- модуль анализа данных для настройки процесса анализа данных.

