

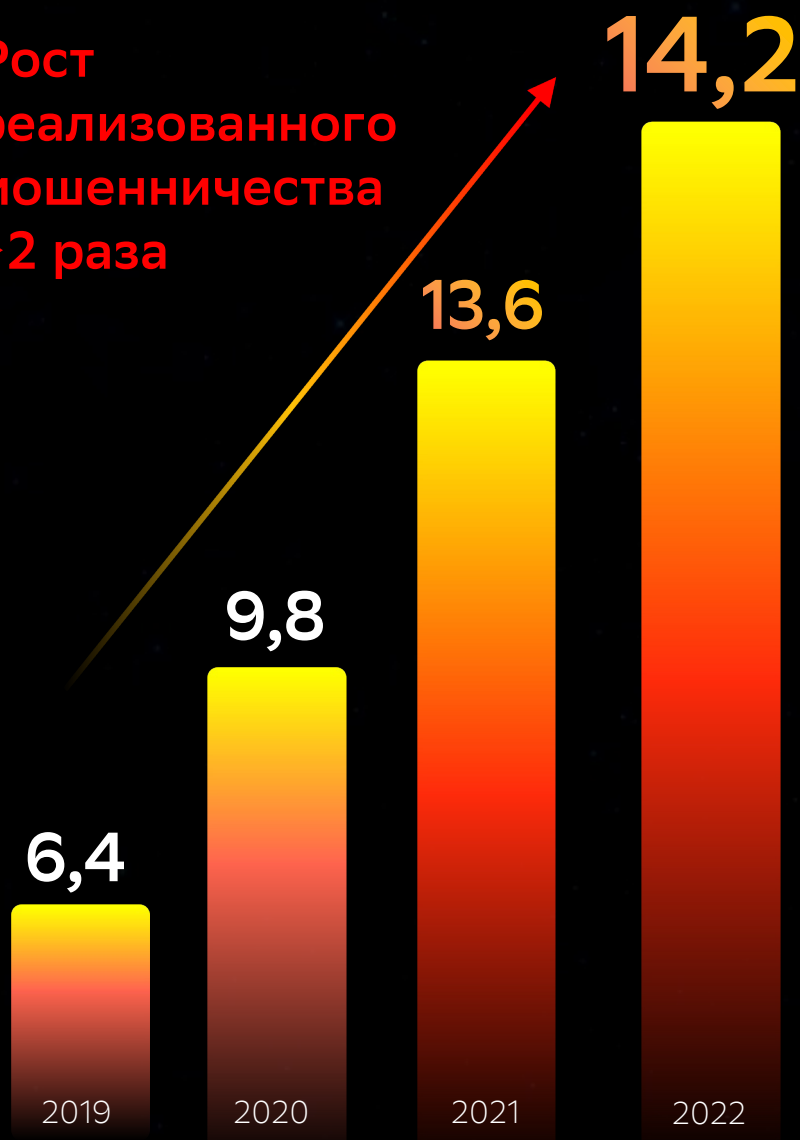


AI модели для выявления мошенничества против клиентов Сбера

Мошенничество - масштабы проблемы



Рост
реализованного
мошенничества
>2 раза



по данным ЦБ РФ (млрд руб.)

5 млрд

попыток
телефонного
мошенничества
в 2022г.

6 млн

рисковых
событий
у операторов
связи в сутки

94%

доля социальной
инженерии во всем
объеме мошенничества

>200 млрд руб.

общая сумма кредитов,
которую граждане РФ пытались
оформить под воздействием
мошенников за 2022 г.

25%

хищений совершается
с использованием
кредитных средств

500 млн руб.

максимальная сумма,
похищенная у одного гражданина

83%

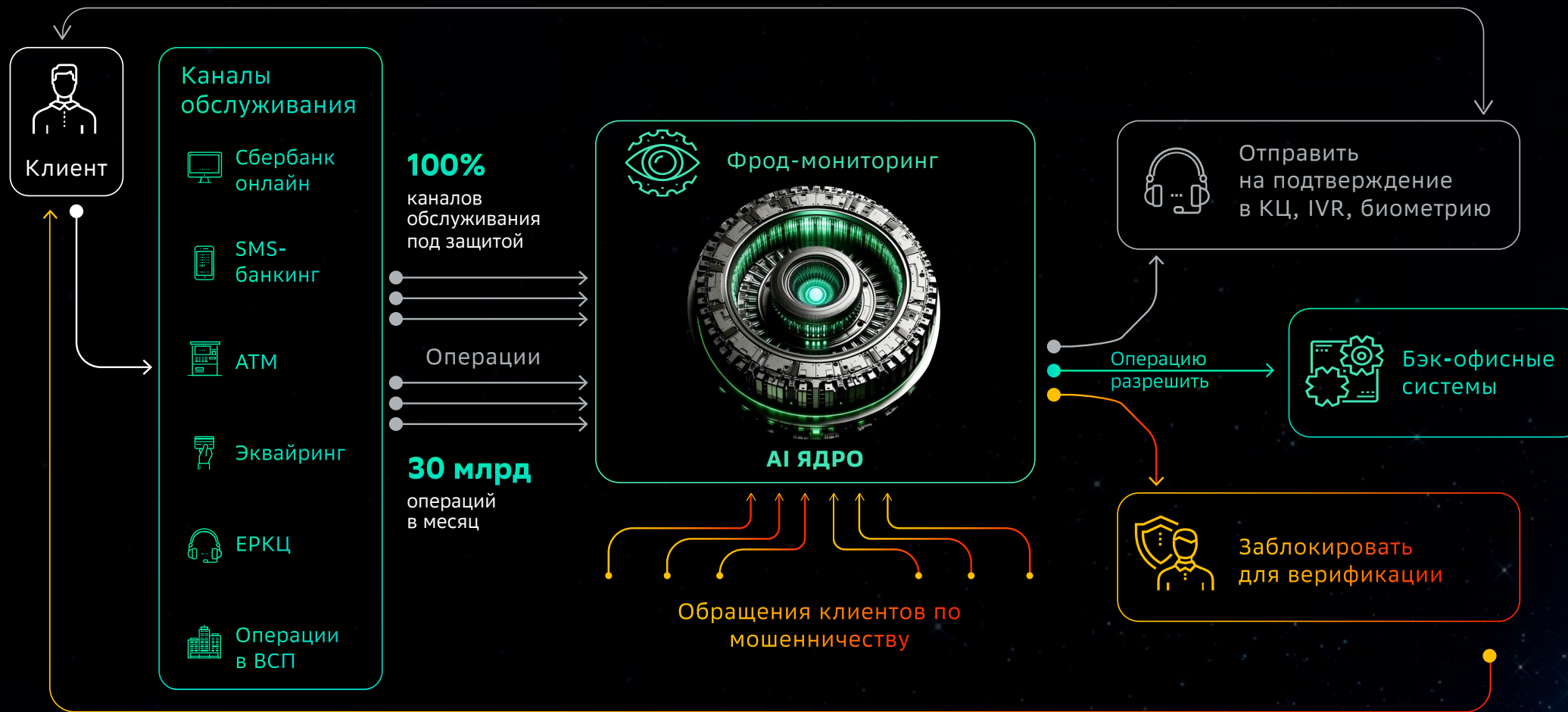
россиян столкнулись
с попыткой
кибермошенничества

Антифрод на страже средств клиентов



Проверка в online-режиме каждой **из 30+ млрд операций/мес с SLA < 50 мс.**

Без AI задачу решить невозможно!



Вызовы при создании AI Ядра Антифрода



Для эффективного решения задачи противодействия мошенничеству система должна:

- выдерживать колоссальную нагрузку: **25 тыс. транзакций/сек с SLA < 50 мс** по каждой из них
- **позволять одновременно работать десяткам AI-моделей** (проведение экспериментов, выделенные модели под определенные каналы/схемы фрода, разные типы моделей)
- комбинировать для оценки риска транзакций как **исторические данные, так и online-контекст проведения операции**
- иметь **низкий T2M фичей (до 48 часов) и моделей (до 10 дней)** и **автоматически дообучаться** по мере поступления новых данных о фроде (мошенники постоянно адаптируются к внедряемым защитным мерам и изменяют схемы фрода)



Используемые модели в AI Ядре



ВЗАИМОДЕЙСТВИЕ С КЛИЕНТОМ

Языковые модели выявляют клиентов под воздействием мошенников



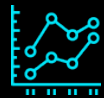
ПОВЫШЕНИЕ КАЧЕСТВА ДАННЫХ

Осуществляют классификацию жалоб для фильтрации бытового мошенничества, ошибочных обращений



REAL-TIME СКОРИНГ ТРАНЗАКЦИЙ

Направлены на оценку риска транзакций с latency в 1 мс



ГРАФОВЫЕ МОДЕЛИ

Выявляют связи между клиентами и другими сущностями, оценивают близость, кластеризуют мошеннические группы



ГЕО-МОДЕЛИ

Выявляют нетипичные геолокационные паттерны перемещений, аномальные скорости передвижений и пр.

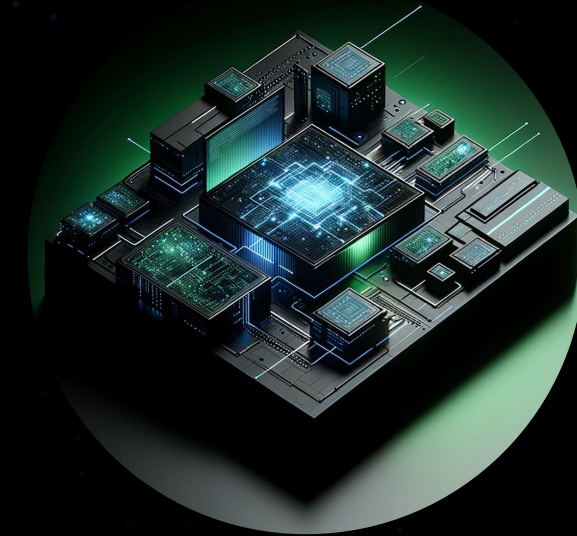


СКОРИНГИ СУЩНОСТЕЙ

Проводят оценку различных сущностей (ФЛ, ЮЛ, телефонов, устройств) по различным негативным аспектам

20+ AI-моделей осуществляют всестороннюю оценку риска транзакций, участвующих сущностей (клиент, устройство, получатель, мерчант, ...), их связей и пр.

Ключевые компоненты AI Ядра



Feature Store:

- система расчета и поставки высококачественных offline-фичей произвольной глубины для AI-моделей
- **150+ витрин с 2000+ фичами по 30+ источникам**
- **от 2-3 часов до 1-2 дней** - проверка гипотез по новым фичам за счет функционала back-filling (эмуляция расчетов по истории)
- **до 24 часов** – T2M новых фичей силами Data Scientis'ов без участия разработчиков/администраторов

Дополнительная обработка:

- **300+ расчетов online-фичей** (24 часа контекста предшествующих операции) в режиме реального времени
- объединяет offline и online-фичи, обогащает ими транзакцию
- осуществляет параллельное исполнение real-time AI-моделей **с SLA < 50 мс**, позволяет проводить A/B тестирование
- **до 7 дней** – T2M новых моделей без участия разработчиков/администраторов

Эффект от использования AI Ядра в Антифрод



200+ спасенных средств
клиентов
млрд руб/год

99,5% Эффективность
системы

0.11 Fraud basis
points*

20+ Одновременно
работающих
AI-моделей



Разработанное E2E решение позволяет Data Scientist'у от создания фичей и обучения моделей до вывода их пройти все шаги БЕЗ участия разработчиков/администраторов

7 часов – средний Time2Market новых фичей

8 дней – средний Time2Market новых моделей

* рассчитывается как отношение объема фрода к обороту