

Отари Меликишвили

Машинное обучение в бизнесе.
Простой и эффективный подход к внедрению.



Время больших данных

Экономический успех уже определяют технологии искусственного интеллекта и машинного обучения³

50%

процессов к 2025г. в сфере обработки и анализа данных будут автоматизированы с помощью ИИ, что снизит острую нехватку высококвалифицированных специалистов¹

80%

времени специалистов по обработке и анализу данных тратится на поиск, очистку и организацию данных²

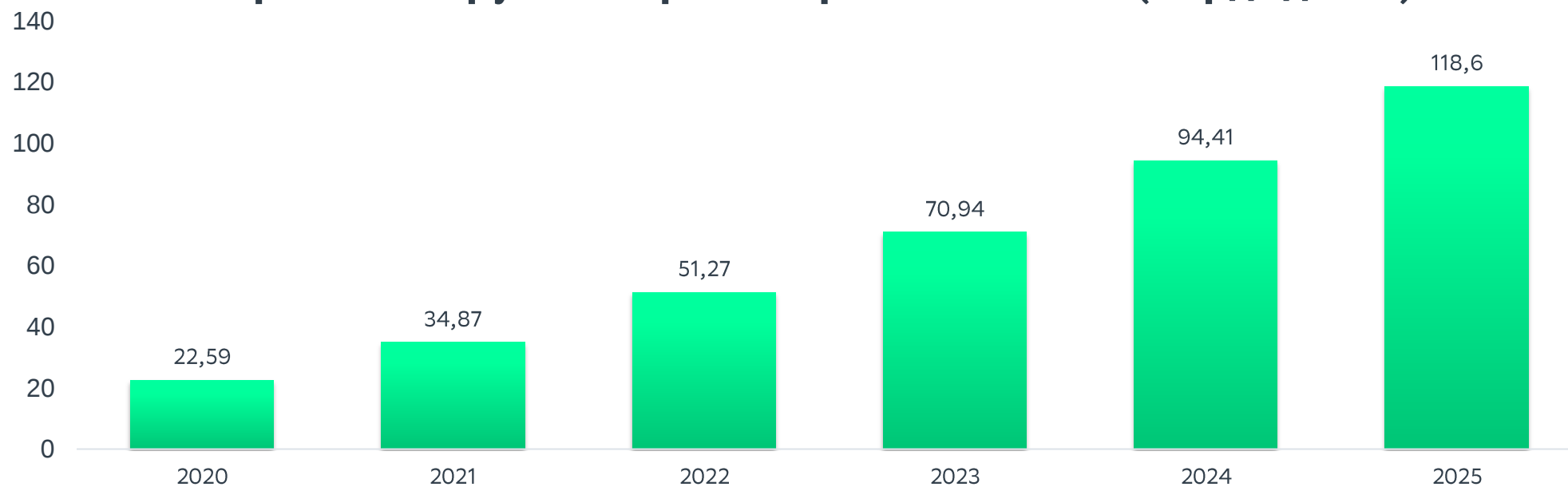
40%

групп разработчиков к 2023 году будут использовать AutoML-сервисы по сравнению с 2% в 2019 году³



Пандемия коронавируса COVID-19 привела к росту расходов на искусственный интеллект в мире⁴

Прогноз выручки мирового рынка AI SaaS (млрд. долл.)



источник: statista.com





ETL

Извлечение, загрузка и трансформация данных.

Достаточно настроить правила, по которым данные будут выгружаться и обрабатываться автоматически



Feature Engineering

Работа с датасетами

Теперь все в едином окне (как Spark кластера, так и любые другие инструменты под рукой)



Machine Learning

Обучение модели.

Обучение моделей либо упрощено (интегрированы мониторинги и др. инструменты), либо выполняется автоматически:
AutoML



Production

Деплой модели и получение результатов бизнес-модели.

Для вывода модели в пром. среду – достаточно пары действий, пользователь получает микросервис с моделью и автомасштабированием



При переходе на AI SaaS у клиентов:



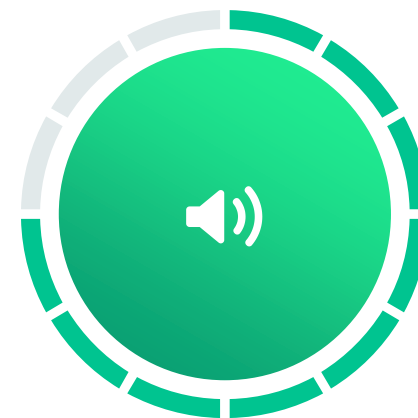
на **82%**

снизился срок
реализации задач с Big
Data



на **57%**

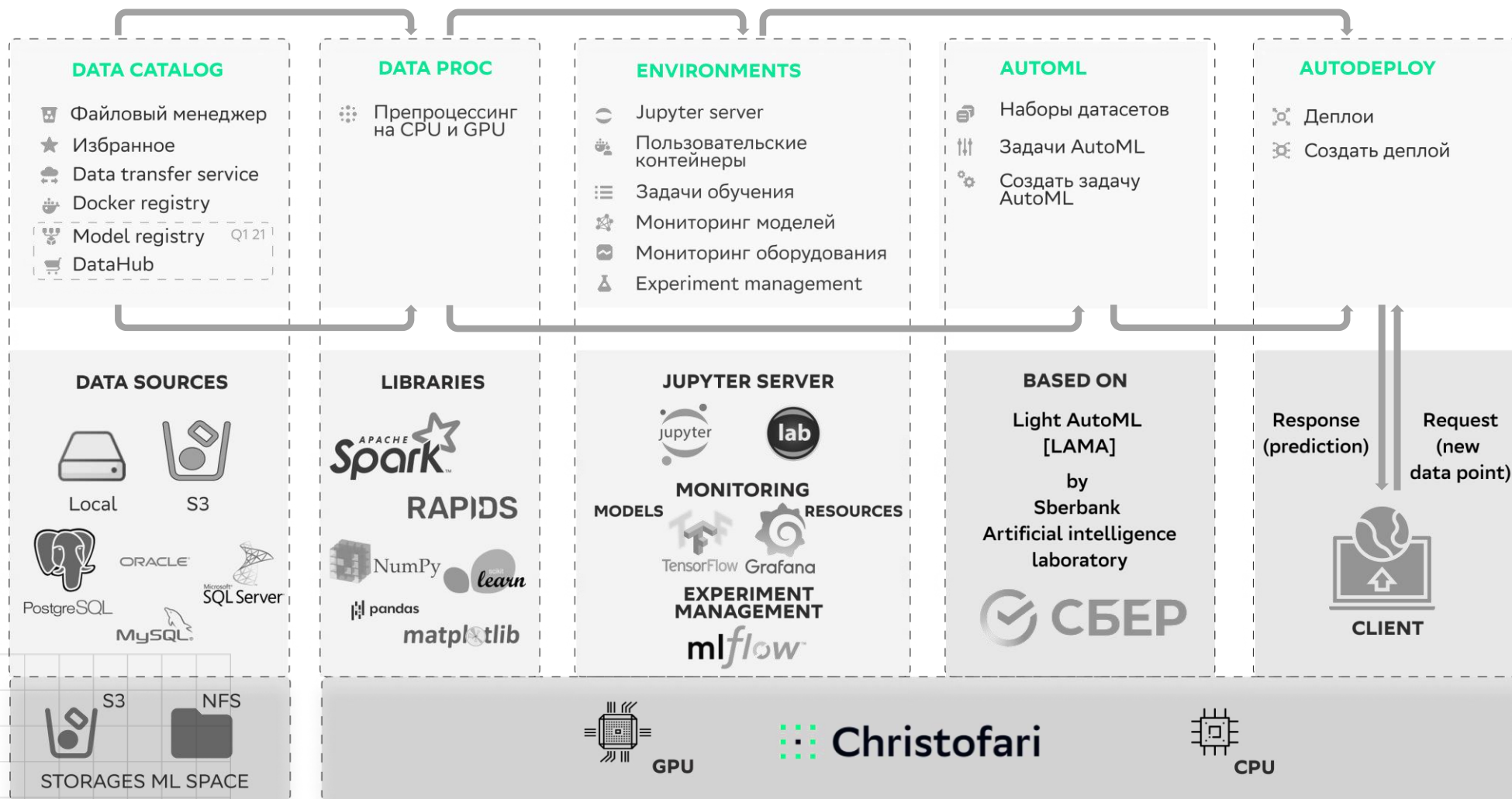
снизились затраты на
ML-разработку



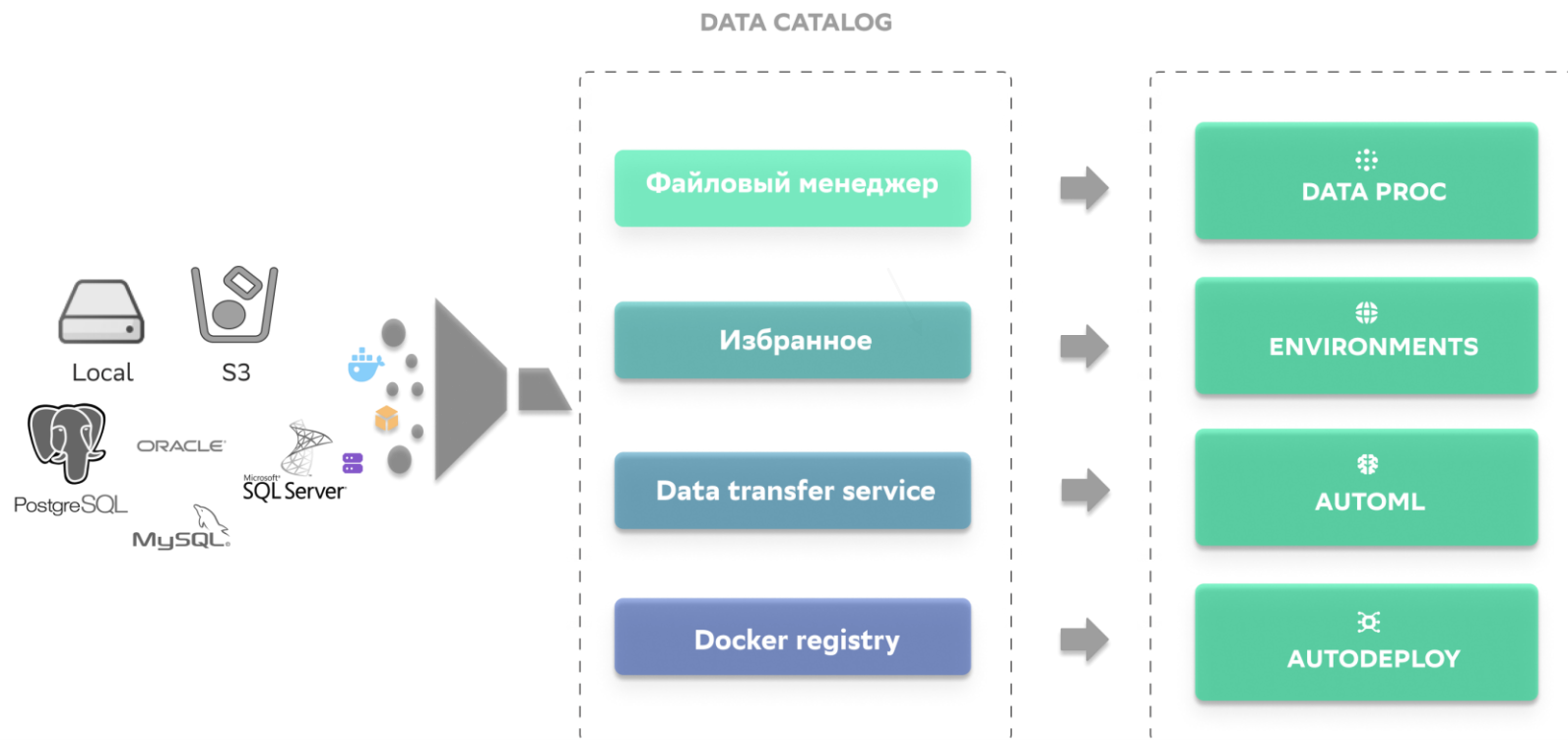
на **70%**

увеличилось количество
R&D задач

ML Space Новая SaaS-платформа



Петабайты данных в совместном управлении

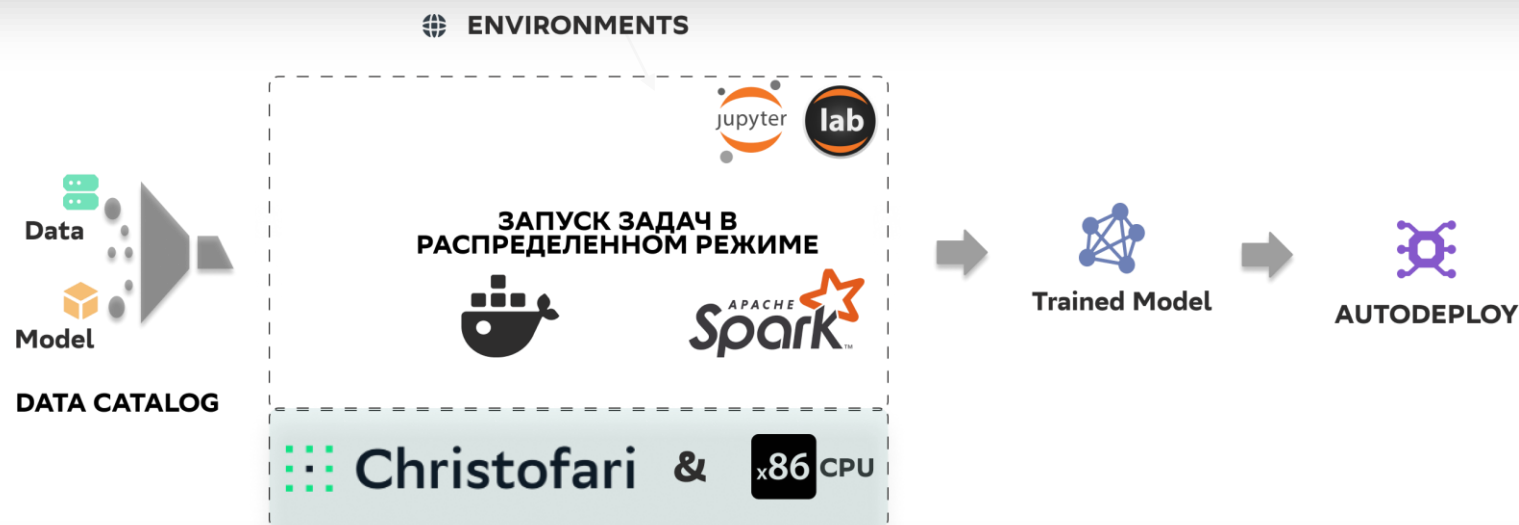


Набор сервисов для загрузки, хранения, анализа, управления доступом и жизненным циклом данных и артефактов машинного обучения (датасетов, моделей, докер контейнеров и др.).

Перенос BigData из любых источников из файловых систем и баз данных по пользовательским правилам.

Готовые среды для обучения моделей с помощью суперкомпьютера Christofari или Spark-кластера

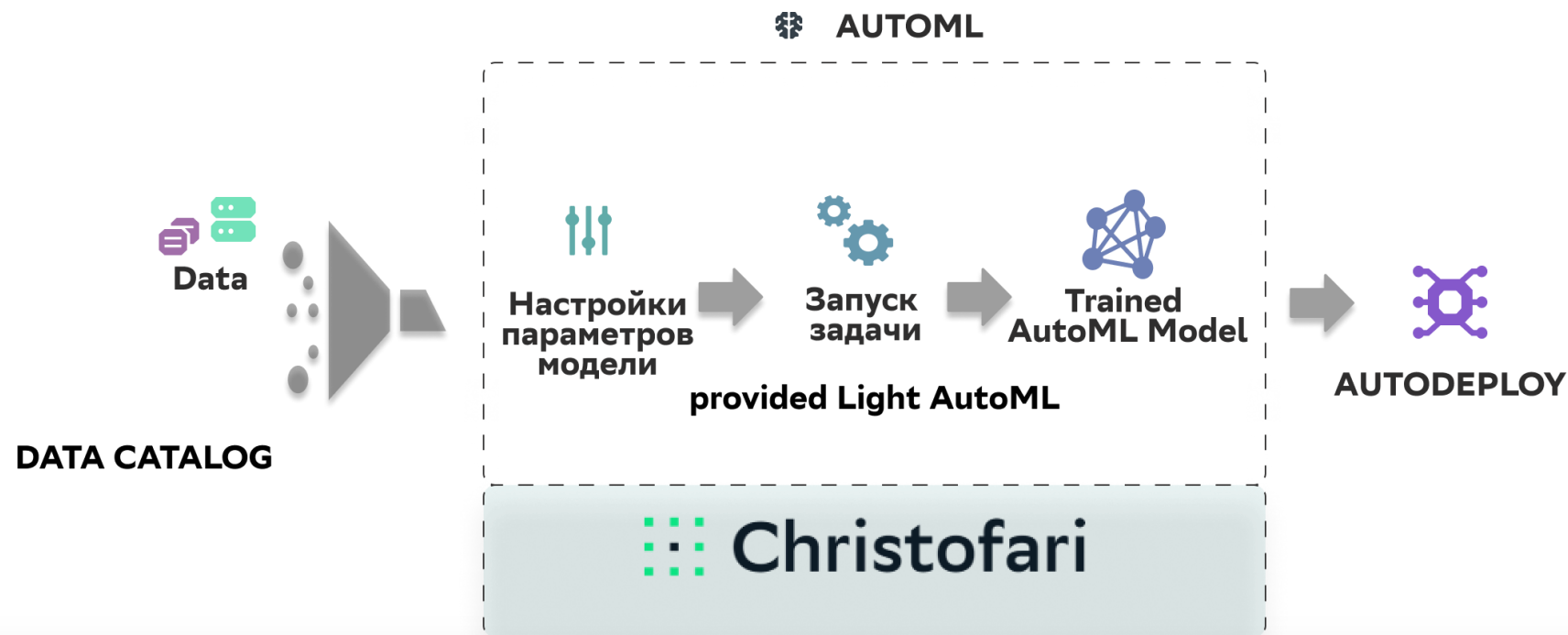
Обучение на привычных Jupyter® Notebook или Jupyter® Lab.
В качестве среды обучения вы можете использовать любой ваш docker-образ.



Есть все необходимые утилиты для мониторинга потребления ресурсов, моделей и эксперимент-менеджмента - Grafana®, TensorBoard® и MLflow®.

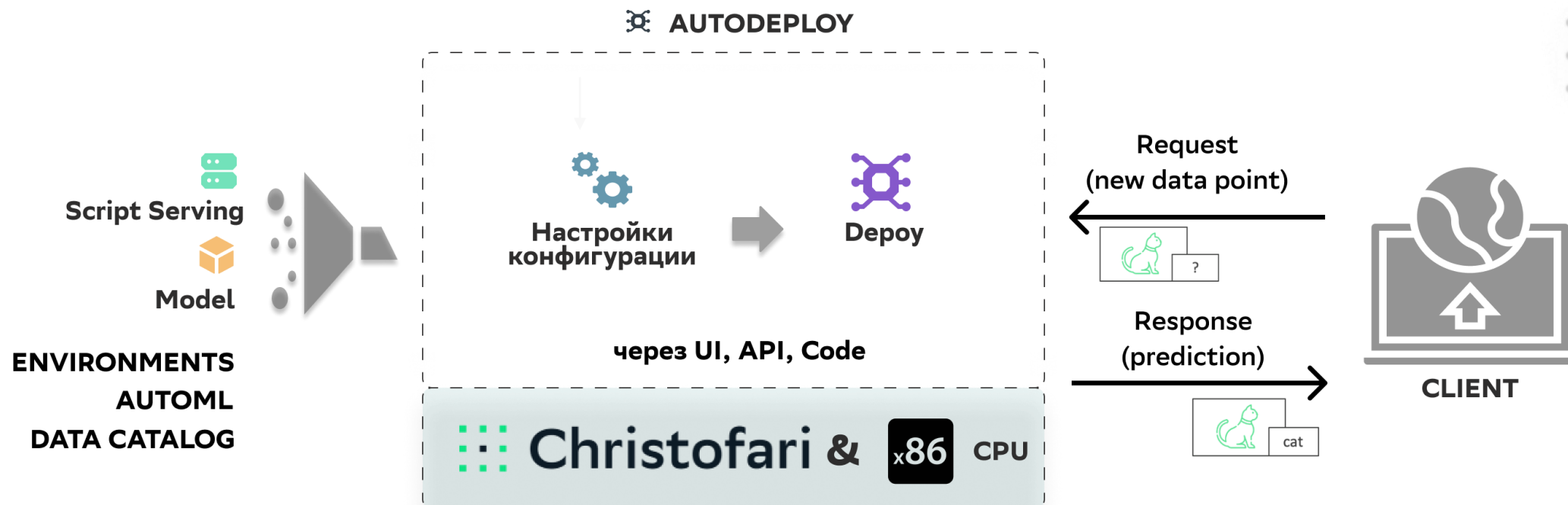
Отслеживайте загрузку ресурсов в разрезах CPU, GPU, RAM в любой момент времени.

Автоматическое построение модели с помощью AutoML от grandmasters Kaggle



AutoML позволяет построить модель с высоким качеством, приняв на вход только данные и минимум параметров обучения. Обученную модель можно так же развернуть в 1 клик.

Применение обученных AI моделей в виде микросервисов для распознавания / прогнозирования новых данных



Автоматический деплой моделей на высокопроизводительных мощностях в два клика - все можно сделать через UI, API, Code. Модуль совместим с любыми фреймворками машинного и глубокого обучения. Вычислительные ресурсы легко и быстро подбираются в соответствии с вашими задачами, потребностями и бюджетом.

Для бизнеса

- Возможность реализации **сложных AI-задач** с Big Data
- **Готовые** сервис и инфраструктура под клиента **в одном окне**
- **Быстрая** реализация AI-задачи **под ключ** (в т.ч. комплексными силами в рамках DS as a Service – AI Cloud + команда клиента)
- **Оптимизация операционных задач** клиента при создании AI-продуктов
- **Сокращение Time-to-market** - быстрый запуск новых AI-продуктов
- **Опережение** конкурентов и прогрессивное **развитие** бизнеса

01

Для Data Scientists

- Возможность **командной работы** с общим хранилищем
- **Широкий выбор ресурсов** для ML задач - 1-1200 GPU
- Все **популярные фреймворки и библиотеки** уже установлены и настроены
- Привычные **Jupyter Notebook и JupyterLab** для работы
- **Оптимизация и упрощение** работы за счет **AutoML и AutoDeploy**

02

А также:

- **Техническая и консультационная** поддержка
- **Круглосуточная** поддержка работоспособности инфраструктуры
- **Защита данных** на уровне **банковских** стандартов безопасности
- Соблюдение **152 ФЗ**

03

Описание

- С помощью ML Space и Christorafi была обучена **впервые на русском языке** трансформерная модель GPT-3 с **760 миллионов параметров**.
- Для обучения использовали **терабайт текстовых данных**.
- В обучающую выборку вошло более **400 тысяч книг, свыше 7 миллионов научных статей, вся Википедия** на русском. А также свыше **35 тысяч проектов с GitHub**.
- Именно поэтому GPT-3 может генерировать **осмысленные** предложения и программный код.

Результат

С помощью GPT-3 стало возможным создавать действительно **прорывные продукты**.

Она позволяет принципиально улучшить понимание машинами **естественного языка** и максимально **эффективно** решать множество задач из области NLP.

Модель выложена в открытом доступе.

GPT-3 от Сбера



Что значит любовь?



Дополнить

Результат:

Что значит любовь?
А как ты узнал, что что-то главное..?
Это состояние души-ЖИЗНЬ)))
Бог есть любовь.
"Любовь долготерпит, милосердствует, любовь не завидует, любовь не превозносится, не гордится." (1Тим. 6:7).
Женщина пребывает в любви, но она не принадлежит никому. Живя в одиночестве, она привязывается к мужчине. Настоящая любовь живёт в людях, которые ею живут.
Людям дана великая благодать.

Описание

Сервис СберЗдоровье и облачная платформа SberCloud, входящие в экосистему Сбербанка, запустили совместный проект по **распознаванию снимков компьютерной томографии легких..**

Любой человек может загрузить на сайт свой снимок КТ и получить его расшифровку в течение 5-10 минут на почту.



Результат

Система использует алгоритмы искусственного интеллекта, разработанные в Лаборатории по искусственному интеллекту Сбербанка и работает на вычислительных мощностях суперкомпьютера «Кристофари», компании SberCloud.

На сегодняшний день она умеет показывать в процентах пострадавшую часть легких, **детектировать** пораженные участки легких (матовое стекло и консолидация), и **рассчитывать их общий объем**. Размеченные DICOM-серии файлов можно скачать и посмотреть в любой программе для просмотра DICOM снимков.

- Такой анализ позволяет пользователю самостоятельно **оценить динамику** при повторном исследовании КТ, чтобы следить за развитием заболевания, не выходя из дома.
- Врачам сервис поможет быстро и просто **расшифровывать** снимки пациентов. На основе результатов автоматического анализа возможно реализовать приоритизацию рабочего списка рентгенолога и категоризировать исследования.

01

02

Описание

ОБРАБОТКА ДОКУМЕНТОВ ДЛЯ ПАО СБЕРБАНК

DS-команда ПАО Сбербанк использовала платформу ML Space и мощности Christofari для решения следующих задач:

- **распознавание текста** на документах различных типов (тренировка обобщенного OCR)
- **выделение необходимых сущностей** из документов - полей, сумм, действующих лиц и т.п. (тренировка моделей NER&ImageNER).

Результат

«Специфика данных задач заключается в том, что для тренировки нейросетевых моделей **необходим перебор большого количества параметров** и **проверка различных гипотез**, а также их создание зачастую требует гигантских подготовленных объемов текстов.

Нам удалось быстро реализовать данные задачи и проверить наши гипотезы за счет следующего функционала ML Space:

- Возможность использования **распределенного обучения при помощи Horovod и Pytorch distributed**, что критически важно для больших моделей и корпусов документов;
- Удобство **проведения и сравнения** множества **экспериментов** благодаря встроенным компонентам **Experiment Management** (MLflow, Tensorboard)».

Галицкий Игорь

Исполнительный директор по исследованию данных

01

02

Описание

SberDevices — компания экосистемы Сбербанка, центр экспертизы по решениям на основе искусственного интеллекта в таких областях применения, как речевые технологии, технологии понимания естественного языка, лицевая и голосовая биометрия.

Компания также фокусируется на создании умных устройств для конечных потребителей и корпоративных клиентов.

Результат

SberDevices разработала следующие продукты с помощью ML Space и Christofari:

■ РАСПОЗНАВАНИЕ РЕЧИ Семейство голосовых ассистентов «Салют»

Для их создания с помощью Christofari были обучены более 70 различных моделей ASR. И еще большее количество моделей Text-to-Speech.

Благодаря этому сегодня каждый из вас может общаться с помощниками - Джой, Афина и Сбер через мобильное приложение Сбербанк-Онлайн или SberBox - умное устройство для телевизора.

01

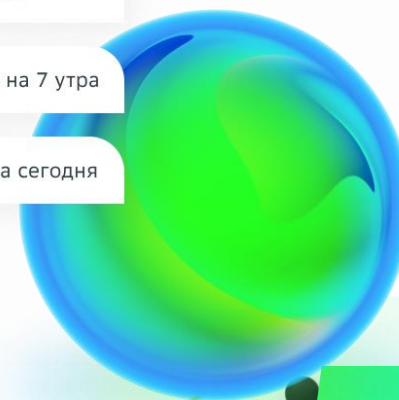


Салют, отправь бабушке денег

Поставь будильник на 7 утра

Погода сегодня

02



Описание

- В ходе ЭКО оплодотворение и рост эмбрионов происходит в искусственных условиях с использованием специальных чашек
- Этапы развития эмбрионов фиксируются специальными камерами с малой периодичностью
- Все снимки специалисту-эмбриологу нужно изучать вручную

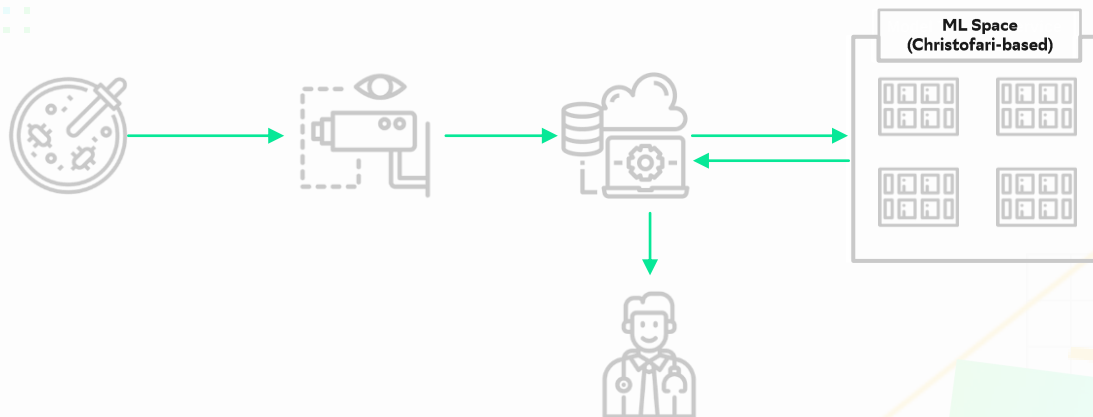
А также

- Большой объем снимков влечет большие трудозатраты
- Возможна человеческая ошибка

01

Результат

- Модель компьютерного зрения **автоматически** изучает все снимки
- Нейронная сеть **эффективно и быстро** обучается и дообучается на суперкомпьютере Christofari на сервисе ML Space Environments
- Эмбриологу поступают **только требующие внимания** снимки, **снижая рутинный труд** и риск человеческой ошибки.
- Повышается **эффективность лечения** бесплодия методами ЭКО



02

ML SPACE



Отари Меликишвили

Лидер направления AICloud | SberCloud | СБЕР

