

Возможности DataSphere для Стартапов

Дмитрий Рыбалко
Консультант по Data Science, Yandex Cloud



Решения по ML-разработке можно разделить на две большие группы

In-house

Требует внутренних мощностей
для обучения

Дорогой и долгий путь: на этапе обучения
модели требуют много ресурсов, позже эти
мощности будут простаивать

Нужно самостоятельно определять
и оптимизировать стек технологий

Cloud

Обучение моделей в облаке

Более простой вариант для старта:
не требует больших финансовых вложений

Можно арендовать виртуальные машины
и настраивать их самостоятельно
или работать с уже оптимизированными
инструментами, такими как DataSphere

Решения по ML-разработке можно разделить на две большие группы

In-house

Требует внутренних мощностей
для обучения

Дорогой и долгий путь: на этапе обучения
модели требуют много ресурсов, позже эти
мощности будут простаивать

Нужно самостоятельно определять
и оптимизировать стек технологий

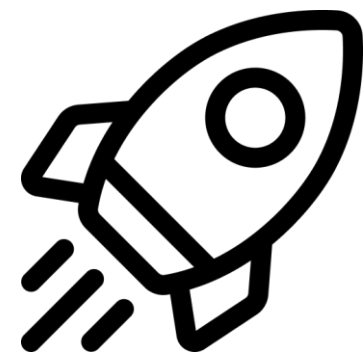
Cloud

Обучение моделей в облаке

Более простой вариант для старта:
не требует больших финансовых вложений

Можно арендовать виртуальные машины
и настраивать их самостоятельно
или работать с уже оптимизированными
инструментами, такими как DataSphere

DataSphere — облачный сервис для разработки ML-решений



Максимально простой старт



Минимум необходимых компетенций внутри компании

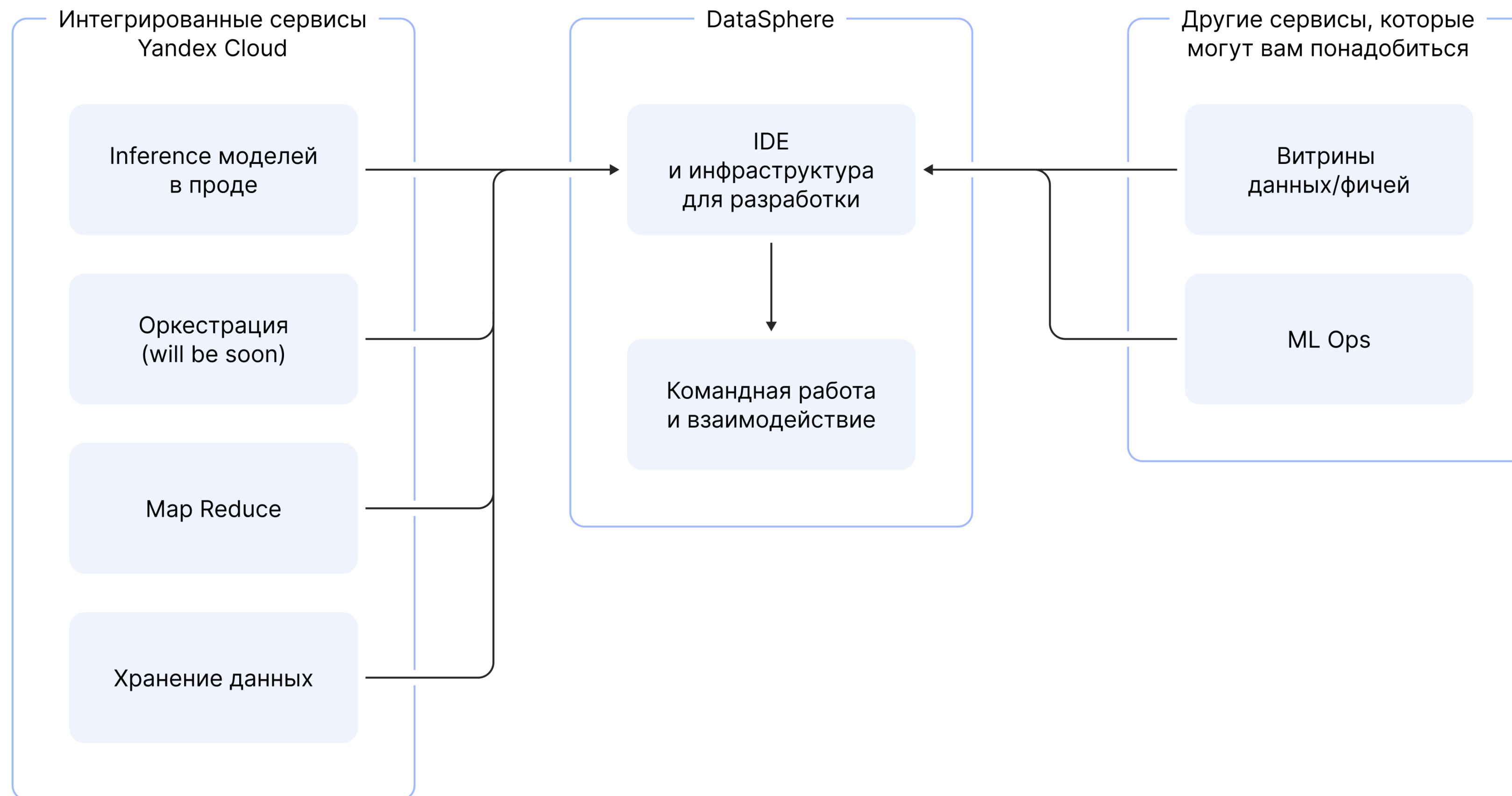


Доступность из любой точки мира



Интеграция в экосистему облачных решений

DataSphere и всё, что нужно для ML-разработки



Ключевые сущности ролевой модели DataSphere

Сообщество

Командное пространство для лёгкого обмена ресурсами между участниками

Проект

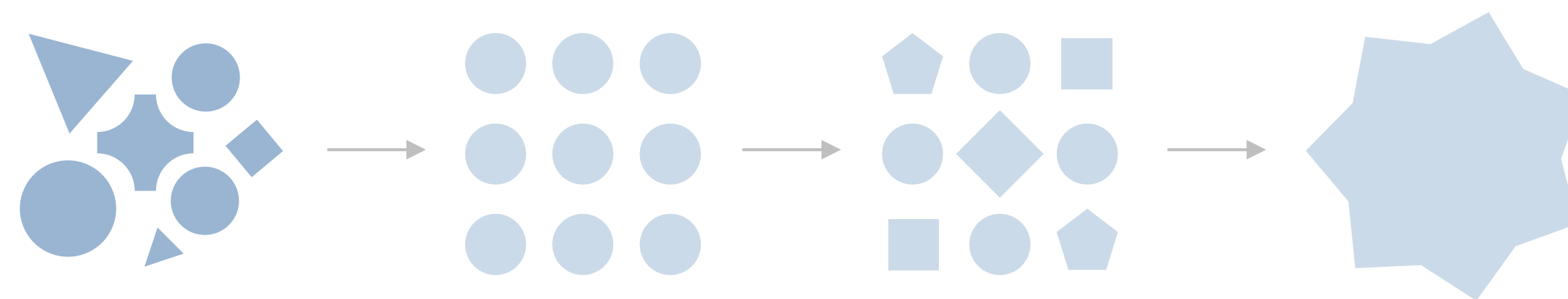
Рабочее пространство каждого участника команды

Необходимость проверки гипотез

Быстрое
прототипирование новых
сервисов и продуктов

- Процесс разработки ML-моделей требует большого количества экспериментов
- Для старта работы с моделями глубокого обучения требуются специализированные ресурсы GPU
- Тяжело заранее оценить потребности в ресурсах
- Закупка и поддержание серверов с GPU - сложный, дорогостоящий и длительный процесс

Быстрая проверка НОВЫХ ГИПОТЕЗ



Как поможет DataSphere

Запуск рабочего места за несколько секунд

Экономия на дорогостоящих
DevOps специалистах и ML
инженерах

Гибкое управление бюджетами

Оплата только за время
вычислений
Экономия на закупке и
обслуживании собственных
вычислительных ресурсов

Совместное использование ресурсов

Вычислительные мощности,
код, результаты
исследований, окружения

Привлечение внешних команд

Разработка совместно с
партнерами без долгой
настройки доступов во
внутренний контур

Yandex [DataSphere](#) позволяет быстро разработать прототип ML-сервиса, чтобы проверить бизнес-идею и принять решение о дальнейшем развитии

В DataSphere есть множество конфигураций на основе CPU и GPU

- Yandex Cloud предоставляет ML/DL-командам мощные конфигурации, содержащие до восьми NVIDIA A100 на каждую виртуальную машину
- Возможность использовать GPU NVIDIA V100 и A100
- Возможность ограничения используемых ресурсов и трат в проекте

Serverless - ключевая особенность в DataSphere

Не нужно держать
ноутбук открытым
всё время
вычислений



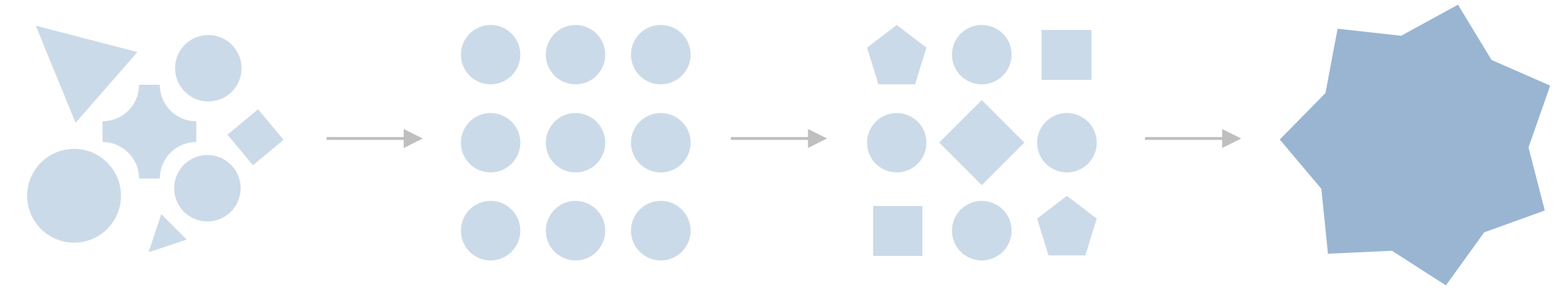
Вы платите только
за время
вычисления:
мощное железо
появляется только
в нужный момент



Гибкое
управление
вычислительными
ресурсами
для максимальной
эффективности



Деплой моделей



Типичные проблемы

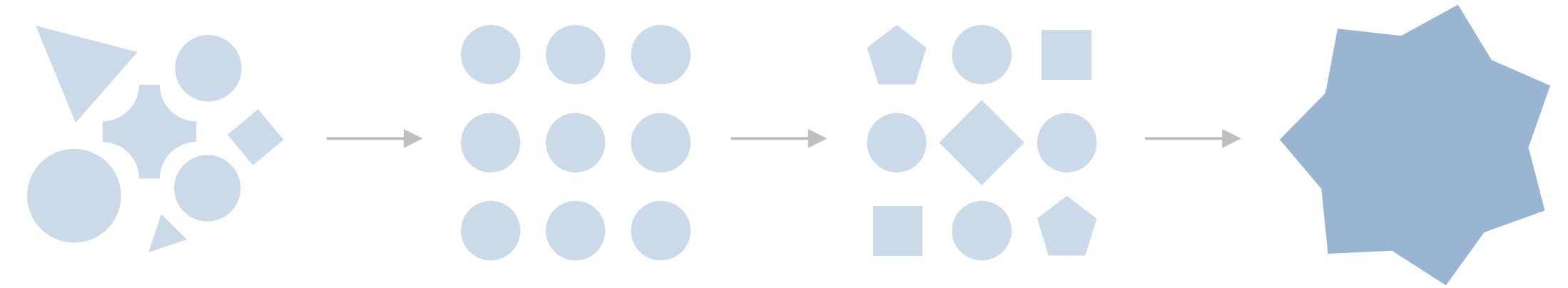
Привлечение разработчиков

Чтобы превратить модель в сервис, нужно привлекать дополнительные ресурсы дорогостоящих специалистов

Содержание инфраструктуры

Необходимо самостоятельно настраивать мониторинг инфраструктуры и своевременно реагировать на рост нагрузки

Деплой моделей



Как поможет DataSphere

Привлечение разработчиков

Чтобы превратить модель в сервис, нужно привлекать дополнительные ресурсы дорогостоящих специалистов

Публикация модели в один клик. Возможность быстро запускать эксперименты, параллельно публикуя нужные версии моделей.

Инструменты для выстраивания релизного процесса: zero downtime с плавным переключением и делением трафика, A/B-тестирование, эксперименты с моделями в проде

Содержание инфраструктуры

Необходимо самостоятельно настраивать мониторинг инфраструктуры и своевременно реагировать на рост нагрузки

Инфраструктура облака работает на вас: она всегда доступна, стабильна и оптимальна по стоимости.

Мониторинг, логирование, балансировка и доставка трафика доступны из коробки

Кейсы использования

Мониторинг промышленных садов

Задача: создать систему прогнозирования урожайности, учитывающую погодные условия и сортовые особенности плодовых культур

Решение: использовали сервисы облачной платформы для хранения и обработки большого количества данных с помощью нейросетей и ML-алгоритмов

Результаты: распределённая система мониторинга насаждений и урожая успешно протестирована на площадках в ботаническом саду МГУ и в экспериментальном саду ФНЦ имени Мичурина

9

облачных сервисов

используются для создания системы прогнозирования

первый опыт

создания систем прогнозирования урожая в России

Охрана Байкала

Задача: создать систему мониторинга изменений состояния фито- и зоопланктона

Решение: развернули облачную инфраструктуру, создали бэкенд и пользовательский интерфейс

Результаты: первый прототип распознаёт несколько видов планктона. Пользовательский интерфейс позволяет сотрудникам НИИ биологии ИГУ самостоятельно дообучать алгоритм и контролировать качество его работы

774
формы

рачков должна
распознавать система

75
лет

анализ проб
производился
вручную

5 млн

записей о еженедельном состоянии
экосистемы уже накоплено

Рекомендательная система на базе облачной платформы

Задача: рекомендовать товары к корзине клиента для увеличения продаж сети Burger King

Решение: рекомендательная система в Yandex DataSphere на основе комбинации бизнес-правил, математических моделей и результатов тестирования

Результаты: быстрая проверка гипотез, внесение изменений, получение результатов заказчиком. Точность рекомендаций выше экспертной

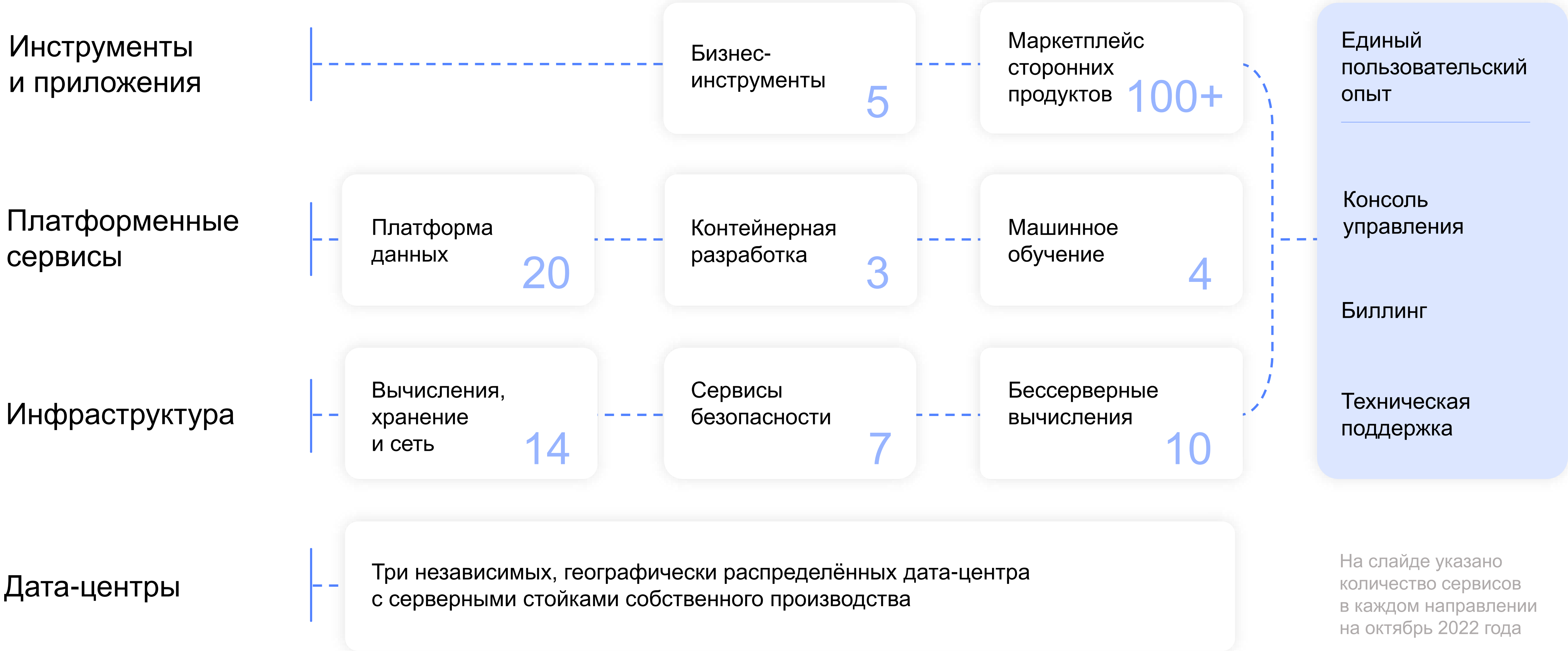
40%

выросла эффективность рекомендаций

менее суток

на внесение изменений
живого тестирования в модель

Преимущества единой облачной экосистемы



ML сервисы Yandex.Cloud

Елена Белоброва,
менеджер по развитию
бизнеса

Yandex Vision — сервис компьютерного зрения для анализа изображений

Варианты использования сервиса

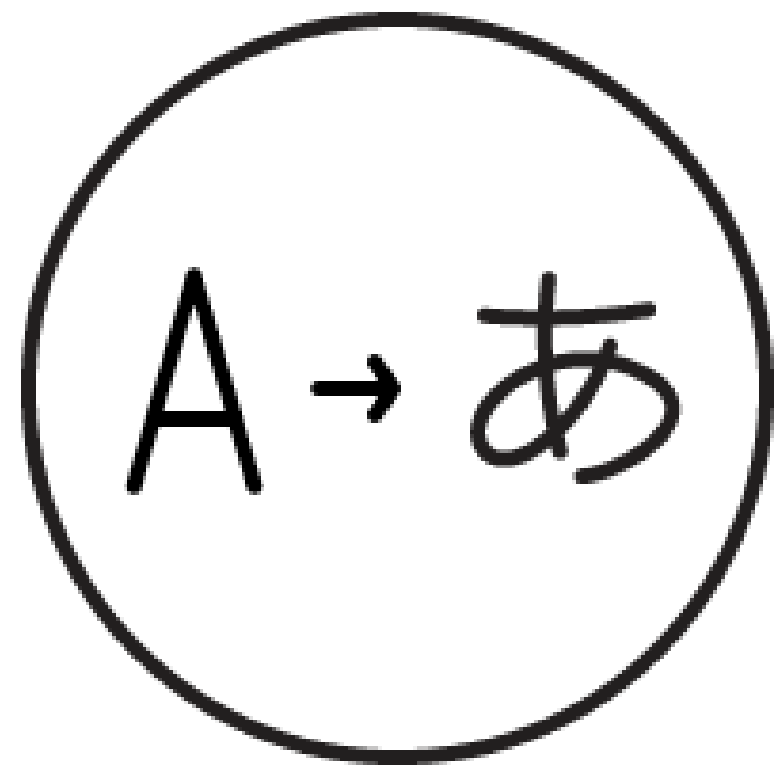
Распознавание
текста
и шаблонов
документов

Поиск дубликатов
и семантически
близких
изображений

Определение
качества
изображений

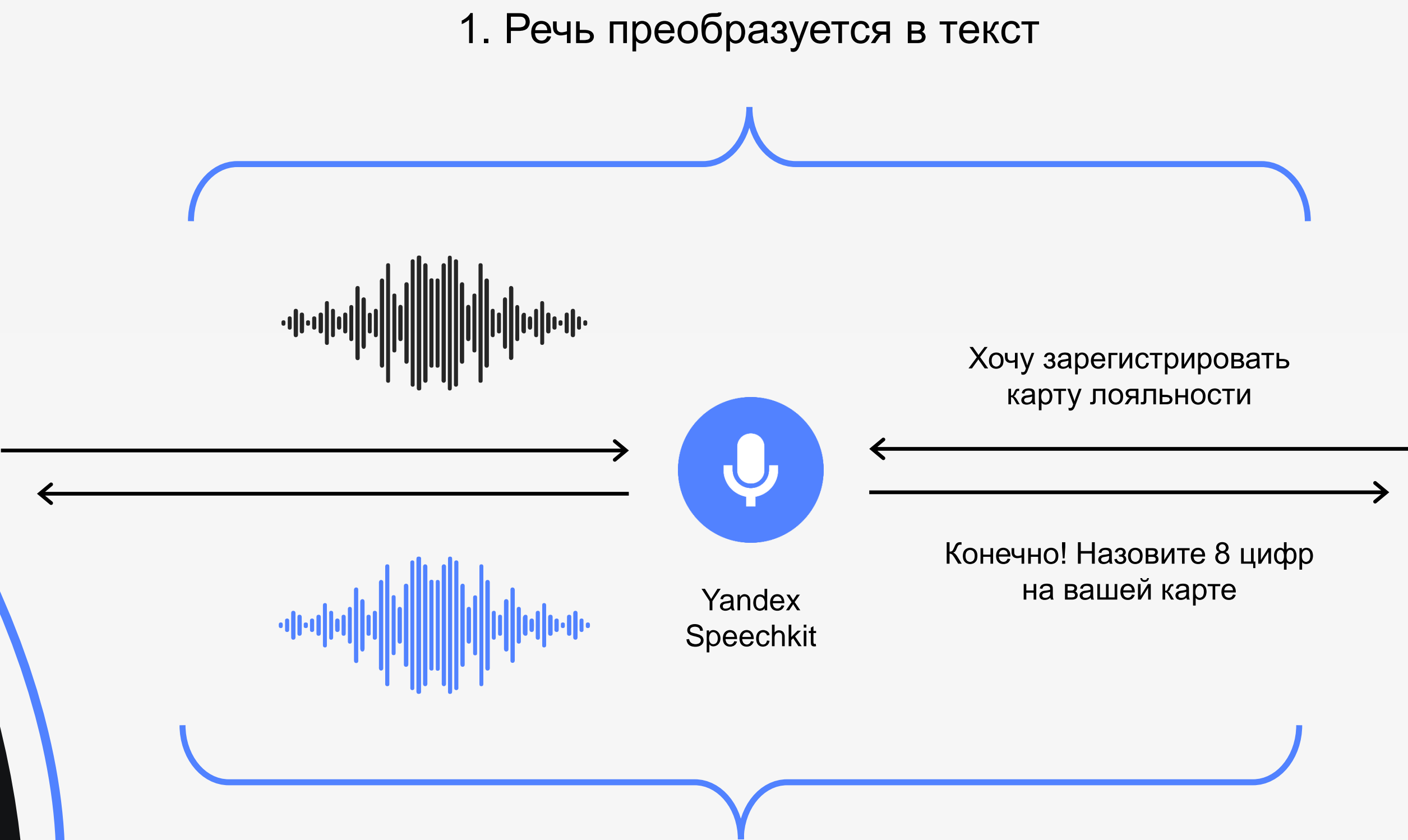
Модерация
пользовательского
контента

Translate API



- › 90+ языков
- › Авто-определение языка
- › Поддержка собственных глоссариев
- › Возможность дообучения модели
- › Коммерческая версия обеспечивает полностью конфиденциальный перевод

SpeechKit позволяет роботам слышать и говорить



2. Текст преобразуется в речь

Распознавание речи.

Разные режимы для разных задач

Роботы и ассистенты

Речевая аналитика и
протоколирование

Озвучка роботов

Озвучка книг и статей

Потоковое распознавание

Ассинхронное и отложенное
распознавание файлов

One-shot синтез, аудио шаблоны

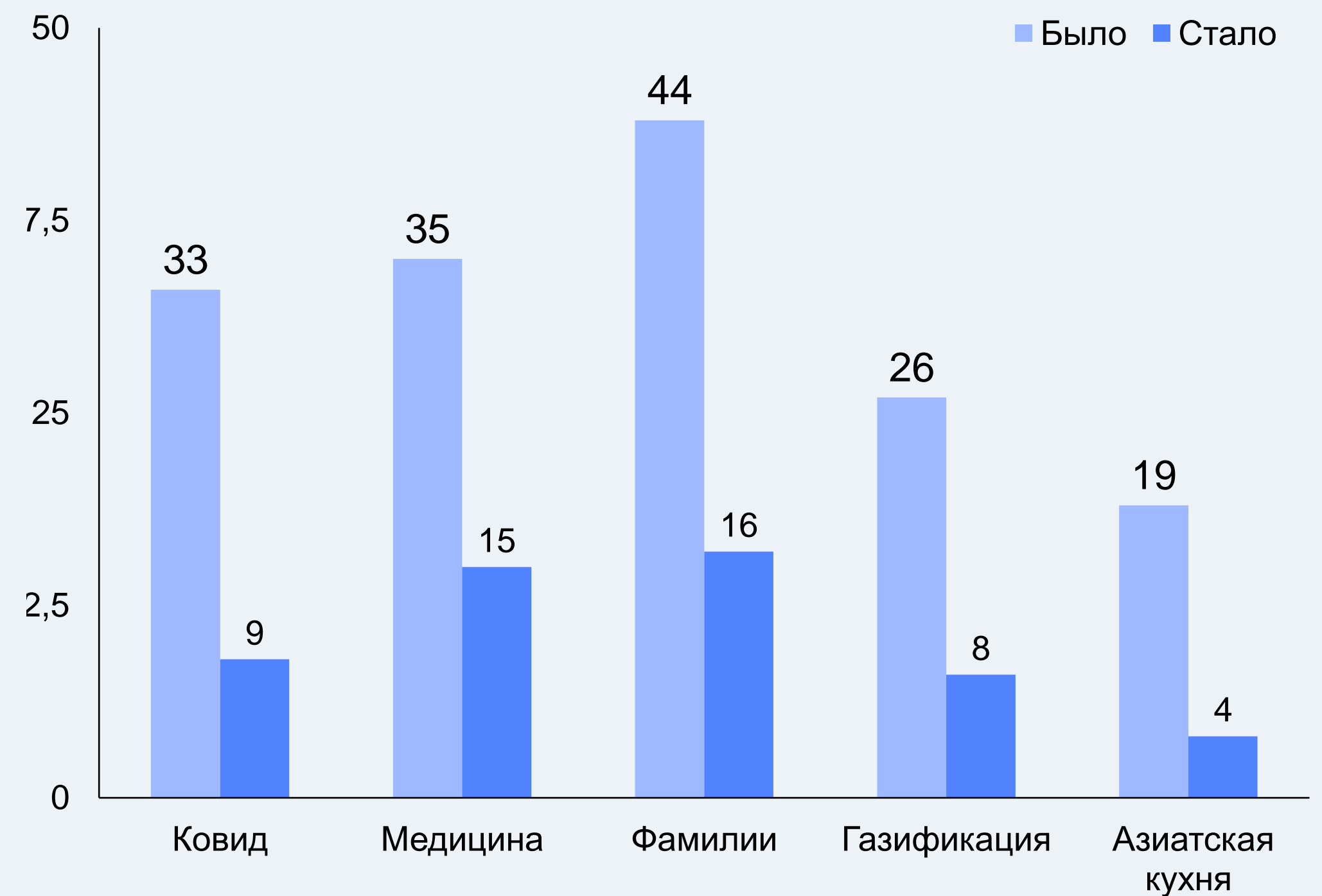
7+голосов для любых сценариев

Кастомизация для предметной области

Готовы адаптировать
для ваших доменов

- На основе аудио
- На основе текстов из предметной области
- На основе шаблонных фраз, ожидаемых от пользователя

Примеры кастомизации на русском



SpeechKit Brand Voice доступен в трёх версиях

🕒 4 месяца

Brand Voice Premium

Полно-текстовый синтез в нескольких амплуа или с необходимыми нюансами речи.

Голосовые помощники
Умный дом
Озвучка контента

🕒 1 месяц

Brand Voice Self Service

Самостоятельно запишите диктора по нашим рекомендациям и получите полнотекстовый синтез.

Общение с клиентами в любых каналах
Озвучка подкастов и передач
Информирование клиентов

🕒 1 минута

Brand Voice Call Center

Моментальный синтез аудио с изменяемыми параметрами (имя, время, адрес и т. д.) на основе короткого аудио-шаблона.

Автоматизация контакт-центров
Мультиголосовые роботы-операторы

Примеры one-shot синтеза

Шаблон

Могу предложить записать вас 11 февраля на {time}

Оригинал

10 часов

Синтез

12 часов

Шаблон

Здравствуйте, Лидия! Меня зовут {name}.
Чем я могу вам помочь?

Оригинал

Ленар

Синтез

Юрий

Спасибо! Люблю отвечать на вопросы

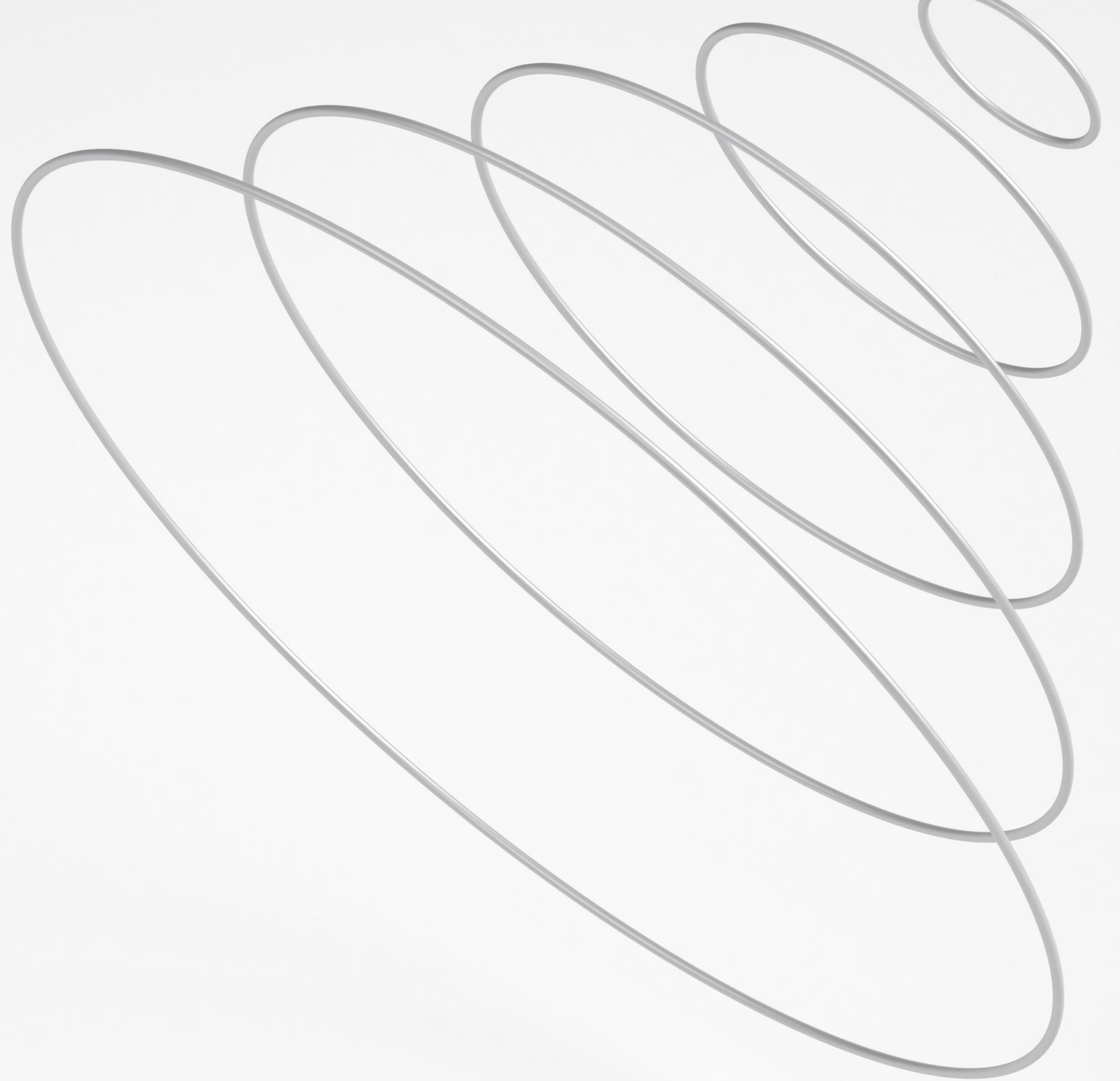


Елена Белоброва,
менеджер по развитию
бизнеса

Yandex  Cloud

Yandex Cloud Boost

Иван Копытов
Программа поддержки стартапов



Что входит в Yandex Cloud Boost

Грант на сервисы Yandex Cloud

400 000 рублей для резидентов Иннополис и участников программ



Технологическая поддержка

Консультации архитектора,
обучение



Инструменты для бизнеса

B2B-сервисы Яндекса



Продвижение

Совместный PR, Marketplace,
мероприятия



Грант на сервисы Yandex Cloud

Infrastructure

100 000 рублей

- Compute Cloud и Monitoring
включая видеокарты
- VPC и Network Load Balancer
- Object Storage
- Marketplace
- DataSphere, SpeechKit, Translate, Vision

Managed DB и другие

300 000 рублей

- Managed Service for PostgreSQL, ClickHouse, MongoDB, Redis, Apache Kafka, Greenplum
- Managed Kubernetes
- Serverless (Functions, Containers, YDB)
- Повышение уровня поддержки

Зачем стартапам облачные ресурсы?

1

Чтобы экономить на покупке серверов и поддержке инфраструктуры

2

Чтобы быстрее проводить эксперименты

3

Чтобы бесконечно масштабироваться при росте бизнеса

Что кроме Yandex Cloud?

Практикум

Помощь в подборе IT- специалистов

Сервис по подбору IT-
и диджитал-специалистов
уровня Junior и Middle
от Яндекс Практикума.
Публикация вакансий
среди выпускников
и подбор кандидатов



Система управления тестированием

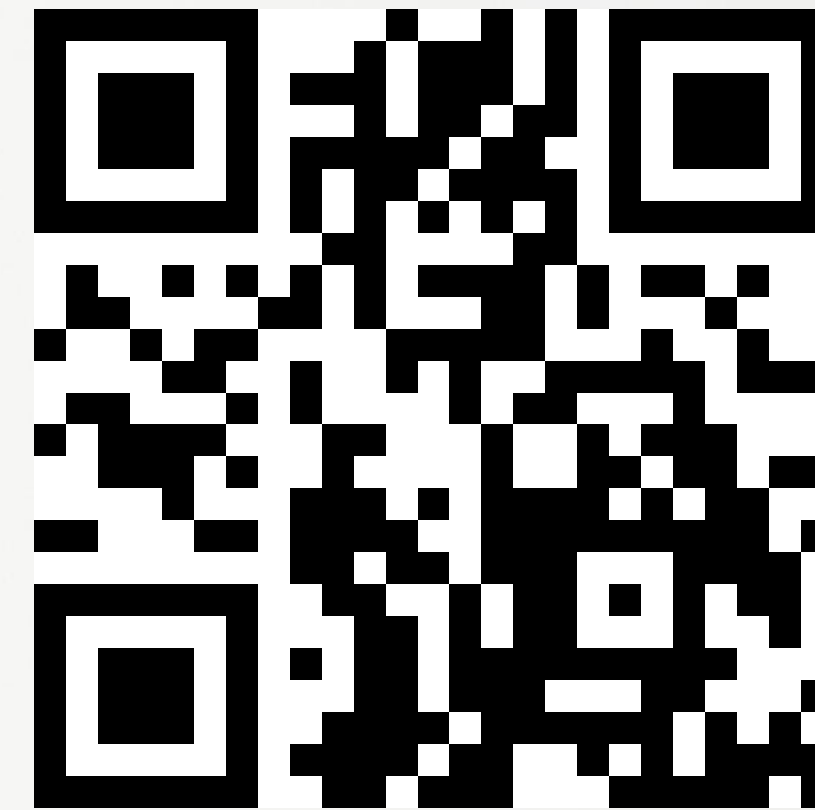
2 месяца бесплатно
от компании TEST IT

Толока

Разметка и сбор данных

\$300 на краудсорсинг-
платформу Яндекс Толока,
классификация 125 000
объектов, сбор 100 000
изображений или сегментация
до 40 000 объектов

Спасибо!



Иван Копытов

Программа поддержки стартапов Yandex Cloud Boost

ivankopytov@yandex-team.ru