

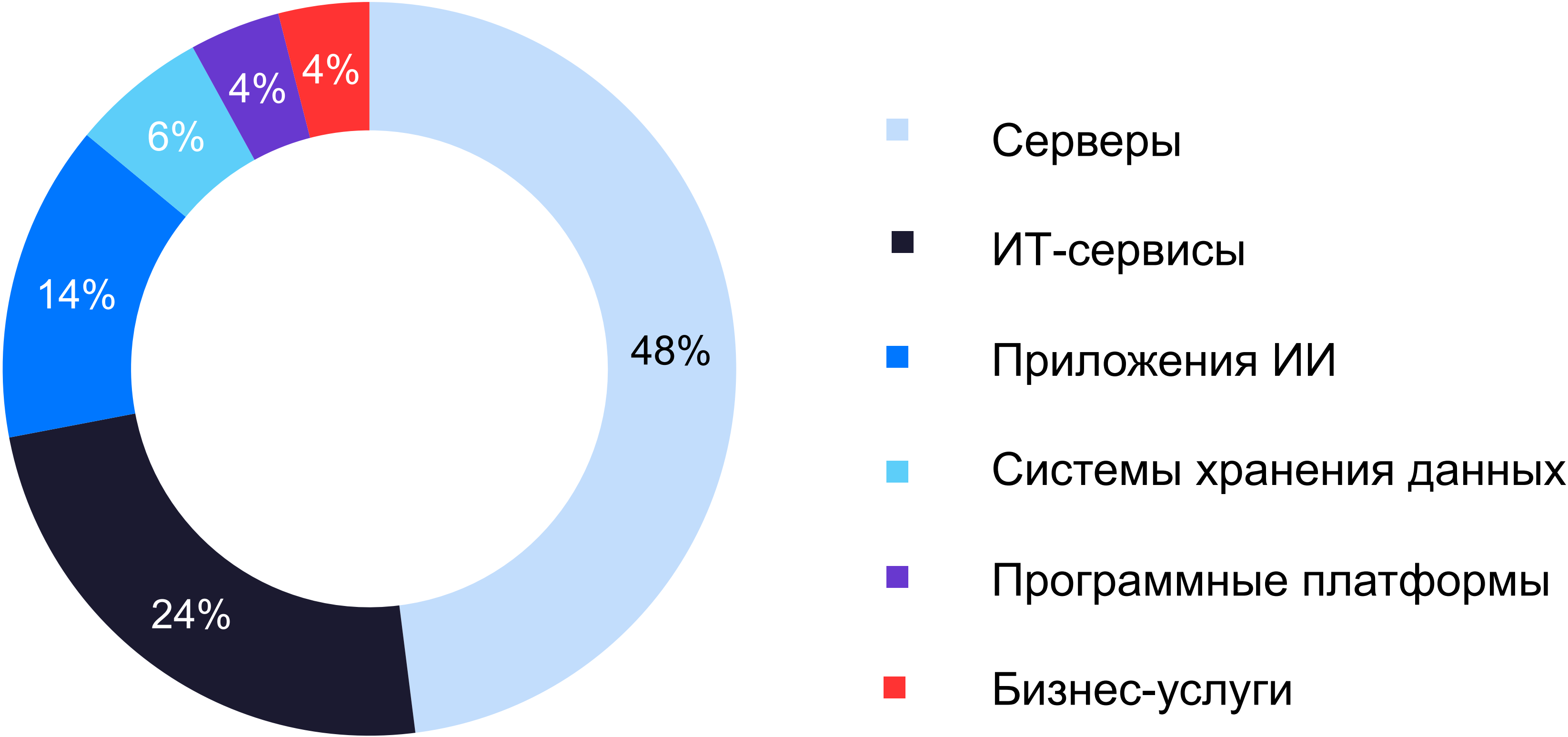
Бизнес в облаках: практические выгоды облаков



Machine Learning

Алёна Дробышевская

Статьи расходов на ML-рынке России

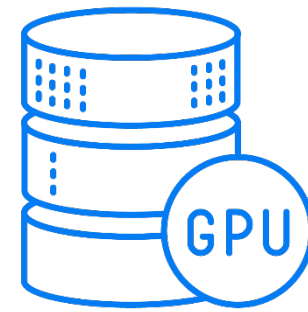


Какие задачи позволяет решить Облако?

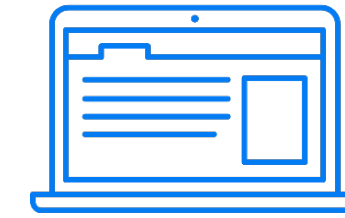
Любые сценарии для оптимизации текущей инфраструктуры и создания новых цифровых сервисов



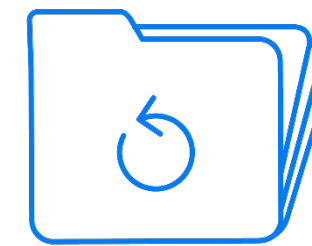
Построение отказоустойчивых веб-приложений



High Performance вычисления (в том числе на базе GPU)



Инфраструктура для виртуальных рабочих мест



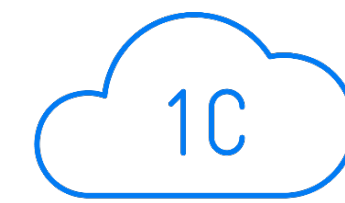
Хранение архивов данных и резервных копий



Test/Dev (в том числе на базе k8s)



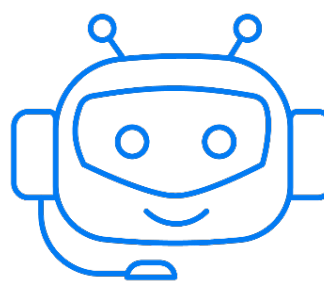
Хранение, обработка, анализ и визуализация данных



1C в Облаке



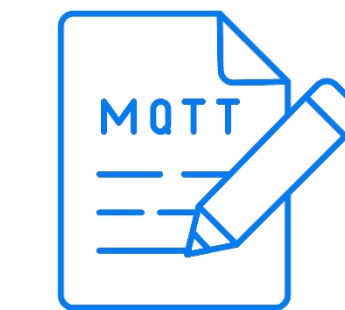
Аналитика логов



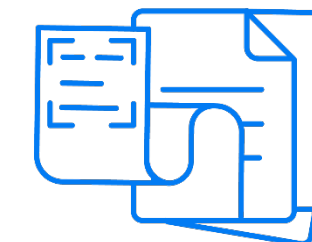
Автоматизация колл-центра



Автоматический перевод текстов



Создание IoT-решений



Оцифровка бумажных архивов

Платформа Яндекс.Облако

SaaS

Партнерские приложения и сервисов

PaaS

Управление данными
и аналитика

Инструменты разработки

Сервисы машинного
обучения

IaaS

Идентификация
и безопасность

Виртуальные машины
и контейнеры

Объектное и блочное
хранилища

Сеть и доставка
контента

Платформа Яндекс.Облако



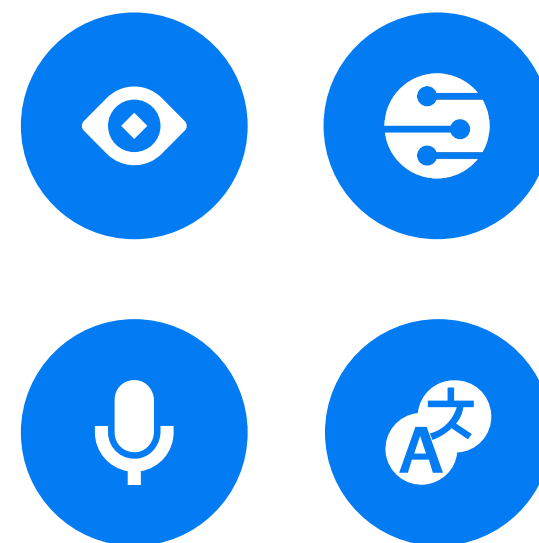
Data Analytics & Management

Сервисы обработки
и управления данными



Cloud-Native Development

Сервисы для создания
облачных приложений



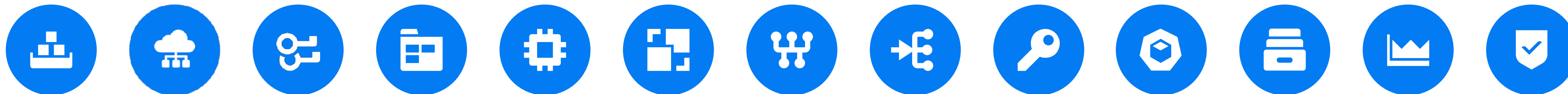
ML & AI

Сервисы машинного
обучения и ИИ



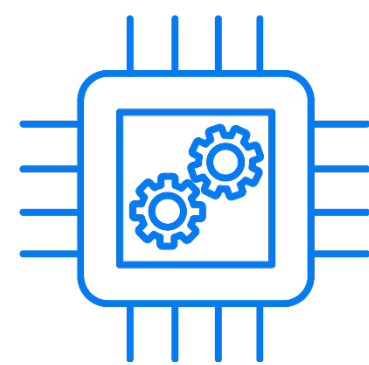
Yandex Cloud Marketplace

Маркетплейс облачных
решений



Hyperscaling Infrastructure

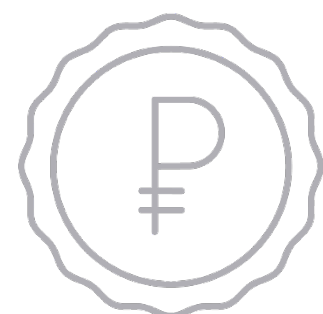
Масштабируемые инфраструктурные сервисы



Уникальный опыт
создания
технологий



Безопасность
и соответствие
требованиям



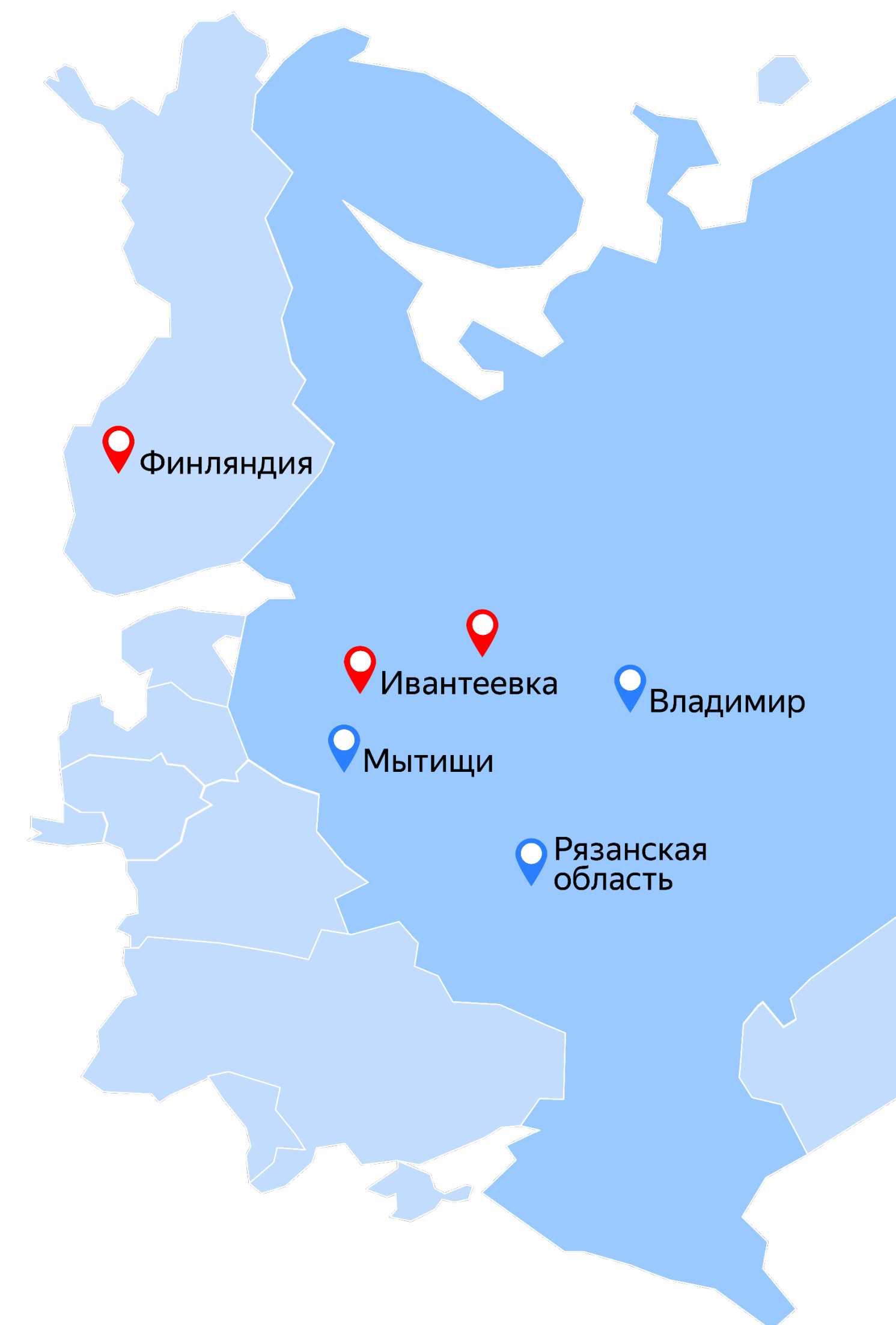
Гибкость
ценообразования

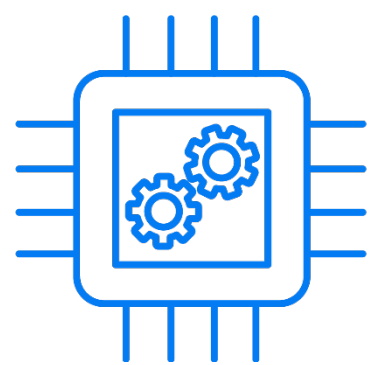


Партнёрская
экосистема

Собственная физическая инфраструктура

- › У Яндекса 5 дата-центров в России и один в Финляндии
- › Яндекс.Облако размещается в трёх зонах доступности
- › Независимая система энергоснабжения у каждого дата-центра
- › ДЦ расположены на расстоянии более 300 км друг от друга
- › Пропускная способность ТБит/сек, благодаря собственной оптоволоконной сети с плотным спектральным мультиплексированием DWDM

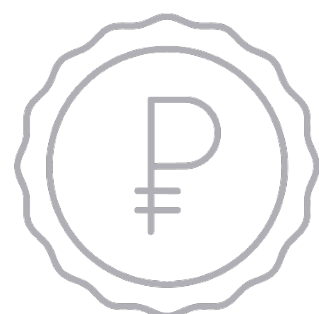




Уникальный опыт
создания
технологий



Безопасность
и соответствие
требованиям



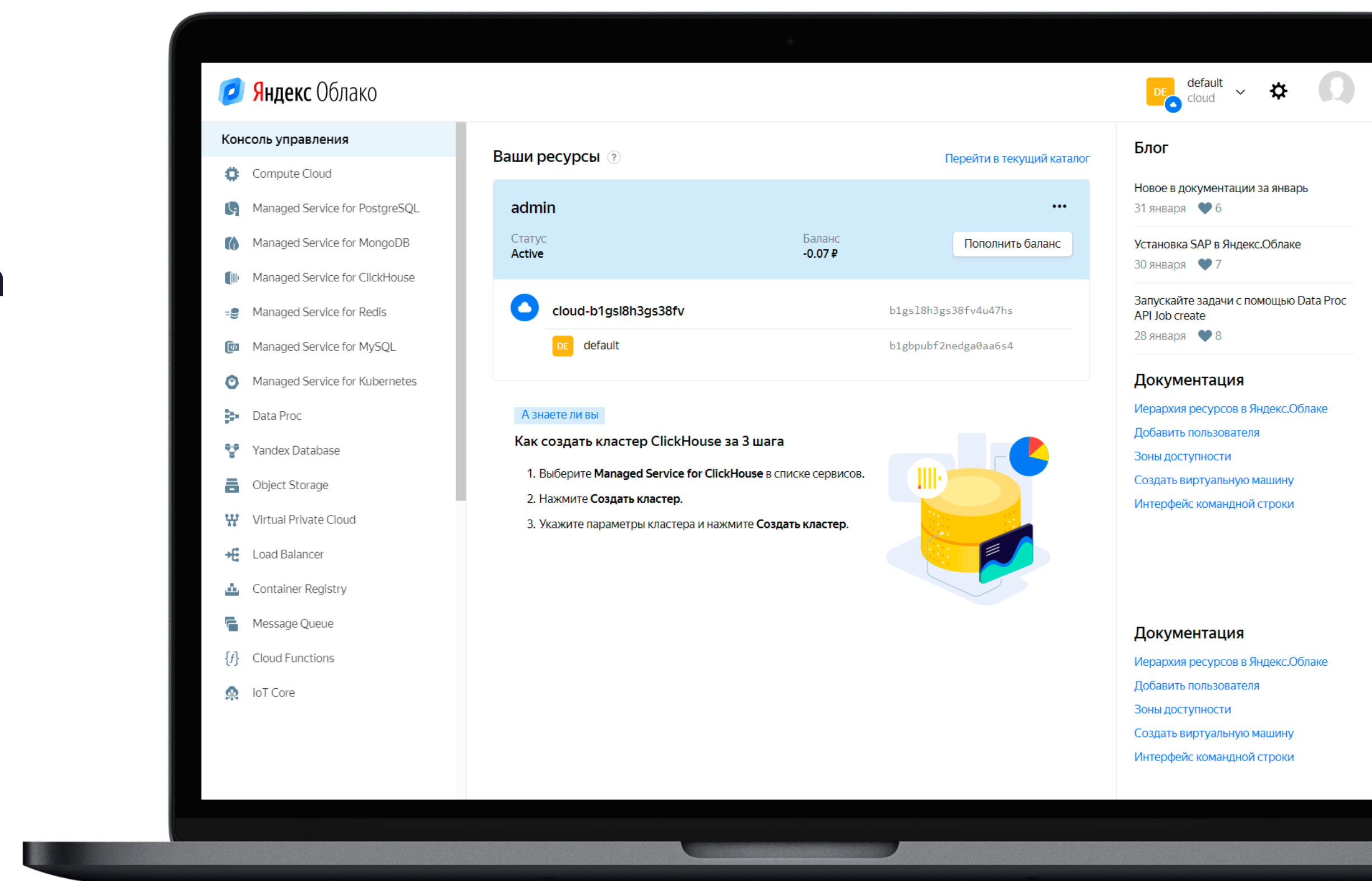
Гибкость
ценообразования

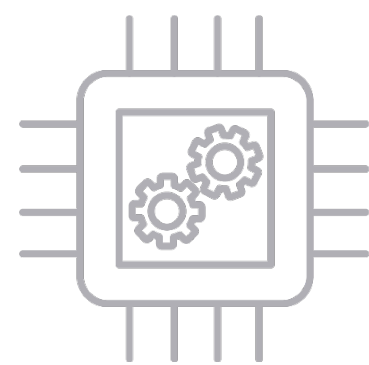


Партнёрская
экосистема

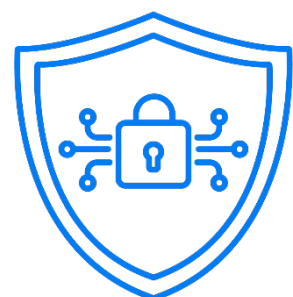
Удобный инструмент для разработчика

- › Полнофункциональная консоль управления
- › Доступ к сервисам через API, CLI, Terraform
- › Русскоязычная документация
- › Несколько уровней технической поддержки

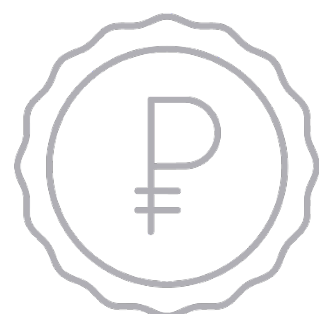




Уникальный опыт
создания
технологий



Безопасность
и соответствие
требованиям

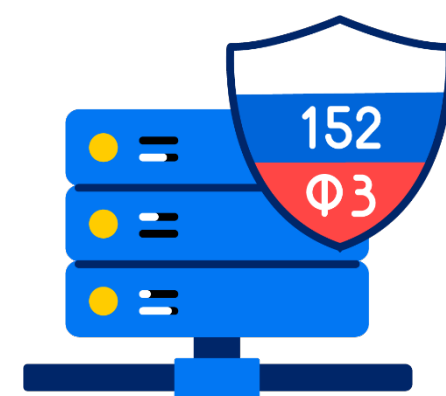


Гибкость
ценообразования



Партнёрская
экосистема

Соответствие законодательным и промышленным требованиям



Соответствует
для РФ



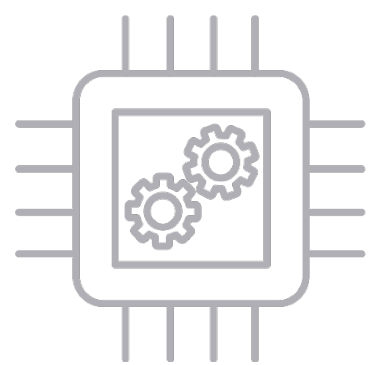
Соответствует
для Евросоюза



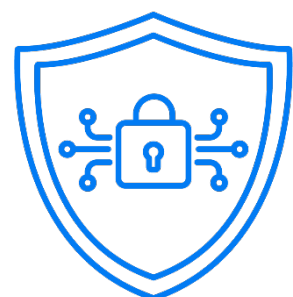
Готово
для ЦОД



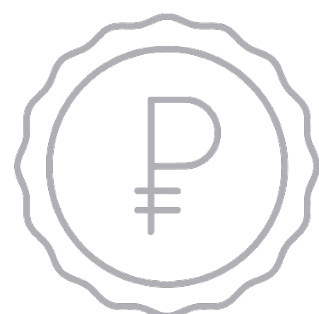
Для облачных
сервисов — Q3 2020



Уникальный опыт
создания
технологий



Безопасность
и соответствие
требованиям

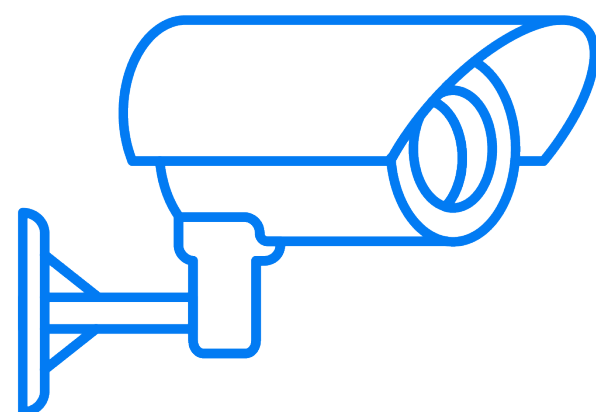


Гибкость
ценообразования



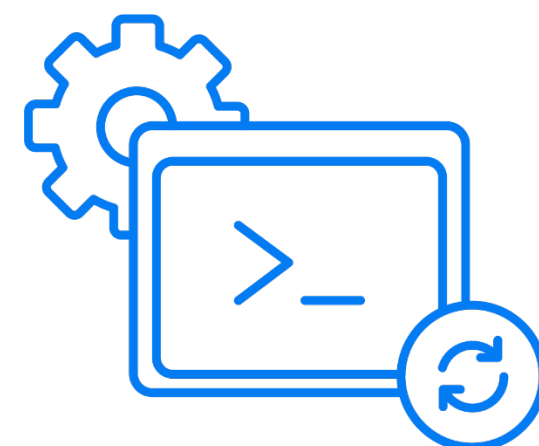
Партнёрская
экосистема

Безопасность на всех этапах создания и эксплуатации Облака



Физическая безопасность

- › Строгие регламенты обслуживания серверов.
- › Все стойки с Облаком под видеонаблюдением.
- › Строгие регламенты уничтожения носителей информации.



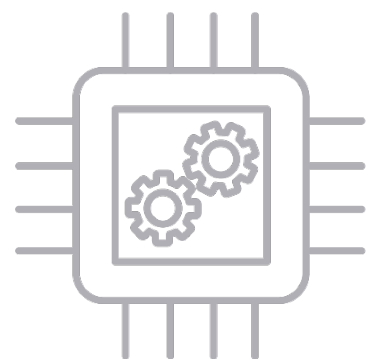
Безопасность разработки

- › Обязательные тренинги по безопасности для разработчиков.
- › Аудит безопасности кода приложений.
- › Регулярные обновления для устранения уязвимостей в сторонних компонентах.



Шифрование данных

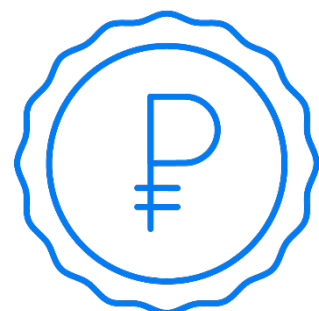
- › Все сервисы облака хранят пользовательские данные в зашифрованном виде
- › Защита данных при передаче по каналам интернет обеспечивается протоколом TLS.



Уникальный опыт
создания
технологий



Безопасность
и соответствие
требованиям

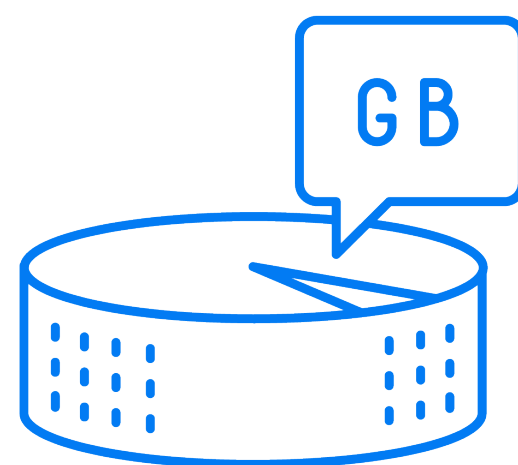


Гибкость
ценообразования



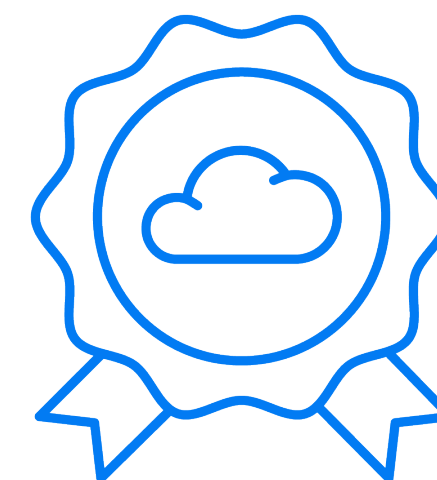
Партнёрская
экосистема

Гибкость ценообразования



Гибкость выбора ресурсов

- › До 64 vCPU
- › До 384 GB RAM
- › До 4 GPU



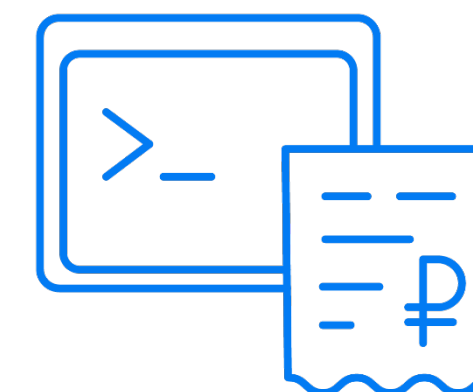
Инвестиции в пилотные проекты

- › Облачные гранты
- › Ресурсы команды архитекторов



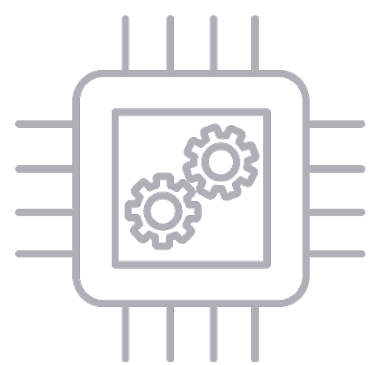
Резервы на 1 или 3 года

- › Compute Cloud
- › Managed Databases



Лёгкое прогнозирование затрат

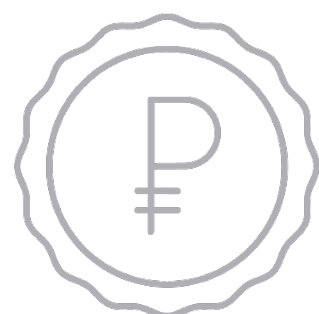
- › Подробная детализация затрат
- › Удобная ресурсная модель (затраты на проект)



Уникальный опыт
создания
технологий



Безопасность
и соответствие
требованиям



Гибкость
ценообразования

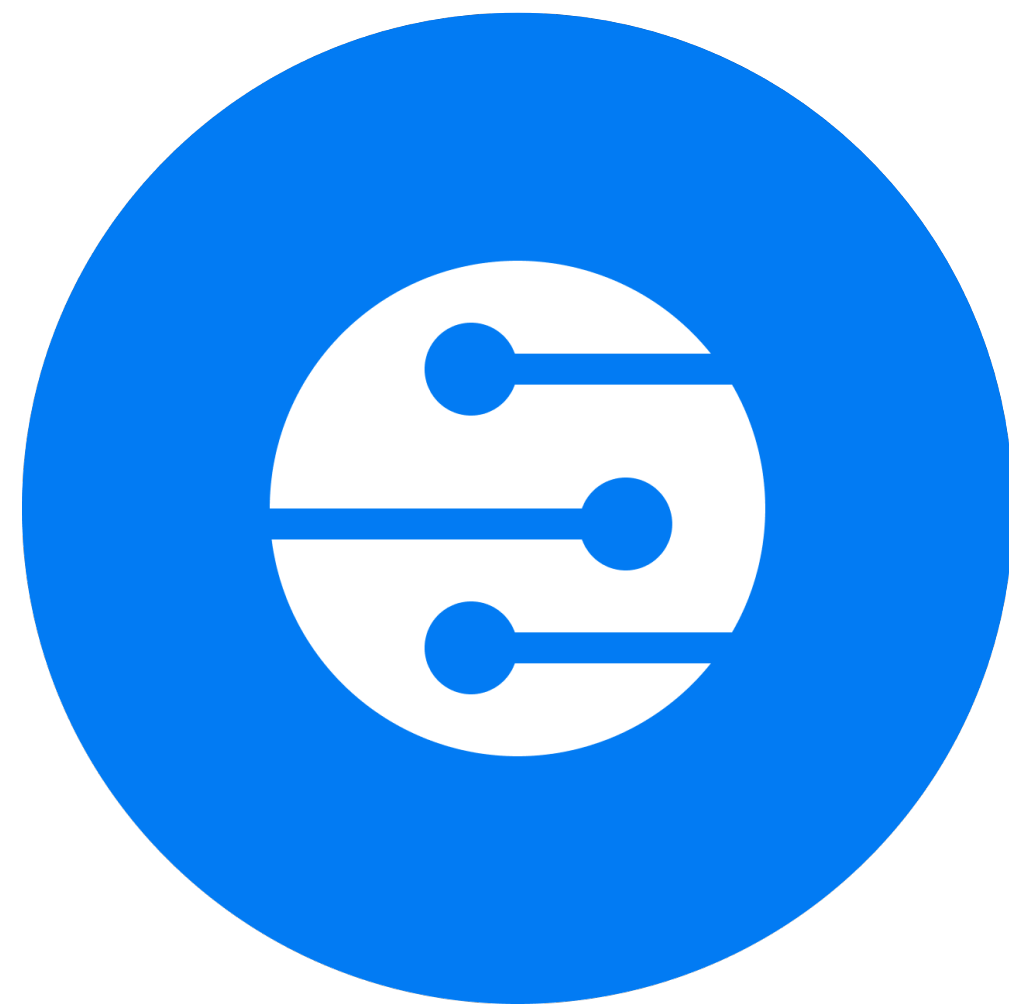


Партнёрская
экосистема

Партнёрская экосистема



Yandex.Cloud для проектов видеоаналитики



Новые начинания проще делать в облаках

- › Облачная инфраструктура позволяет обойтись без закупок специализированного железа для ML
- › Мы предоставляем не только инфраструктуру, но и сервисы, что позволяет экономить на настройке и обслуживании ПО
- › Мы создаём экосистему партнёров для упрощения поиска экспертов в ML
- › Vision позволяет использовать готовое API (например, детектирование лиц)
- › DataSphere существенно понижает порог входа и стоимость прототипирования для проектов, связанных с ML разработкой

Yandex Compute Cloud

BM

- › Два поколения процессоров Intel Broadwell и Cascade Lake
- › **BM — зональный объект**
- › Гибкие параметры выбора размера сервера
- › Burst и Прерываемые BM
- › GPU Tesla V100
- › 512 RAM BM **NEW**
- › Группы размещения **NEW**

Block Storage

- › Реплицируемое сетевое блочное хранилище HDD и SSD
- › Возможность создавать снимки дисков
- › Поддержка образов Linux и Windows
- › Возможность загрузить свой образ
- › **Диск — зональный объект**
- › Образы и снимки — **региональные объекты**

Тарификация

- › Поминутная тарификация для CPU / RAM / HDD
- › Встроенная тарификация Windows
- › On-Demand & BYOL варианты лицензирования

Yandex Object Storage

Функциональность

- › **Региональный** сервис
- › S3 совместимое API (поддержка boto3)
- › Стандартный и холодный типы хранения
- › ACL
- › Lifecycle
- › Возможность использовать собственный домен с HTTPS **NEW**
- › В планах автоматизация создания сертификатов на основе Lets Encrypt **NEW**

Архитектура

- › Данные реплицируются в каждую зону доступности автоматически

Тарификация

- › За объем хранимых данных
- › За запросы (PUT/GET/POST)
- › За исходящий в интернет трафик
- › Standard оптимизирован по цене под большое число запросов
- › Cold оптимизирован по цене под большой объем данных

Yandex Vision

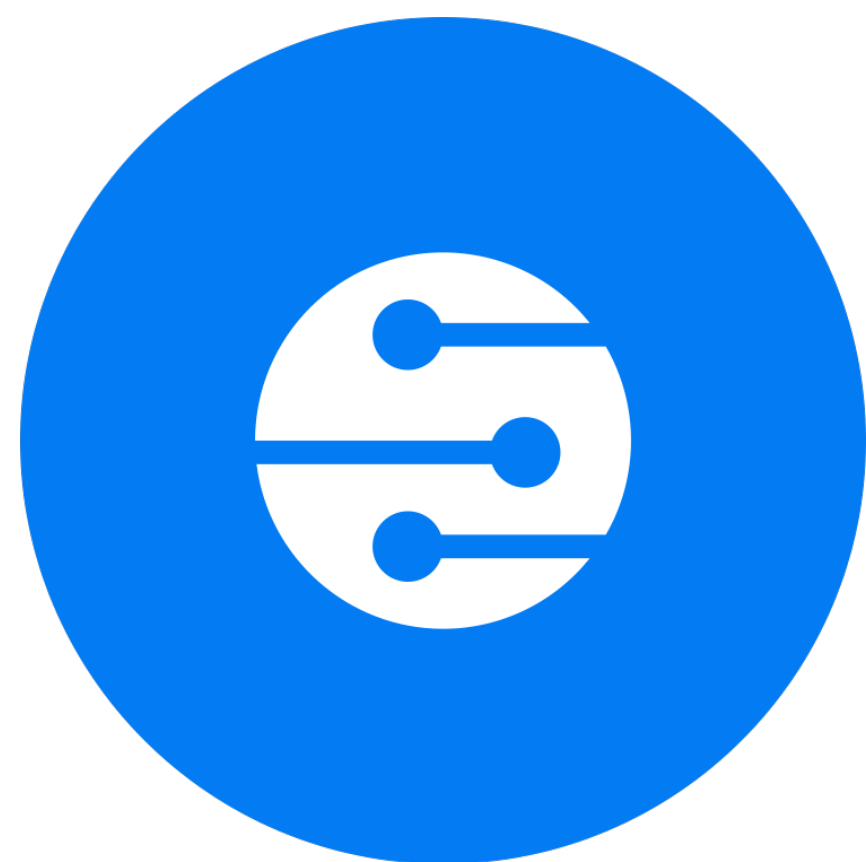


- › Распознавание текста (OCR)
- › Прочие стандартные функции анализа изображения
- › Кастомизируемые классификаторы
- › Технология DeepHD – это технология улучшения изображения основана на генеративно-сопоставительных сетях (англ. GAN — Generative Adversarial Network)

Yandex DataSphere — интегрированная среда для облачной ML-разработки «всё в одном»



Yandex DataSphere – интегрированная среда для разработки ML "все в одном"



- › Serverless computing на мощностях Яндекс.Облака
- › Среда исследования и разработки: привычный интерфейс Jupyter Notebook
- › Бесшовное переключение разных типов ресурсов прямо во время работы, без перезапуска проекта
- › Распределенное исполнение (SPARK, Нирвана)
- › Сквозное версионирование хода разработки
- › Интеграция с сетью пользователя и единое управление правами доступа
- › Inference и регулярное исполнение

Компании, которые нам доверяют



RENAULT



PORSCHE

Yandex Cloud

Спасибо!

Алёна Дробышевская

Руководитель направления по развитию бизнеса ML-сервисов

alenadrob@yandex-team.ru

+7 (985) 761-26-24