

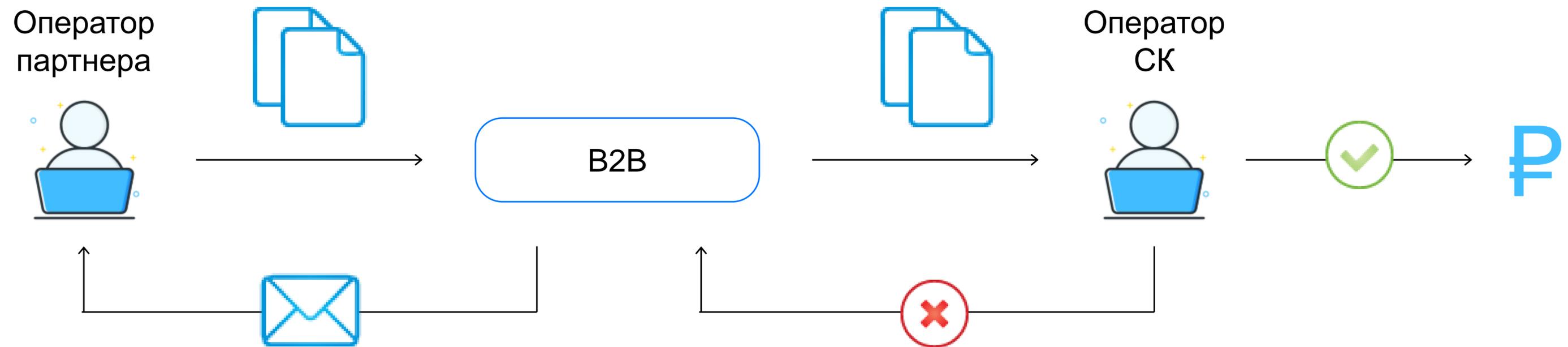


Машинное обучение и Kubernetes на службе документооборота

Задача

- › Ускорить процесс проверки клиентского пакета документов
- › Сократить количество ошибок
- › Уменьшить количество задействованных сотрудников

Исходный процесс



Декомпозиция задачи

Обязательно:

- › Проверить наличие всех типов документов в комплекте
- › Проверить наличие всех страниц документа
- › Проверить наличие всех подписей

Желательно:

- › Разделить документы в пакете по типам в виде отдельных файлов
- › Привести в порядок ориентацию страниц
- › Упорядочить страницы внутри документа

Проблематика

- › Подписи могут быть в разных местах документа
- › Сканы могут быть различных форматов (pdf, jpg, pdf, tiff и т.д.)
- › Комплект может состоять из нескольких файлов разного формата
- › Различное качество изображений
- › Многообразие типов устройств печати
- › Многообразие типов устройств оцифровки изображений

Решение №1 – «В лоб»

Компания А



4 000 000 ₺

Компания В



3 800 000 ₺

Компания С



3 000 000 ₺

Инфраструктура



2 000 000 ₺

Есть ли другой путь?

Viva la QR code!



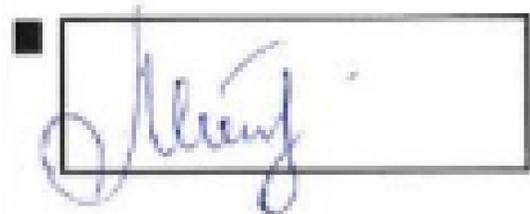
- › Библиотека для работы с документами MS Office без MS Office
- › Библиотека для работы с PDF и изображениями
- › В QR-коде информация о типе документа, подписях и страницах
- › Веб-сервис (.net core)
- › Мы умеем поворачивать, сортировать, разделять
- › Мы умеем проверять комплектность!



Подписи – это котики



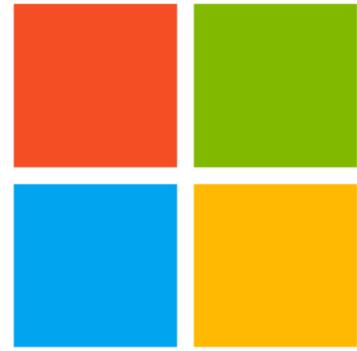
- › Обучающий набор – 2 000 изображений с подписями
- › Компетенции в машинном обучении и Python



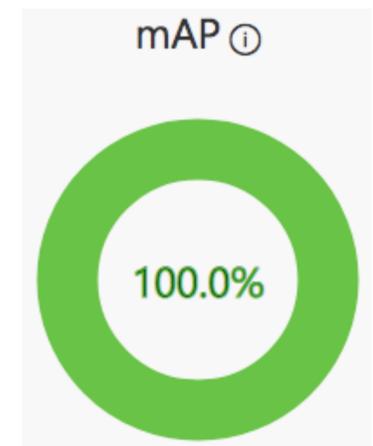
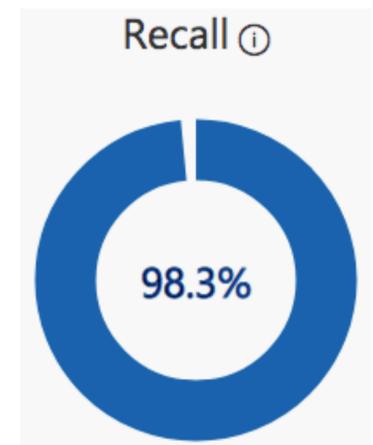
- › Область с маркером на месте подписи
- › Мощности GPU для обучения модели
- › 4 часа на один цикл обучения
- › Результат 59%



А можно ещё проще? - Да!



- › Обучающий набор – 2 000 изображений с подписями
- › Удобный пользовательский интерфейс
- › 0 строк кода для обучения модели
- › Веб-сервис из «коробки»
- › 15 минут на один цикл обучения
- › Результат ~95%

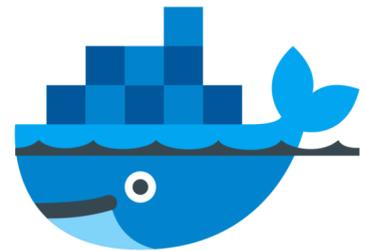


Зачем «железо», когда есть облака



› Три независимых дата-центра

› Автоматическая балансировка нагрузки



› Автоматическое восстановление в случае сбоя

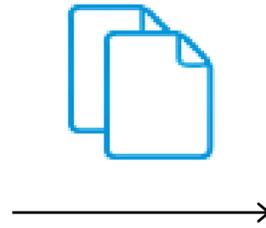
› Возможность «сине-зеленого» развертывания

› Отсутствие необходимости в собственном SRE

› Гибкое управление расходами

Финальный процесс

Оператор
партнера



Загруженные документы:

Название	Статус	Договор	Страницы	Подписи
Памятка (4086404183_МЕМО_2019_11_05_00_16_09).pdf	Корректен	4086404183, 3063013840/4086404183	4 из 4	1 из 1
Договор (3063013840_POL_2019_11_05_00_16_10).pdf	Корректен	4086404183, 3063013840/3063013840	5 из 5	
Договор (4086404183_POL_2019_11_05_00_16_11).pdf	Корректен	4086404183, 3063013840/4086404183	5 из 5	
4086404183_3063013840_Undefined_2019_11_05_00_16_20.pdf	Корректен			

Подробная информация доступна в блоке "Вложения".

Если необходимо, внесите исправления в документы и повторно загрузите в систему.

Если вы абсолютно уверены, что сканы не содержат ошибок, нажмите кнопку "Я уверен в корректности документов".

Я хочу внести исправления

Я уверен в корректности документов

Решение №2 – «Новая эра»

Библиотеки



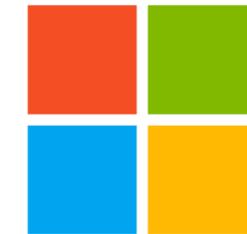
70 000 ₽/год

Разработка



150 часов

Машинное обучение



180 000 ₽/год

Инфраструктура



180 000 ₽/год

Спасибо!

Александр Мольский

Директор ДИТ

Ренессанс Жизнь

Остались вопросы?