

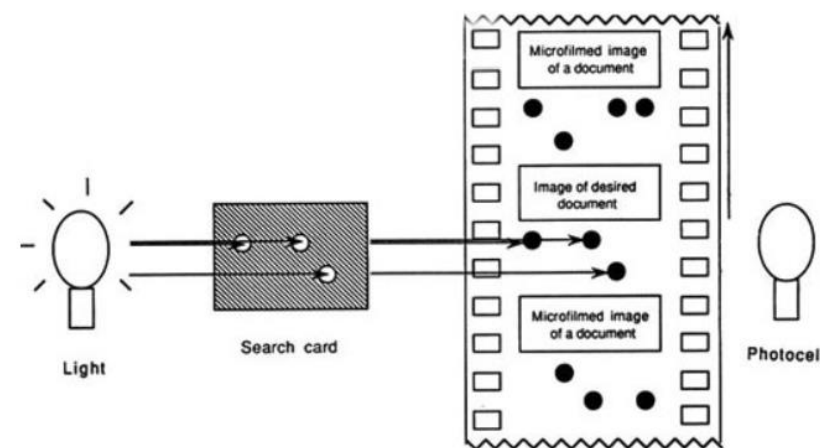


# Как приручить распознавание документов

**Александр Костюченко**  
Директор по разработке продуктов  
ITFB Group



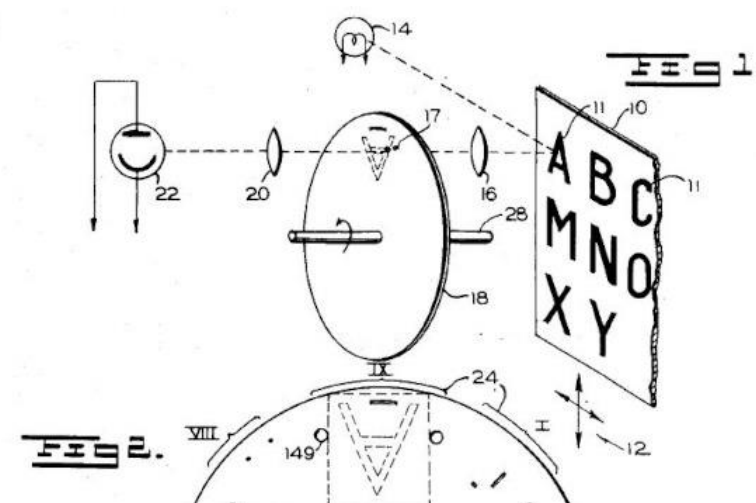
# 100 лет распознавания



## Поиск по микроплёнкам

Эмануэль Голдберг, Германия

**1927**



## Запись на перфокарты

Дэвид Шепард, США

**1953**



## Сканеры Штрих-кодов

Дэвид Колинс, Джорж Лорер, США

**1974**

**1913**

## Оптофон

Фурнье Д'Альба, Ирландия

## Читающая машина

Густав Таушек, Австрия

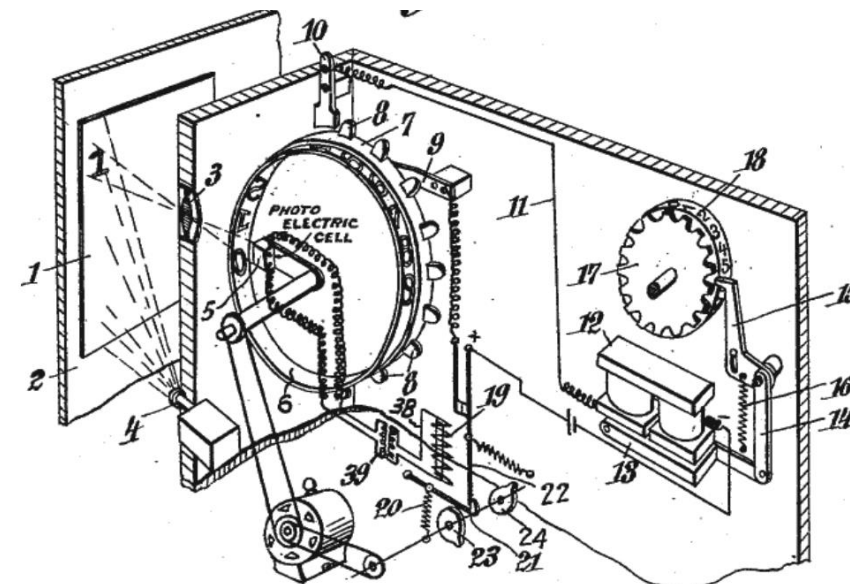
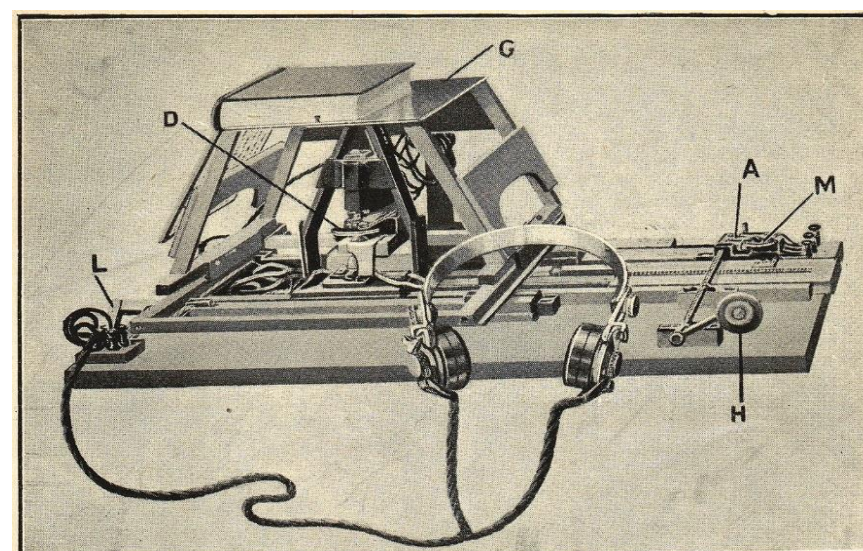
**1955**

## Первое внедрение

Ридерс Дайджест, США

**2000**

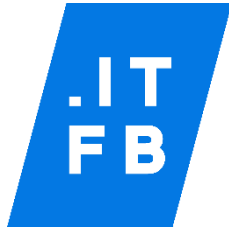
## Оцифровка документов



UNION SUGAR COMPANY  
SUGAR STOCK AND SALES REPORT  
AS OF MARCH 31, 1955

REPORTING COMPANY	1954-55	1953-54	1952-53	1951-52	1950-51
Amstar Sugar Corp., Houston	55,000	60,000	60,000	60,000	60,000
Amstar Sugar Corp., Houston	959,409	1,038,421	1,038,421	1,038,421	1,038,421
Amstar Sugar Corp., Houston	238,559	238,559	238,559	238,559	238,559
Amstar Sugar Corp., Houston	321,909	321,909	321,909	321,909	321,909
Amstar Sugar Corp., Houston	1,096,052	1,237,858	1,237,858	1,237,858	1,237,858
Amstar Sugar Corp., Houston	295,000	295,000	295,000	295,000	295,000
Amstar Sugar Corp., Houston	88,285	88,285	88,285	88,285	88,285
Amstar Sugar Corp., Houston	321,624	321,624	321,624	321,624	321,624
Amstar Sugar Corp., Houston	988,000	988,000	988,000	988,000	988,000
Amstar Sugar Corp., Houston	897,600	897,600	897,600	897,600	897,600
Amstar Sugar Corp., Houston	145,000	145,000	145,000	145,000	145,000
Amstar Sugar Corp., Houston	1,802,600	1,802,600	1,802,600	1,802,600	1,802,600
Amstar Sugar Corp., Houston	392,600	392,600	392,600	392,600	392,600
Amstar Sugar Corp., Houston	282,600	282,600	282,600	282,600	282,600
Amstar Sugar Corp., Houston	125,250	125,250	125,250	125,250	125,250
Amstar Sugar Corp., Houston	81,477	81,477	81,477	81,477	81,477
Amstar Sugar Corp., Houston	31,200	31,200	31,200	31,200	31,200
Amstar Sugar Corp., Houston	34,400	34,400	34,400	34,400	34,400
Amstar Sugar Corp., Houston	-	180,000	-	-	-
TOTAL REPORTING COMPANIES	7,777,024	8,202,000	8,202,000	8,202,000	8,202,000
UNION SUGAR COMPANY	2,021,290	2,021,290	2,021,290	2,021,290	2,021,290
UNION SUGAR COMPANY	8,747,955	10,223,290	10,223,290	10,223,290	10,223,290
UNION SUGAR COMPANY	12,208,365	13,684,580	13,684,580	13,684,580	13,684,580
UNION SUGAR COMPANY	1,121,476	1,121,476	1,121,476	1,121,476	1,121,476
UNION SUGAR COMPANY	1,121,476	1,121,476	1,121,476	1,121,476	1,121,476





# Современные системы распознавания

01

## Предобработка

Улучшения качества оригинальных изображений, коррекция шума

02

## Распознавание

Извлечение текстового слоя из документов на разных языках

03

## Классификация

Определение типа документа с использованием нейронных сетей

04

## Извлечение

Применение шаблонов распознавания и технологий NLP

05

## Верификация

Подтверждение автоматически распознанных атрибутов

06

## Сравнение

Поиск различий в отсканированных и оригинальных документах

07

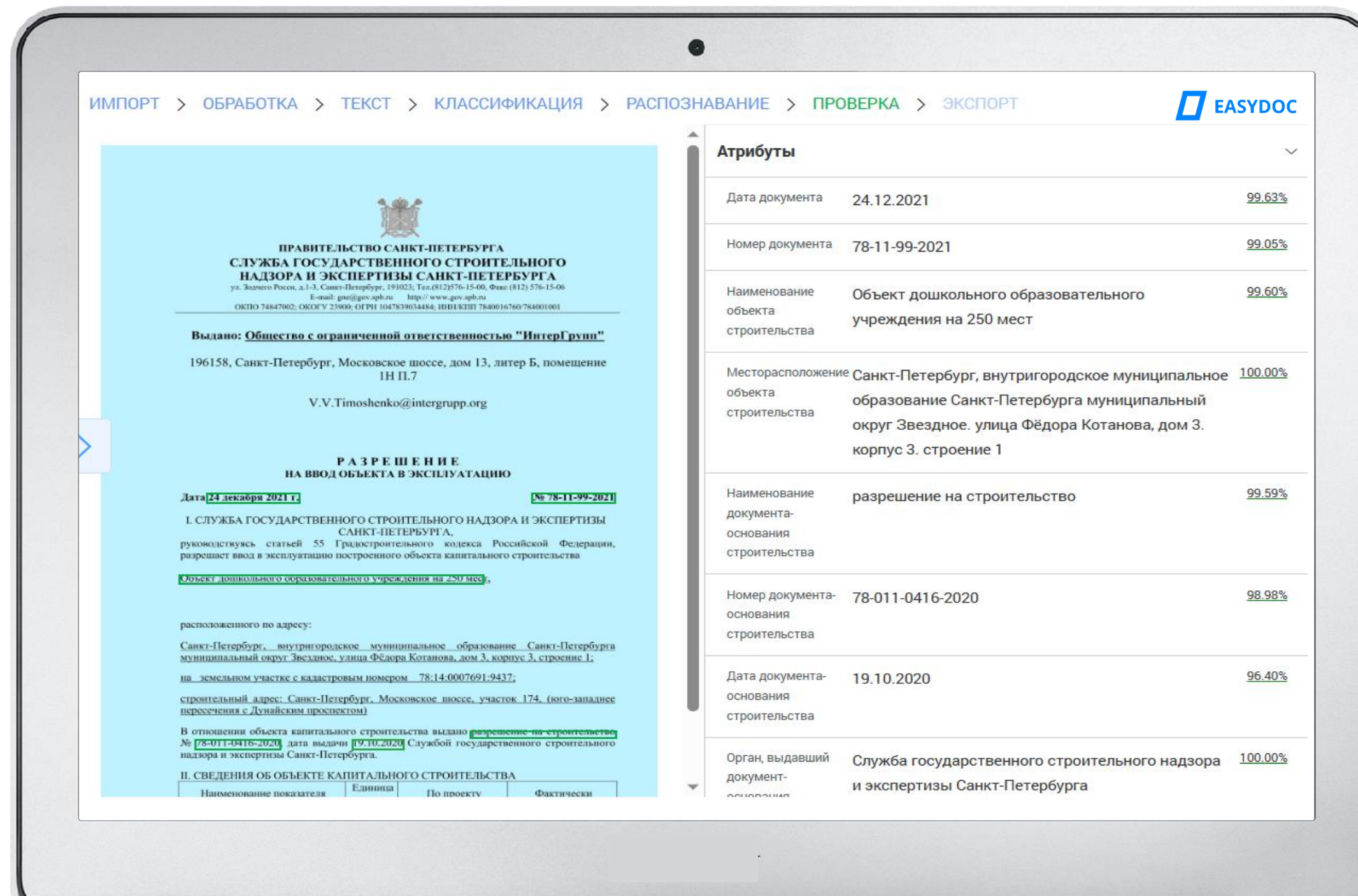
## Кросс-проверки

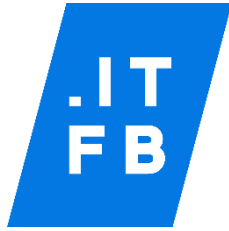
Междокументарные проверки атрибутов разных документов, проверка комплектности

08

## Внешние справочники

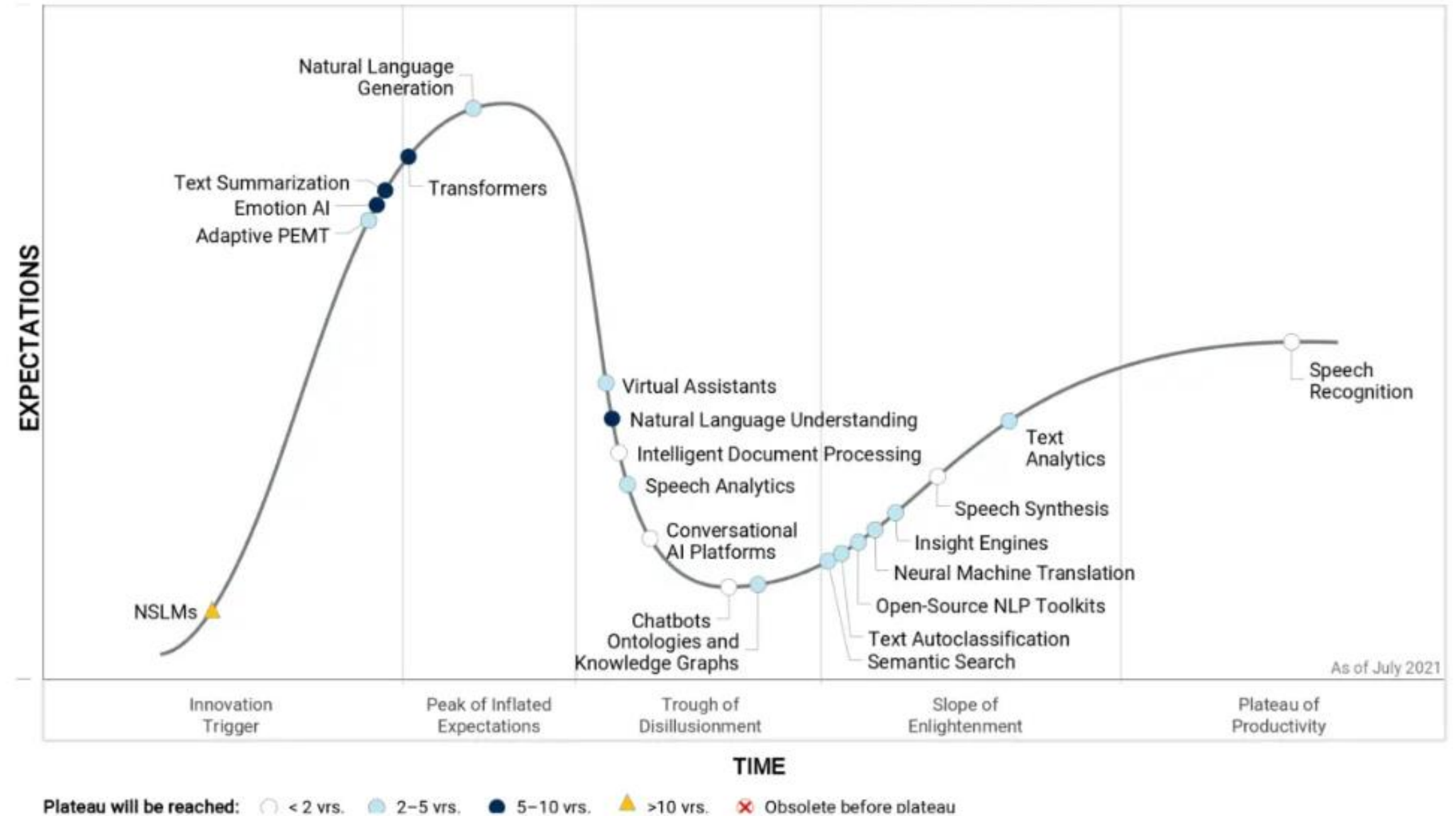
Сравнение атрибутов с внешними справочниками



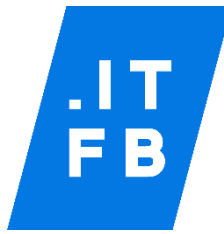


# Направления развития OCR в мире

- Облачные и мобильные решения
- Мультиязычность
- Распознавание не только текстовых образов
- Взаимодействие с RPA
- Использование современных AI подходов



Источник: Gartner



# Направления развития OCS в мире и в России

- Облачные и мобильные решения
- Мультиязычность
- Распознавание не только текстовых образов
- Взаимодействие с RPA
- Использование современных AI подходов
- **Импортозамещение**



# .IT FB ITFB – опытный интегратор и вендор



10 лет

на рынке OCR

100+

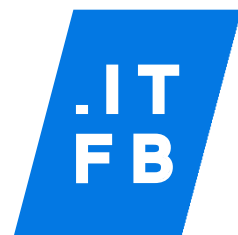
заказчиков, крупнейших компаний РФ

1100+

успешно внедренных ИТ-проектов

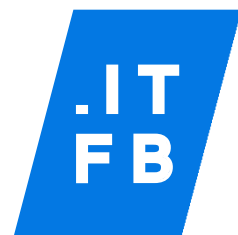
R&D

Распределенная команда разработки продуктов



# Наши решения в сфере импортозамещения

КЛАСС СИСТЕМ	ИМПОРТНЫЙ ПРОДУКТ	ОТЕЧЕСТВЕННОЕ ПО / OPEN SOURCE
CRM / BPM: клиентский опыт и автоматизация бизнес-процессов	Oracle CRM / CRM SalesLogix	ELMA365 / BPM Soft / Bitrix24 / TWORK / SuiteCRM / SugarCRM / Activiti BPM / Camunda BPM / JBOSS Drools
OCR / NLP: системы распознавания текста	ABBYY / Kofax / IBM / EMC	ITFB EasyDoc
BI: бизнес-анализ	Qlik / Tableau / Oracle BI	Visiology / Insight / Форсайт / FineBi
СЭД / ECM: управление электронными документами	Documentum / Opentext	Elma ECM / ITFB Симфония
ERP: управления ресурсами предприятия	SAP	1С



# Наши решения в сфере импортозамещения

КЛАСС СИСТЕМ	ИМПОРТНЫЙ ПРОДУКТ	ОТЕЧЕСТВЕННОЕ ПО / OPEN SOURCE
CRM / BPM: клиентский опыт и автоматизация бизнес-процессов	Oracle CRM / CRM SalesLogix	ELMA365 / BPM Soft / Bitrix24 / TWORK / SuiteCRM / SugarCRM / Activiti BPM / Camunda BPM / JBOSS Drools
OCR / NLP: системы распознавания текста	ABBYY / Kofax / IBM / EMC	<b>ITFB EasyDoc</b>
BI: бизнес-анализ	Qlik / Tableau / Oracle BI	Visiology / Insight / Форсайт / FineBi
СЭД / ECM: управление электронными документами	Documentum / Opentext	Elma ECM / ITFB Симфония
ERP: управления ресурсами предприятия	SAP	1С





# Технологии EasyDoc

   
Java python™  
Open Source

+

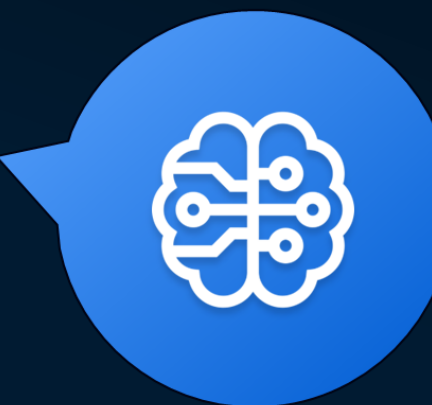
ОСОБЫЕ  
ПОДХОДЫ  
К ОБРАБОТКЕ



COMPUTER VISION



NATURAL LANGUAGES  
PROCESSING



DEEP LEARNING



LARGE LANGUAGE  
MODELS



# Области применения EasyDoc

## Открытие счета

Паспорта, заявления, учредительные документы

## Обработка обращений

Классификация обращений на сайте, в мессенджерах, в СЭД

## Судебные документы

Судебные приказы, решения

## Первичка

Счёт-Фактуры, Торг12, УПД, Акты, КС-2

## Сопровождение договоров

Договоры, уставы, соглашения

x1000r/min

## Контроль проектирования

Чертежи, Спецификации, Схемы, РВЭ

## Кадровый учет

Паспорта, ИНН, СНИЛС, трудовые книжки, 2-НДФЛ, диплом, военный билет, резюме

## Зарубежные партнеры

Сертификаты, инвойсы, накладные

## Кредитование

Паспорта, заявления, 2-НДФЛ, учредительные документы



# КЕЙС «Распознавание бухгалтерской первички»

## Задача

- Автоматизировать передачу данных из филиалов в единый центр обслуживания
- Повысить уровень точности и достоверности финансовых данных

## Решение

- Разработаны шаблоны распознавания для 170 типов документов
- Настроены многоуровневые проверки распознанных данных между документами

СЧЕТ-ФАКТУРА № 57 « 03 » июля 2017 года (1)  
 ИСПРАВЛЕНИЕ № — от « — » — (1a)

Продавец ООО «Продавец» (2)  
 с Ярославская обл., г. Переславль-Залесский, ул. Менделеевская, д. 90 (2a)  
 ИНН/КПП продавца 7608123456/760801001 (2б)  
 Руководитель и его адрес — (3)  
 Грузополучатель и его адрес — (4)  
 платежно-расчетному документу № 701 от 29 июня 2017 года (5)  
 Покупатель ООО «Покупатель» (6)  
 Адрес г. Москва, ул. Шота Руставели, д. 12 (6a)  
 ИНН/КПП покупателя 7715123456/111501001 (6б)  
 Валюта: наименование, код российский рубль, 643 (7)

Наименование товара (описание выполненных работ, оказанных услуг), имущественного права	Единица измерения		Количество (объем)	Цена (тариф) за единицу измерения	Стоимость товаров (работ, услуг), имущественных прав без налога — всего	В том числе сумма акциза	Налоговая ставка	Сумма налога, предъявляемая покупателю	Стоимость товаров (работ, услуг), имущественных прав с налогом — всего
	к о д	условное обозначение (национальное)							
1	2	2a	3	4	5	6	7	8	9
Предоплата по договору № 45 от 02.06.2017	643	руб.	—	—	—	—	18/118	18 000,00	118 000,00
Всего к оплате						X		18 000,00	118 000,00

Руководитель организации или иное уполномоченное лицо Иванов (подпись) Иванов И.И. (ф.и.о.) Главный бухгалтер или иное уполномоченное лицо Петрова (подпись)

Атрибуты	
Номер документа	57
Продавец	ООО "Продавец"
ИНН продавца	7608123456
КПП продавца	760801001
Грузополучатель и его адрес	—
Покупатель	ООО "Покупатель"
ИНН покупателя	7715123456
КПП покупателя	111501001
Всего к оплате (сумма налога)	18 000,00
Всего к оплате (стоимость)	118 000,00

# КЕЙС «Распознавание бухгалтерской первички»

## Задача

- Автоматизировать передачу данных из филиалов в единый центр обслуживания
- Повысить уровень точности и достоверности финансовых данных

## Решение

- Разработаны шаблоны распознавания для 170 типов документов
- Настроены многоуровневые проверки распознанных данных между документами

## Результат

**1850+**

атрибутов  
автоматически  
извлекает система

в **5** раз

сокращение времени  
на извлечение  
информации из пакета  
документов

**85-95** %

среднее качество  
распознавания  
документов

**98-100** %

достоверность  
финансовых  
данных

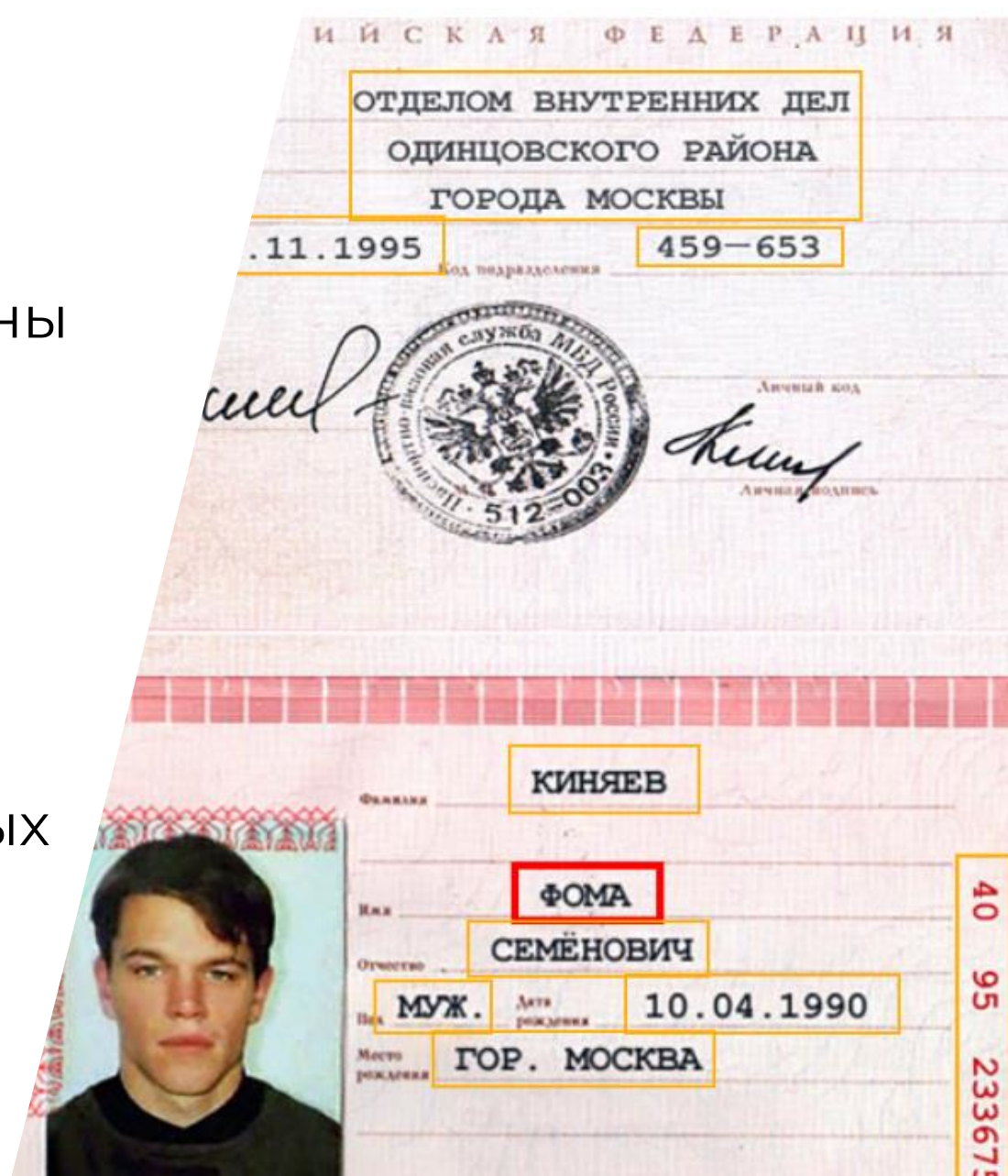
# КЕЙС «Открытие счёта»

## Задача

- Сократить время обработки клиентских документов
- Автоматизировать обработку документов

## Решение

- Разработаны шаблоны распознавания для 80 типов документов
- Настроены многоуровневые проверки распознанных данных между документами



Атрибуты	
Серия	4095
Номер	233675
Фамилия	КИНЯЕВ
Имя	ФОМА
Отчество	СЕМЕНОВИЧ
Ф.И.О.	КИНЯЕВ ФОМА СЕМЕНОВИЧ
Пол	М
Дата рождения	10.04.1990
Дата выдачи	11.11.1995
Код подразделения	459-653
Кем выдан	ОТДЕЛОМ ВНУТРЕННИХ ДЕЛ ОДИНЦОВСКОГО РАЙОНА ГОРОДА МОСКВЫ

# КЕЙС «Открытие счёта»

## Задача

- Сократить время обработки клиентских документов
- Автоматизировать обработку документов

## Решение

- Разработаны шаблоны распознавания для 80 типов документов
- Настроены многоуровневые проверки распознанных данных между документами

## Результат

**550+**

атрибутов  
автоматически  
извлекает система

**85-95%**

среднее качество  
распознавания  
документов

**5-10** минут

сокращение времени  
на извлечение  
информации из пакета  
документов

**98-100%**

достоверность  
информации  
о клиенте

# КЕЙС «Распознавание судебных документов»

## Задача

- Сократить время и затраты на обработку судебных документов
- Увеличить скорость взыскания дебиторской задолженности

## Решение

- Разработаны шаблоны распознавания судебных приказов и решений
- Настроена интеграция с учетной системой

ФИКАЦИЯ > РАСПОЗНАВАНИЕ > ПРОВЕРКА > ЭКСПОРТ

судок № 4 города Златоуста Челябинской области,  
Челябинская область, город Златоуст, ул.Бушуева, д.4  
http://zlatoustchel4.chel.msudrf.ru  
E-mail: zlatoust4@mirovoy-sud74.ru

Именем Российской Федерации  
**СУДЕБНЫЙ ПРИКАЗ**

Дело № 2-654/2088

Судья судебного участка №4 г.Златоуста Челябинской области Лидова Л.А., рассмотрев заявление публичного акционерного общества «Банк Уралсиб» о вынесении судебного приказа о взыскании с Цветкова Павла задолженности по кредитному договору, руководствуясь ст.ст. 98, 121, 122 Гражданского процессуального кодекса Российской Федерации, ст.ст. 309, 310, 807-811, 819 Гражданского процессуального кодекса Российской Федерации,

**ПОСТАНОВИЛ:**

Взыскать с Цветкова Павла Игоревича в пользу публичного акционерного общества «Банк Уралсиб» задолженность по кредитному договору №8505-N93/02088 от 29 апреля 2018 года за период с 30 апреля 2018 года по 10 января 2022 года в размере **80 461 рубль 59 копеек**, из которых:

- 5 896 рублей 79 копеек – задолженность по кредиту;
- 3 204 рубля 71 копейка – задолженность по уплате процентов;
- 11 360 рублей 09 копеек – пени,

Место рождения должника	Златоуст Челябинской области	
Дата рождения должника	19.06.1912	
СНИЛС должника	67986090666	
Номер паспорта должника	888436	
Серия паспорта должника	7517	
Дата выдачи паспорта должника	01.03.2001	
Реквизиты взыскателя	ПАО Банк Уралсиб	
Номер кредитного договора	8505-93/02088	
Дата заключения кредитного договора	29.04.2013	
Круглая печать суда	нет	
Предмет исполнения	Установил	Сумма
Итоговая сумма		
Предмет исполнения	Решил	Сумма
задолженность по кредитному договору	100.00%	80 461 рубль 59 копеек
задолженность по кредиту	100.00%	5 896 рубль 79 копеек
задолженность по уплате процентов	100.00%	3 204 рубль 71 копейка
пени	100.00%	11 360 рубль 09 копеек
расходы по уплате государственной пошлины	100.00%	

# КЕЙС «Распознавание судебных документов»

## Задача

- Сократить время и затраты на обработку судебных документов
- Увеличить скорость взыскания дебиторской задолженности

## Решение

- Разработаны шаблоны распознавания судебных приказов и решений
- Настроена интеграция с учетной системой

## Результат

на **40%**  
сократилось время  
на подготовку к суду

**0%**  
ошибок  
ручного ввода

на **80%**  
снизились затраты  
на обработку  
документов



квалифицированные  
сотрудники заняты  
профильной работой



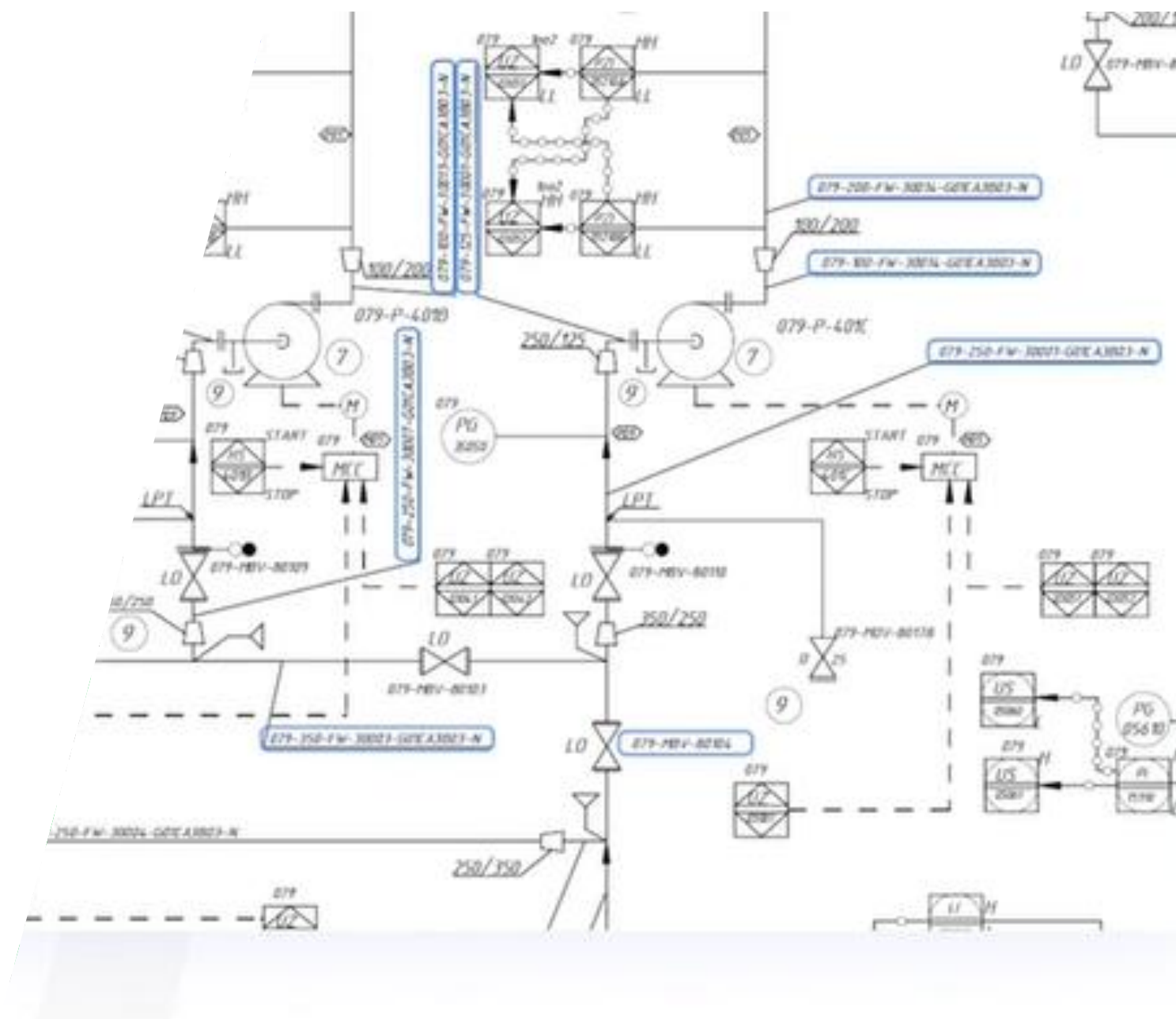
# КЕЙС «Распознавание элементов чертежей»

## Задача

- Найти на чертежах все определённые элементы и сравнить со спецификацией
- Извлечь значения тегов, соответствующие элементам схемы
- Распознать спецификацию с чертежа и сравнить с проектной документацией

## Решение

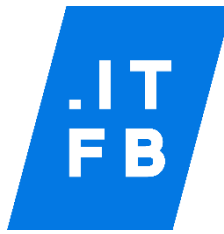
- Разработана модель Computer Vision для детекции элементов схем
- Реализовано извлечение характеристик элементов



- 079-250-FW-30002-GO
- 079-125-FW-30007-GO
- 079-250-FW-30007-GO
- 079-100-FW-30013-GO
- 079-125-FW-30001-GO
- 079-200-FW-30014-GO
- 079-100-FW-30014-GO
- 079-250-FW-30001-GO
- 079-MBV-80104
- 079-350-FW-30003-GO

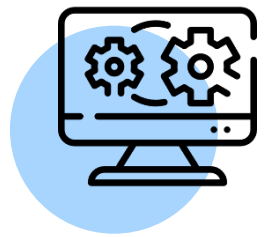


Если вы хотите  
сами обучать



# EasyDoc 3.0

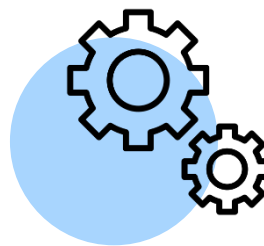
## платформа интеллектуальной обработки документов



Архитектура не зависит от импортного ПО, поддерживает российские ОС



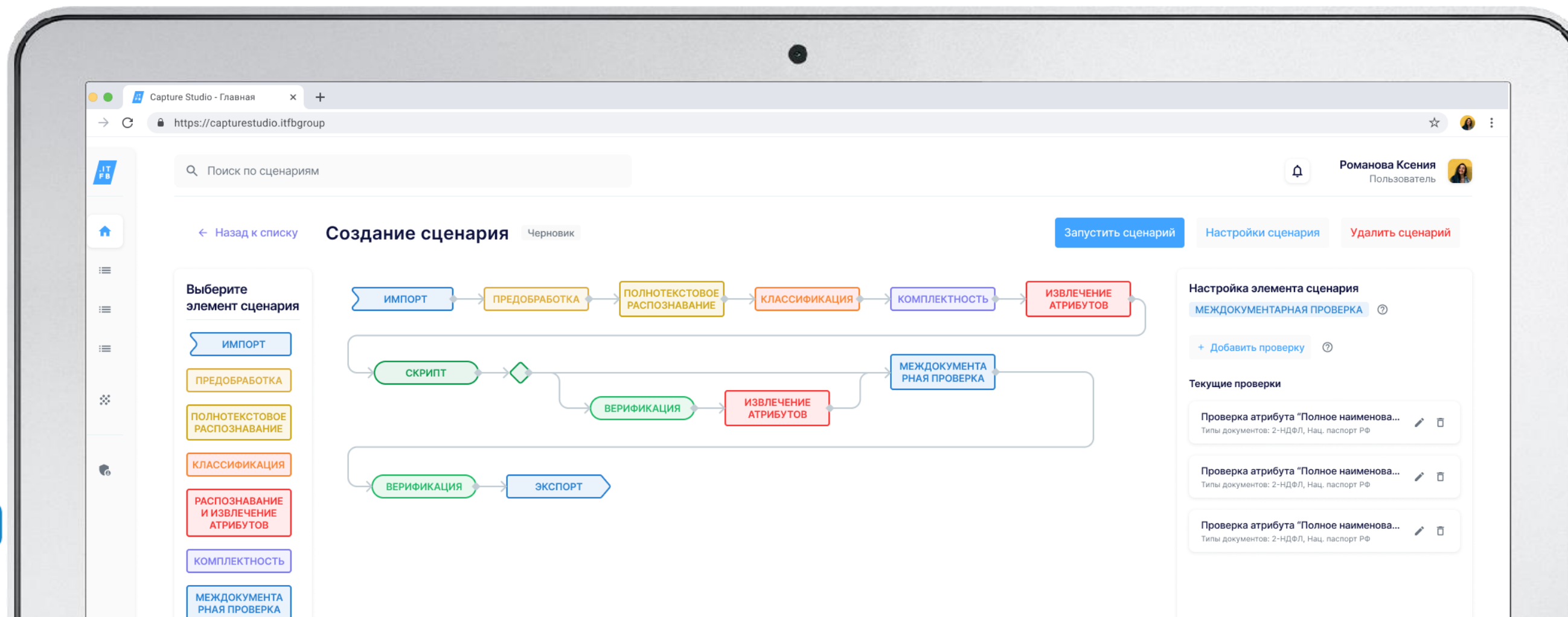
Настройка AI без навыков программирования



Масштабируемая распределенная обработка документов



Удобные веб-приложения не требуют локальной инсталляции и обновления





**ИСПОЛЬЗУЙТЕ  
ПРАВИЛЬНЫЕ  
ОТЕЧЕСТВЕННЫЕ  
ТЕХНОЛОГИИ**

**Александр Костюченко**  
Директор по разработке продуктов  
ITFB Group

