

**Smart Document Engine** — российская система автоматического анализа и распознавания документов для **десктопных, серверных и мобильных платформ**. Smart Document Engine – это программный инструмент для бизнеса и разработчиков, обеспечивающий высокоточное и высокоскоростное распознавание текстовых и иных данных первичных, деловых, уставных, бухгалтерских, налоговых, нотариальных, юридических, страховых и банковских документов, а также типовых анкет и форм строгой отчетности.

Научноемкая технология оптического распознавания символов GreenOCR®, разработанная нашими учеными, позволяет точно распознавать текст любых документов на более чем 100 языках, включая кириллицу, латиницу, арабский, персидский, урду, японский, китайский, корейский и другие. GreenOCR® обеспечивает высокую точность распознавания печатного текста (OCR), рукопечатного и рукописного заполнения полей (ICR), а также распознавание меток и чекбоксов (OMR).

**Smart Document Engine** является уникальным инструментом, позволяющими гибко интегрировать распознавание документов в текущую деятельность компании. Программа легко и удобно встраивается в уже существующие бизнес-процессы, не нарушая привычного опыта использования. Решение доступно для **интеграции с различными информационными системами, например, ECM, CRM, RPA, 1C, АБС и т.д.**

#### **Для чего нужно распознавание документов Smart Document Engine?**

Программный продукт Smart Document Engine предназначен для автоматизации обработки и оцифровки бумажных документов, документооборота и оптимизации процессов массового ввода документов.

Smart Document Engine решает даже на обычных мобильных телефонах задачи оптического распознавания деловых документов, которые раньше были доступны только высокопроизводительным серверам.

#### **Конфиденциальность и безопасность**

Программное обеспечение Smart Document Engine:

- НЕ использует код Open Source и иностранные программные компоненты, обеспечивая технологический суверенитет
- НЕ передает личные данные ваших клиентов на обработку в сторонние онлайн-сервисы и/или третьим лицам для ручного ввода
- НЕ сохраняет данные: вся обработка ведется в локальной оперативной памяти устройства – 100% on-premise
- НЕ требует сетевого соединения
- НЕ использует технологии: HITL(human in the loop), CrowdSourcing, Mechanical Turk и аналоги, удаленных, облачных или внешних верификаторов

Важным аспектом **Smart Document Engine** является возможность создания специализированных решений для конкретных заказчиков, позволяющих производить автоматическую обработку, классификацию, распознавание и анализ бумажных и отсканированных pdf документов и форм любой сложности. Система понимает все востребованные типы документов в виде графических файлов (jpg, tiff, png, и др). Распознавание документов в PDF и других форматах выполняется после конвертации в один из поддерживаемых форматов растровых данных.

# Основные возможности системы распознавания Smart Document Engine



## Классификация и распознавание жестких форм

Жесткими формами называют документы, разные экземпляры которых совпадают “на просвет” при удалении реквизитов. Классическими примерами жестких форм являются больничный лист или бланк ЕГЭ. Система распознавания документов Smart Document Engine включает в себя state-of-the-art алгоритмы быстрой локализации и типизации жестких форм на сканах, фотографиях и в видеопотоке, что позволяет осуществлять мгновенное наведение и разбор таких документов, а также верификацию их типа.

## Классификация и распознавание гибких форм

Гибкими формами называют документы, элементы и реквизиты которых могут менять взаимное расположение на бланке. Классическими примерами гибких форм являются широко распространенные налоговые и бухгалтерские документы, такие как справка о доходах физического лица установленного образца (ранее 2-НДФЛ), бухгалтерский баланс предприятия, отчет о финансовых результатах, платежное поручение и т.п. При помощи системы Smart Document Engine можно быстро идентифицировать документы подобного рода, локализовать значимые реквизиты, провести их распознавание и анализ.



## Анализ документов произвольного вида

Система распознавания документов Smart Document Engine позволяет создавать решения для распознавания и анализа документов произвольного вида, такие как доверенности, согласия, договора. При помощи Smart Document Engine вы можете производить классификацию таких документов, извлекать текстовые или иные реквизиты, верифицировать наличие подписей и печатей, и др.

### Поддержка сканов, фотографий и видео

Smart Document Engine может распознавать документы и формы как на изображениях, полученных с планшетных и протяжных сканеров, так и с фото и видео, полученных при помощи смартфонов, планшетов и других мобильных устройств.



2350	(113)	(86)
2400	2 027	1 880

Продолжение на странице 15

**СТРАНИЦА 15**

↓

Копировать (подпись) Копировать Марфа Наумова (расшифровка подписи)

января 2020 г.

**СТРАНИЦА 16**

### Анализ многостраничных документов

Smart Document Engine позволяет классифицировать и распознавать как отдельные изображения документов, анкет и форм, так и документы с множеством страниц, либо последовательности страниц содержащих несколько документов. Обработка последовательности страниц позволяет упростить [процесс потокового сканирования](#), производить сортировку потока и проверять наличие необходимых страниц и документов (проверка комплектности).

### Распознавание текстовых реквизитов

Стек технологий, реализованный в платформе Smart Document Engine, позволяет быстро и точно распознавать однострочные и многострочные текстовые поля на более чем 100 языках мира, извлекать реквизиты из сплошного текста, машинописные и рукопечатные поля в разграфке, и многое другое.

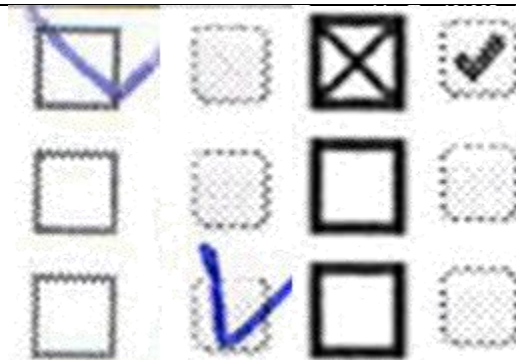
3 1 0 7 2 0 2 0

Общество с ограниченной ответственностью «Сладкая  
Организация Забегаловка»

Дата рождения 10.02.1970

Имя Ефросиния

0 7 0 4 1 9 8 2



### Распознавание меток и чекбоксов

Система Smart Document Engine позволяет точно и надежно распознавать метки и чекбоксы, выполненные как в печатном варианте, так и ручкой.

## Поиск и распознавание печатей

Система Smart Document Engine оснащена модулями быстрого детектирования, локализации и классификации печатей. Помимо поиска и верификации наличия печати, платформа позволяет реализовать распознавание отдельных текстовых компонентов печатей.



АКТИВ					Код показателя
1					2
I. ВНЕОБОРОТНЫЕ АКТИВЫ					
Нематериальные активы					110

Месяц	Код дохода	Сумма дохода	Код вычета	Сумма вычета
1	2000	90 948.00		
2	2000	90 948.00		
3	2000	30 260.00		
4	2000	6 564.00		
5	2000	96 564.00		
6	2000	96 500.00		
7	2000	96 564.00		

## Распознавание таблиц

Для полноценной обработки бухгалтерских, налоговых, банковских и других документов в системе Smart Document Engine реализован поиск и распознавание табличных данных. Поддерживаются как таблицы с жесткой структурой, реляционные и не реляционные, таблицы с переменным количеством столбцов, и структурированные наборы реквизитов с разграфкой.

## Контроль заполнения документа

Система Smart Document Engine оснащена модулями быстрого детектирования, локализации и классификации печатей. Помимо поиска и верификации наличия печати, платформа позволяет реализовать распознавание отдельных текстовых компонентов печатей.

Налоговый орган (1 - налоговый орган, 2 - уполномоченный предприниматель) \_\_\_\_\_

Ильин Иван Иванович 2017 ☒

Наличие документа, подтверждающего полномочия предпринимателя ☐

\* Отметка устанавливается при наличии

Приложение № 1  
к Постановлению Правительства РФ от 15.10.2014 № 1303-ПП

СПРАВКА О ДОХОДАХ И СУММАХ НАЛОГОВ ФИЗИЧЕСКОГО ЛИЦА  
№ 2017, от 15.03.2017

1. Данные о налогоплательщике  
Код по ОКТМО: 4000000 ☒ ИНН: 7707 000 00-00 ☒ КПП: 7707003034 ☒  
Полный адрес: 125040, г. Москва, ул. Мухоморова, д. 1, стр. 1  
Коды идентификации налогоплательщика: ИНН: 7707003034  
Организационно-правовая форма: ООО

2. Данные о физическом лице - налогоплательщике  
Имя: Иван Иванович Фамилия: Ильин  
Дата рождения: 15.10.1974 Пол: Мужской  
Сторона регистрации: 1 Страна: Россия  
Вид документа: декларация о доходах Серия и номер документа: 01.04.2017

компетентному государственному органу (или иной надлежущей уполномочен  
Стандарта Автоматического Обмена Налоговой Информацией или Common Reportin  
компетентными органами иностранных государств (территорий)», а также  
в Сосвете Женераль С. А. публичную компанию с ограниченной ответственности  
и уничтожение персональных данных Клиента) - основному акционеру Банка?  
Я ПОДТВЕРЖДАЮ, ЧТО ВСЕ ВНЕСЕННЫЕ СВЕДЕНИЯ ЯВЛЯЮТСЯ КОРРЕКТНЫМИ, ДОС

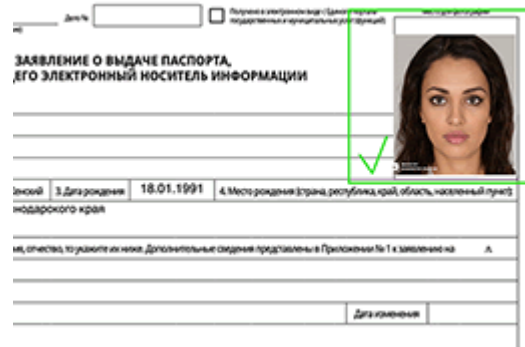
государственному органу (или иной надлежущей уполномочен  
Стандарта Автоматического Обмена Налоговой Информацией или Common Reportin  
компетентными органами иностранных государств (территорий)», а также  
в Сосвете С. А. публичную компанию с ограниченной ответственности  
и уничтожение персональных данных Клиента) - основному акционеру Банка?  
Я ПОДТВЕРЖДАЮ. ☒

## Проверка наличия ключевых слов и словосочетаний

В процессах анализа сложно структурированных документов для идентификации значимых фраз и параграфов, а также для обнаружения некорректных изменений в тексте печатных документов система Smart Document Engine позволяет проверять наличие ключевых слов, словосочетаний, фраз, и контролировать целостность контрольных параграфов.

### Контроль логотипов и прочих графических элементов

Smart Document Engine позволяет детектировать, локализовать и контролировать наличие графических элементов, таких как логотипы компаний, а также важных графических полей документов, таких как фотография, вклеенная в анкету.



Дата:  ☐ Получен в электронном виде (Данное заявление передано в электронном виде)

**ЗАЯВЛЕНИЕ О ВЫДАЧЕ ПАСПОРТА, ЕГО ЭЛЕКТРОННЫЙ НОСИТЕЛЬ ИНФОРМАЦИИ**

Искать: 3. Дата рождения: 18.01.1991 4. Место рождения (страна, республика, край, область, населенный пункт)

м, отчество, то укажите их ниже. Дополнительные сведения представлены в Приложении № 1 к заявлению на

Дата изменения:

Месяц	Код дохода	Сумма дохода	Код вычета	Сумма вычета
6	4800	1 300.00		
7	2000	107 731.53		
7	2002	2 366.96		
7	4800	1 300.00		
8	2000	25 691.64	105 000.03	
8	4800	33 036.00		
9	2000	76 213.19		
9	4800	54 887.54		
9	2012	65 143.83		
9	4800	1 300.00		
10	2000	52 409.77		
10	4800	1 300.00		

### Контроль отсутствия помарок, исправлений, и прочих пометок

Система Smart Document Engine позволяет детектировать, локализовать, и распознавать помарки, зачеркивания, исправления и прочие пометки на документе с целью контроля его достоверности и извлечения дополнительной информации.

### Проверка наличия и распознавание рукописных пометок и подписей

Smart Document Engine позволяет извлекать рукописные поля и подписи, распознавать цифровые поля, выполненные от руки в свободной форме, а также верифицировать наличие рукописных отметок и подписей.

Заведующая лабораторией И.А. Айболит /И.А. Айболит/

я, Иванов Иван Иванович



### Контроль цветности документа

Система Smart Document Engine позволяет определять как цветность документа в целом, так и его отдельных элементов (печатей и подписи), что позволяет распознать черно-белую копию документа даже при повторном сканировании или фотографировании.

### Анализ атрибутов текстовых полей

Помимо распознавания текстовых полей, Smart Document Engine позволяет анализировать атрибуты полей и отдельных текстовых символов, такие как характеристики шрифта (наличие засечек, курсив, жирность), оценки однородности и монотонности и др.

