

# ABBYU®

## NLP в задачах финансовых организаций.

Контроль затрат, управление  
рисками и привлечение клиентов

---

Ольга Морозова

*Директор по работе с финансовым  
сектором, ABBYU Россия*

Дмитрий Черноус

*Заместитель директора по консалтингу  
и развитию бизнеса, ABBYU Россия*



# Спикеры вебинара:



**Ольга Морозова**

*Директор по работе с  
финансовым сектором,  
ABBYY Россия*



**Дмитрий Черноус**

*Заместитель директора  
по консалтингу и развитию  
бизнеса, ABBYY Россия*

# Программа вебинара

- ИИ в финансовой отрасли: прогнозы рынка и технологические тренды.
- Возможности NLP в задачах банков.
- Успешные кейсы применения NLP-технологий в компаниях финансового сектора.
- Демонстрация работы технологии на примере обработки документов от федеральных органов исполнительной власти (ФОИВ).



# Лидеры цифровой трансформации в России

95%

Финансовый сектор

62%

Ритейл и FMCG

59%

Телекоммуникации

*\* Источник: Digital Leader, PwC, IDC, КРОК "Тренды и технологии 2030", 2020*

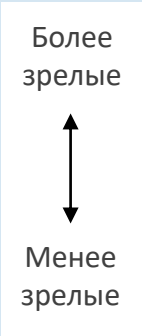
*Опрос проводили среди 160 сотрудников из ведущих компаний по цифровой трансформации: Intel, PwC, Pure Storage, VMWare, КРОК и др.*

Во время пандемии цифровые каналы стали основным способом внутренней и внешней коммуникации с клиентами, поэтому вопрос технологической трансформации в финансовом секторе стоит особенно остро



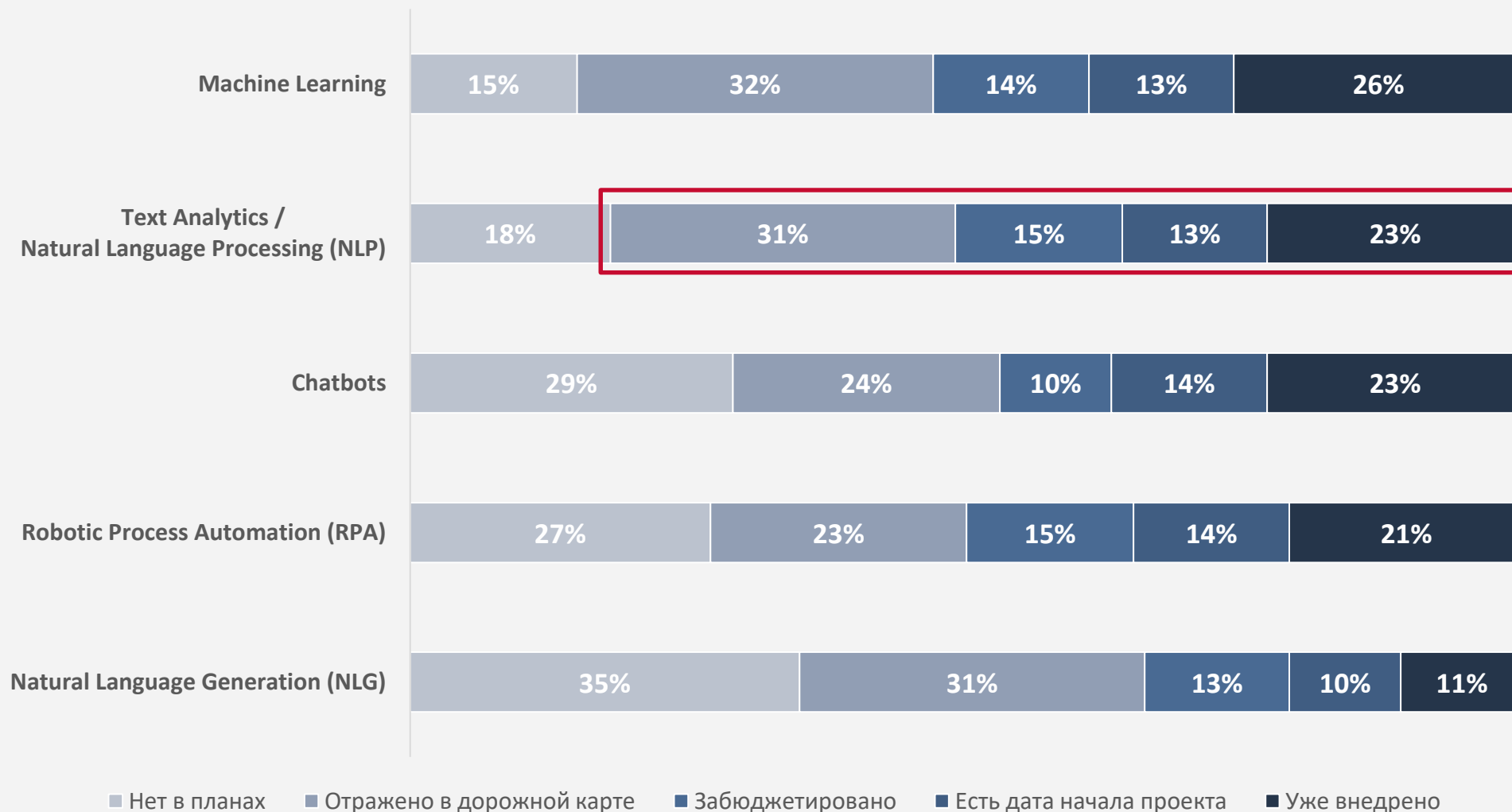
# Банки могут сэкономить \$447 млрд к 2023 году, если будут использовать ИИ

Использование ИИ в финансовом секторе				
Канал	Фронт-офис		Мидл-офис	Бэк-офис
Возможная экономия	 \$199 МЛРД		 \$217 МЛРД	 \$31 МЛРД
Ключевые бизнес-сценарии	Коммуникации с клиентами		Борьба с мошенничеством и управление рисками	Кредитный андеррайтинг
	Биометрические технологии	Персональные предложения	Борьба с отмыванием денег / Процедура "Know Your Customer"	Смарт-контракты



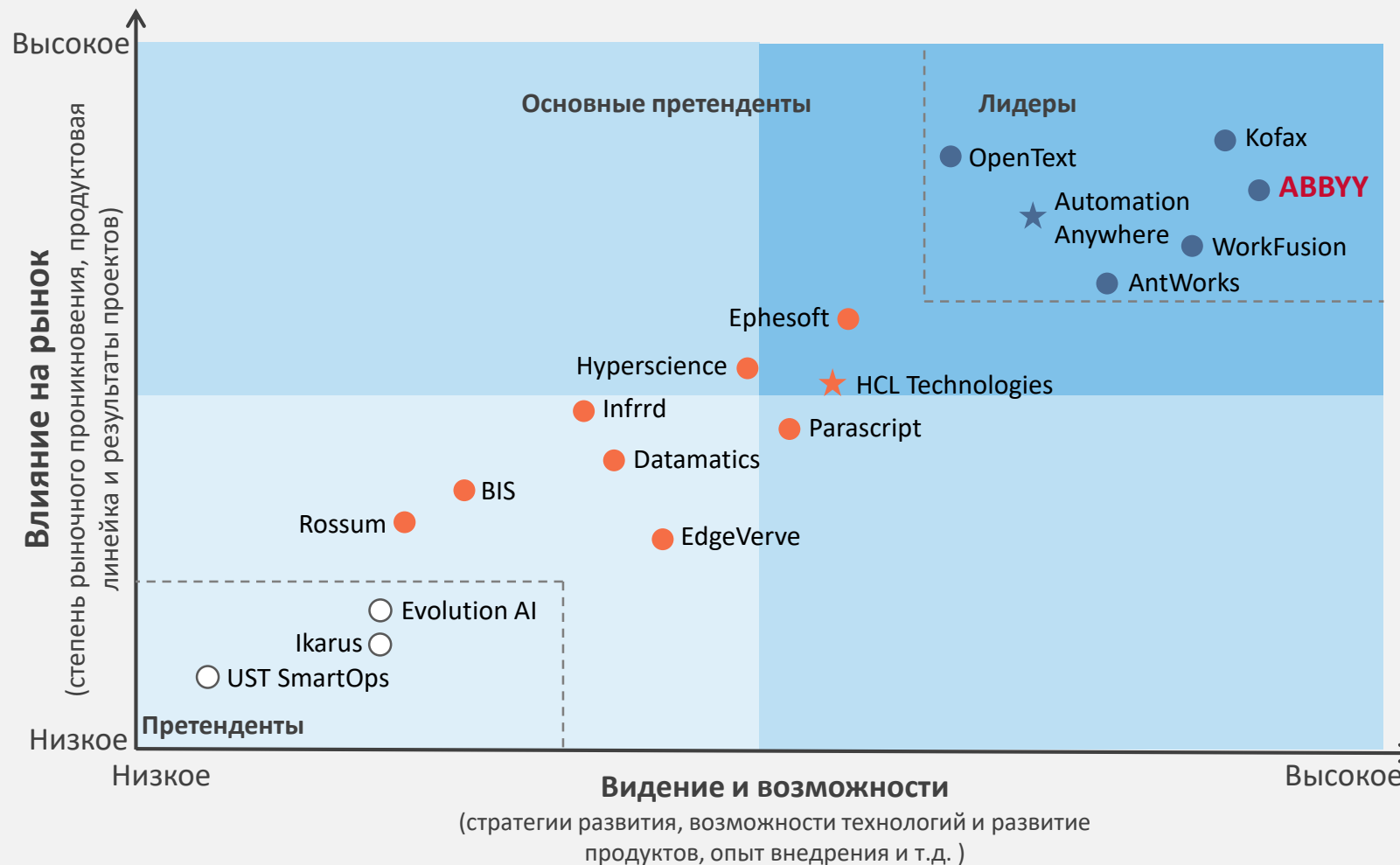
\* Источник: Business Insider "Artificial Intelligence in Banking", 2019

# Востребованность ИИ-технологий в финансовом секторе



82% банков внедрили или планируют внедрить NLP в свои бизнес процессы

# Рынок интеллектуальной обработки документов



- Лидеры
- Основные претенденты
- Претенденты
- ☆ Звезды



# Основные сферы применения технологий искусственного интеллекта на сегодня



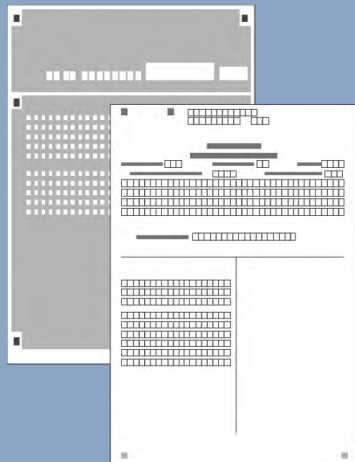


Что такое NLP технологии и какие задачи они решают?

# Технологии NLP расширяют возможности извлечения данных

## Извлечение данных с помощью NLP

### Структурированные документы



- счета-фактуры, платежные поручения, анкеты, ЕГРЮЛ, выписки

### Слабоструктурированные документы



- заявления, счета, документы по форме банка, документы удостоверяющие личность

### Неструктурированные документы



- письма, запросы гос. органов, договоры, уставы, доверенности, приказы

Традиционные технологии извлечения данных

## Классификация

- Кластеризация
- Тегирование и идентификация темы
- Оценка тональности
- Сегментация

## NER (Named Entity Recognition)

- Извлечение именованных сущностей:
- Люди
  - Организации
  - Адреса

## ML+NLP

Извлечение любых пользовательских сущностей с помощью ML с учителем на размеченных документах

## Определение фактов

- Можно извлекать такие факты, как: Мистер X был назначен генеральным директором
- Президент X посетил страну Y
  - Компания X была оштрафована на \$1млн

Извлечение данных из неструктурированных документов & текстовая аналитика



# Успешное извлечение данных с помощью ИИ



$$= \text{NLP} + \text{ML}$$

Извлечение данных из неструктурированных документов

Предоставляет возможности, помогающие сократить количество обучающей выборки и повысить качество

Размеченная обучающая выборка



- “Быстрый старт” для начала работы требуется всего 100 вместо 10 000 образцов документов;
- Комплексные варианты использования: извлечение фактов



ПОНИМАНИЕ  
СМЫСЛА  
ДОКУМЕНТОВ

## СЦЕНАРИИ



Обработка неструктурированных документов:  
уставы, протоколы, контракты, акты, доверенности,  
резюме и т.д.



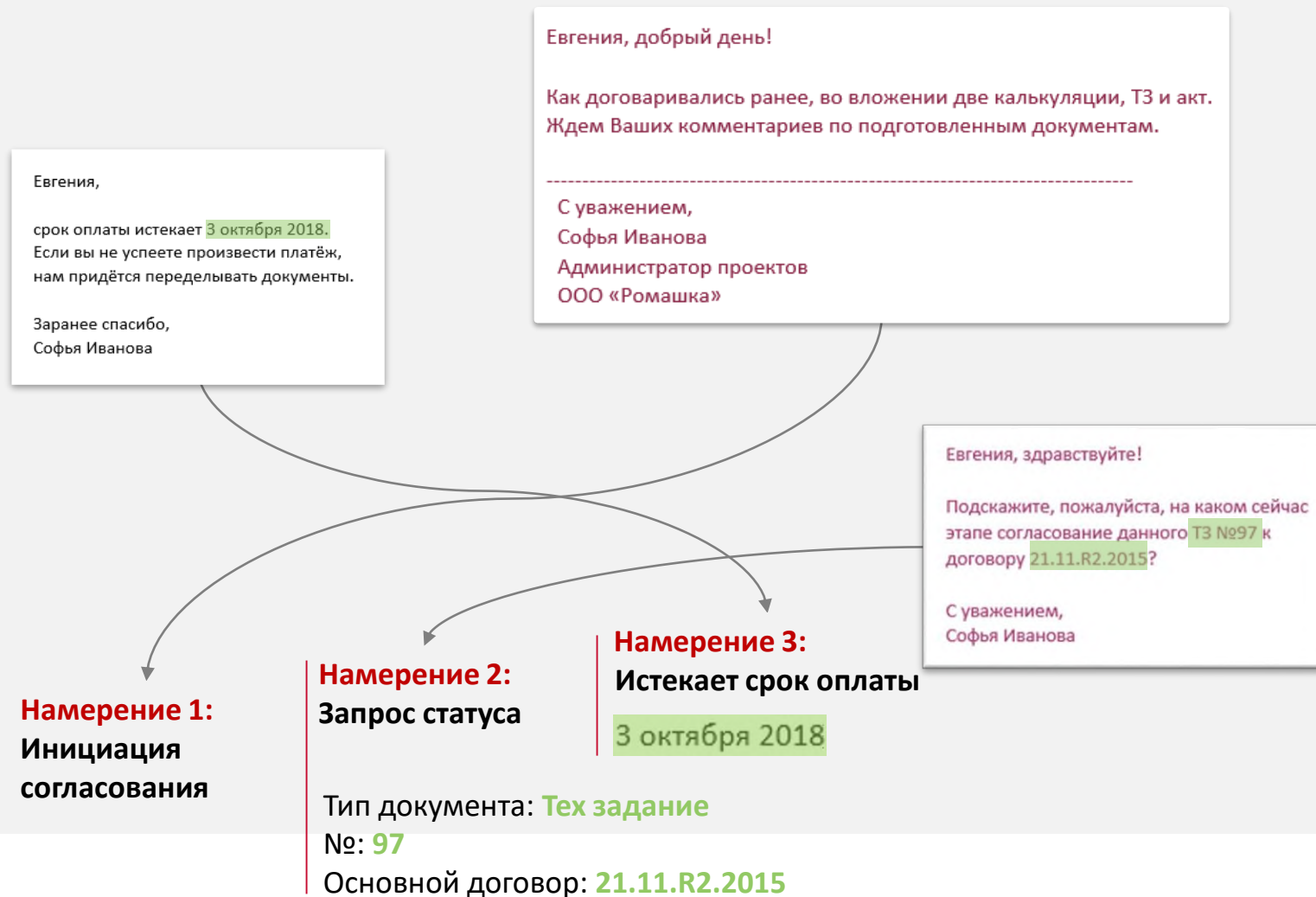
Классификация входящих обращений для  
автоматического ответа



Анализ новостей и оперативных сводок.

# Пример использования NLP

1. Классификация сообщений;
2. Извлечение именованных сущностей (NER);
3. Извлечение информации по пользовательской онтологии.



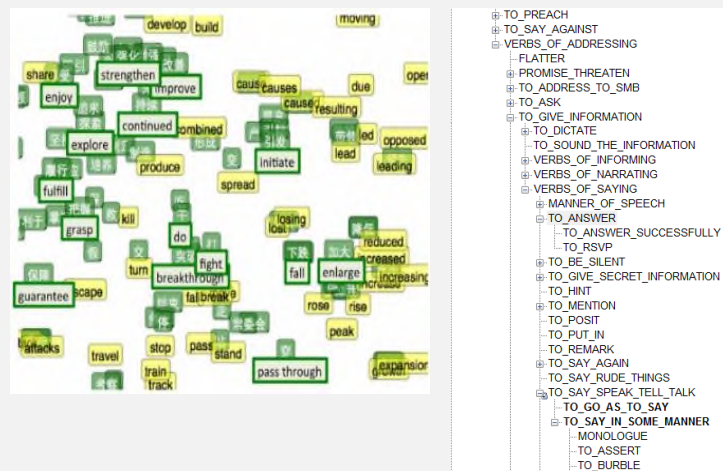
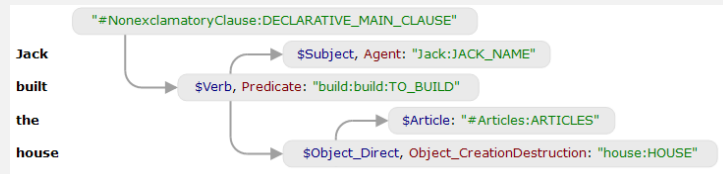
# Пример использования NLP

Мы, гражданин РФ **Иванов Иван Иванович**, 11 марта 1911 года рождения, место рождения: г. Москва, пол мужской, паспорт РФ 11 11 11111, выдан Отделением УФМС России по гор. Москве по району Отрадное 11.11.2015 г., код подразделения 112-011, зарегистрированный по адресу: Москва, ул. Отрадная, дом 2Б, именуемый в дальнейшем «ПРОДАВЕЦ», с одной стороны, и

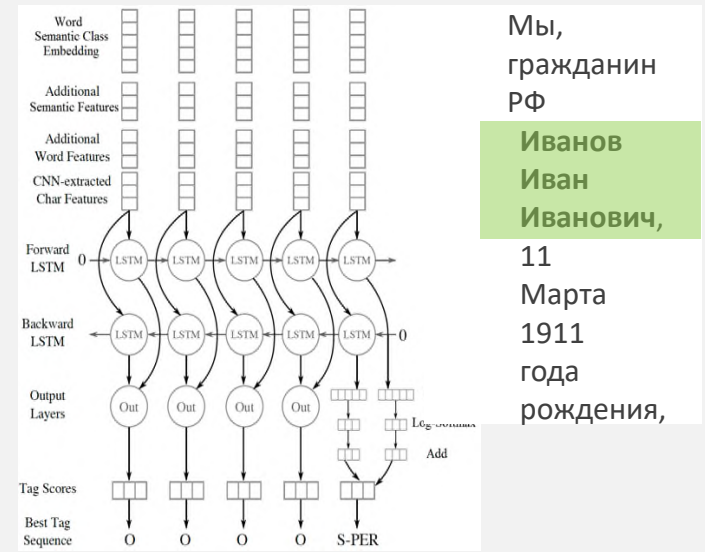
## Онтология

Продавец [Персона]  
Покупатель [Персона]

## Семантический разбор



## Тегирование слов



Мы,  
гражданин  
РФ  
**Иванов  
Иван  
Иванович,**  
11  
Марта  
1911  
года  
рождения,

## Сегментирование

**ДОГОВОР КУПЛИ-ПРОДАЖИ КВАРТИРЫ**

Город Москва, третьи октябрь две тысячи восемнадцатого года

Мы, гражданин РФ **Иванов Иван Иванович**, 11 марта 1911 года рождения, место рождения: г. Москва, пол мужской, паспорт РФ 11 11 11111, выдан Отделением УФМС России по гор. Москве по району Отрадное 11.11.2015 г., код подразделения 112-011, зарегистрированный по адресу: Москва, ул. Отрадная, дом 2Б, именуемый в дальнейшем «ПРОДАВЕЦ», с одной стороны, и

гражданин РФ **Петров Петр Петрович**, 22 марта 1922 года рождения, место рождения: г. Москва, пол мужской, паспорт РФ 22 22 22222, выдан Отделением УФМС России по гор. Москве по району Алексеевский 22.22.2015 г., код подразделения 112-022, зарегистрированный по адресу: Московская обл., город Чехов, улица Новомосковская, дом 2, именуемый в дальнейшем «ПОКУПАТЕЛЬ», с другой стороны, вместе именуемые «Стороны»,

находясь в здравом уме, ясной памяти, действуя добровольно, заключили настоящий договор о нижеследующем:

- ПРОДАВЕЦ продает ПОКУПАТЕЛЮ, а ПОКУПАТЕЛЬ покупает в собственность квартиру, расположенную по адресу: Московская область, город Балашиха, улица Весенняя, дом 33 (Тринадцать три), квартира 4 (Сорок четыре), далее – «КВАРТИРА».
- Указанная КВАРТИРА расположена на 11 (Одиннадцатом) этаже, состоит из 1 (Одной) жилой комнаты, имеет общую площадь 33,7 (Тринадцать три целых семь десятых) кв.м.
- Указанная КВАРТИРА принадлежит на праве собственности ПРОДАВЦУ на основании следующего документа: Договора № 1234567890 участия в долевом строительстве многоквартирного дома от 01.01.2014 г., зарегистрированного Управлением Федеральной службы государственной регистрации, кадастра и картографии по Московской области 02.02.2014 г. за № 098/7654321; Передаточного акта от 03.05.2014 г.; Разрешения на ввод объекта в эксплуатацию от 04.04.2014 г. №2015, что подтверждается Свидетельством о государственной регистрации права, бланк серия 50-50 2015, выданным Управлением Федеральной службы государственной регистрации, кадастра и картографии по Московской области 05.05.2015 г., о чем в Едином государственном реестре прав на недвижимое имущество и сделок с ним 06.05.2015 г. сделана запись регистрации № 50-50-2015.
- По соглашению сторон ПОКУПАТЕЛЬ купил у ПРОДАВЦА указанную КВАРТИРУ за 3 000 000 (Три миллиона) рублей 00 копеек. ПОКУПАТЕЛЬ обязуется выплатить вышеуказанную сумму ПРОДАВЦУ в течение 1 (Одного) дня после государственной регистрации перехода права собственности по настоящему Договору в Управлении Федеральной службы государственной регистрации, кадастра и картографии по

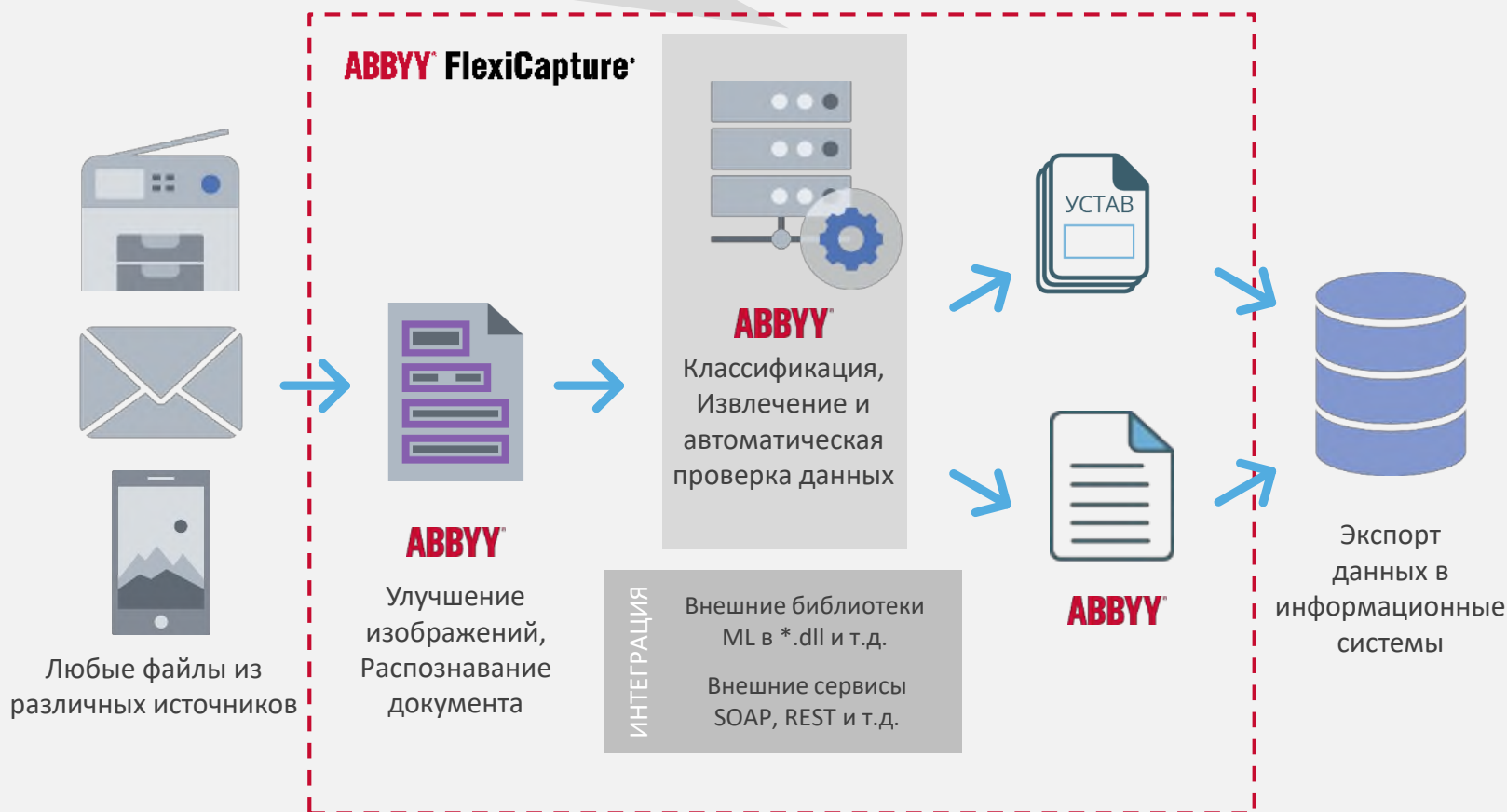


# Схема интеллектуальной обработки информации



## Извлечение данных из неструктурированных документов

С помощью технологий NLP, платформа умеет находить в текстах документов важные сущности, такие как имена, названия организаций и адреса, а также выявлять значимые факты и связи между ними.



## Эффект от внедрения:

- Снижение издержек на обработку информации в 2-5 раз
- Увеличение скорости поиска и обработки информации в 3–10 раз
- Повышение точности информации, введенной в ИС до 100%
- Минимизация риска утери информации
- Возврат инвестиций на внедрение 3-9 месяцев



Готовая enterprise-платформа, поддерживающая полный цикл обработки документов



NLP-технология работает со сложными неструктурированными документами



Готовая NER-технология позволяет извлекать именованные сущности, без дополнительных настроек и кастомизации



Использование широкого спектра встроенных алгоритмов позволяет работать с документами, у которых неидеальное качество распознавания



Не требуется большого количества примеров, для обучения требуется всего 20-50 примеров документов одной выборки;



Разметка и обучение модели проходит в «одном окне», >> нет потребности привлекать дата-сайентистов для разметки документов



# Успешные кейсы применения NLP-технологий

# Популярные сценарии по работе с неструктурированными данными в банках





## Задачи

- Оперативная обработка архивных юридических документов корпоративных клиентов банка.
- Отправка структурированных данных в систему применения правил автоматизированной юр. экспертизы.
- Минимальное время оформления кредита по «best offer».

## Решение

NLP технологии ABYY:

- Анализируют структуру документов.
- Определяют тип документов.
- Извлекают необходимые атрибуты.

## Результаты



Показатель качества извлечения данных **95%**.



Сокращение времени на оформление кредита до 7 минут.



> 23 типов обработанных документов. Из них 13 - в неструктурированном виде.



Получено 1227 атрибутов.

## Задачи

- Автоматизация поиска в новостях негативных событий с участием банков-контрагентов.
- Сокращение трудозатрат на обработку сообщений.
- Оценка факторов риска в реальном времени.

## Решение

NLP технологии ABBYU:

- Классифицируют новости и извлекают факты и связи между ними из текстов.
- Определяют рисковые индикаторы и связанные с ними организации и персоны.

## Результаты



Оперативное реагирование на значимые события.



Извлечение 9 риск-факторов из новостей в онлайн-режиме с точностью – 92% и полнотой – 87,8%.

## Задачи

- Увеличение количества новых клиентов.
- Оперативный ответ на вопросы пользователей.
- Автоматическое распределение клиентских обращений по их тематикам.

## Решение

NLP технологии ABBYY:

- За 1 секунду анализируют содержание запроса, определяют его тематику.
- Проставляют теги в карточке разговора в CRM.

## Результаты



Повышение скорости обработки запросов на 15%.



Увеличение количества новых клиентов на 10%.



Уменьшение времени на классификацию резюме разговоров до 1 секунды.

# Автоматизация работы с обращениями гос. органов

## Задачи

- Увеличение скорости обработки обращений от государственных органов.
- Снижение риска просрочки ответа на важные запросы регуляторов.

## Решение

NLP технологии ABBYY:

- Классифицируют все входящие запросы.
- Извлекают данные, необходимые для подготовки ответа по запросу.

## Результаты



Оптимизация времени на обработку одного запроса до 30 секунд.



> 4000 обработанных документов в месяц.



Повышение скорости ответов на обращения регуляторов в 5 раз.





# Демонстрация работы NLP в решении АБВУУ FlexiCapture



## Контакты

---

**Ольга Морозова**

*Директор по работе с финансовым сектором, АБВУУ Россия*

[olga.morozova@abbyy.com](mailto:olga.morozova@abbyy.com)

**Дмитрий Черноус**

*Заместитель директора по консалтингу и развитию бизнеса, АБВУУ Россия*

[dmitry.chernous@abbyy.com](mailto:dmitry.chernous@abbyy.com)

