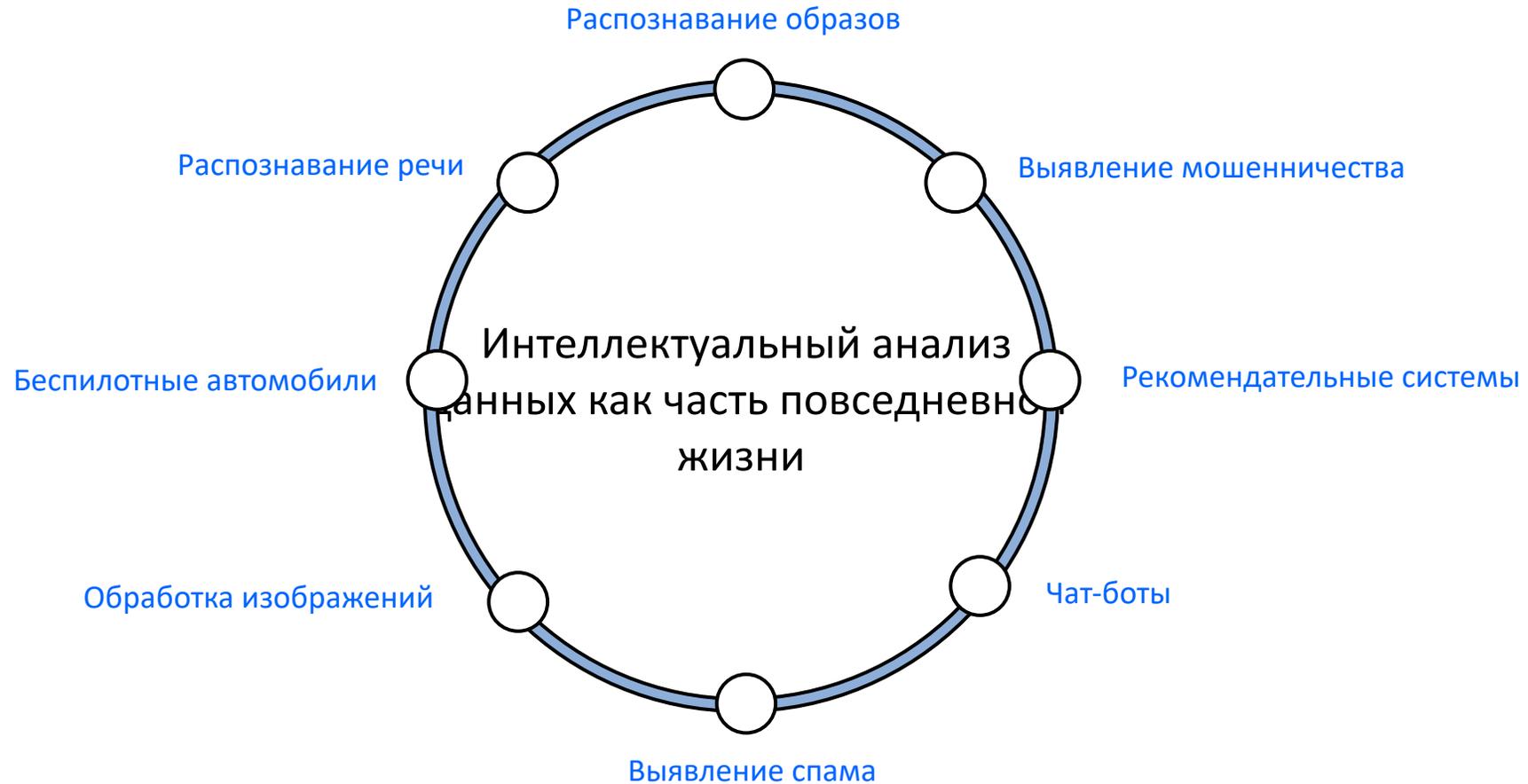


# AI for Social Good

**Кошелева Евгения | Станислав Кладов**  
18.09.2019

<b>Введение</b>
<b>Основы машинного обучения</b>
<b>IBM Science for social good</b>
<b>Конкурс проектных идей Smart Social</b>
<b>Кейсы клиентского центра</b>

# Современные тенденции в области искусственного интеллекта



# Когнитивные системы — партнерство человека и технологий

## Сильные стороны человека

- Здравый смысл
- Мораль
- Воображение
- Сравнение
- Абстрагирование
- Выбор
- Мечты
- Обобщение

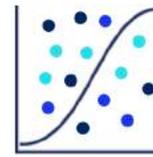


## Сильные стороны КОГНИТИВНЫХ СИСТЕМ

- Извлечение знаний из данных
- Распознавание повторяющихся структур
- Запросы на естественном языке
- Машинное обучение
- Непредвзятость
- Быстродействие

<b>Введение</b>
<b>Основы машинного обучения</b>
<b>IBM Science for social good</b>
<b>Конкурс проектных идей Smart Social</b>
<b>Кейсы клиентского центра</b>

# Введение в машинное обучения



## Машинное обучение -

подраздел  
искусственного  
интеллекта, которые  
учит компьютер  
самостоятельно  
находить решения  
различных задач.



**Данные/  
Признаки**



**Алгоритмы**



## **Может**

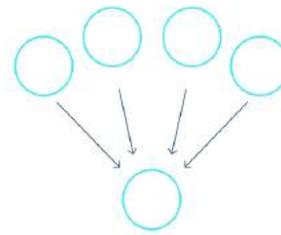
- Предсказывать
- Воспроизводить
- Запоминать
- Выбирать лучшее

## **Не может**

- Выйти за рамки задачи
- Резко поумнеть
- Создавать совсем новое



# Классификация/регрессия



## Определение

### Что это?

- Подход, при котором человек собирает исходные данные и размечает желаемую целевую переменную, а алгоритм учится находить связь между ними

### Когда использовать?

- Если вы знаете как классифицировать исходные данные и хотите автоматизировать предсказание целевой переменной применительно к новым данным

## Принцип работы



## Примеры использования

### Классификация:

- Спам-фильтры
- Поиск похожих документов
- Анализ тональности текста
- Определение подозрительных транзакций

### Регрессия:

- Анализ спроса, объема продаж
- Прогноз стоимости ценных бумаг

# Кластерный анализ



## Определение

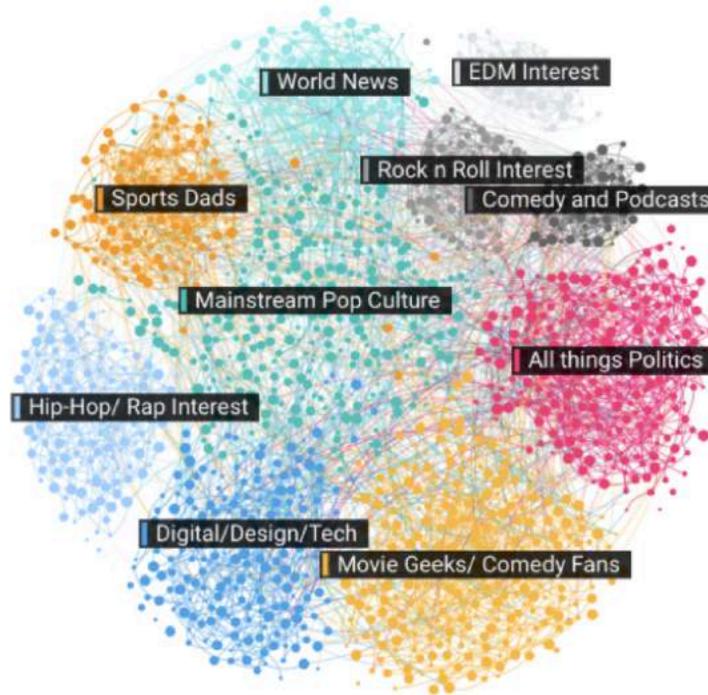
### Что это?

- Подход при котором алгоритм получает неразмеченные данные , исследует их и возвращает группы, которые демонстрируют схожее поведение

### Когда использовать?

- Вы не знаете, как классифицировать данные, и вы хотите, чтобы алгоритм нашел шаблоны и классифицировал данные для вас

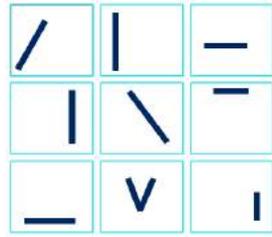
## Принцип работы



## Примеры использования

- Сегментация рынка (типов покупателей, лояльности)
- Объединение близких точек на карте
- Тематическое разделение новостей по содержанию
- Детекция аномального поведения

# Нейронные сети



## Определение

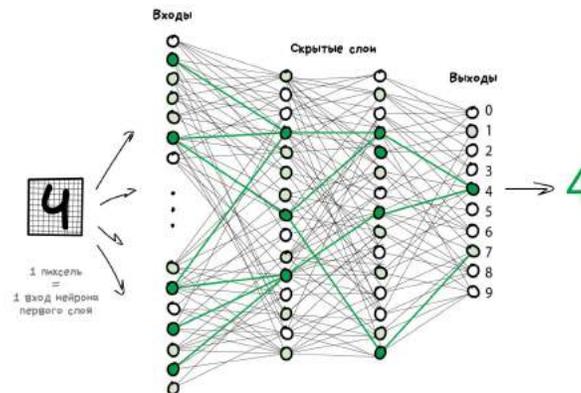
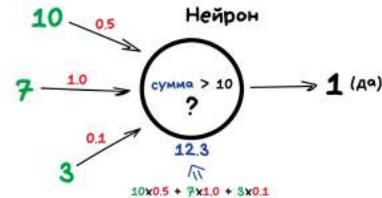
### Что это?

- система соединенных и взаимодействующих между собой простых функций (нейронов). Будучи соединёнными в достаточно большую сеть, эти нейроны вместе способны выполнять довольно сложные задачи.

### Когда использовать?

- Когда есть коллекция изображений или видео и нужно определить что там изображено
- Если данные представлены в виде последовательностей: аудиозаписи, текст...

## Принцип работы



## Примеры использования

### Сверточные нейронные сети

- Диагностика заболеваний по медицинским снимкам
- Обнаружение дефектных продуктов при производстве
- Распознавание лиц

### Рекуррентные нейронные сети

- Машинный перевод текстов
- Компьютерный синтез речи
- Автоматическое создание подписей к картинкам
- Генерация текста (например для чат-ботов)

<b>Введение</b>
<b>Основы машинного обучения</b>
<b>IBM Science for social good</b>
<b>Конкурс проектных идей Smart Social</b>
<b>Кейсы клиентского центра</b>

# IBM Science for Social Good

- Запущенная в 2016 году, совместная с IBM Research программа, которая предоставляет наработки компании в сфере ИИ для проведения социальных научных исследований.
- Примеры тем проектов: предсказание развития эпидемий, исследования причин возникновения опиоидной зависимости, выявление ксенофобных высказываний в социальных сетях



# Neighborhood Trust Financial Partners\* & IBM Research

## Проблема

Работники с небольшой заработной платой часто бывают финансово неграмотны, что приводит их к частому оформлению небольших кредитов

## Решение

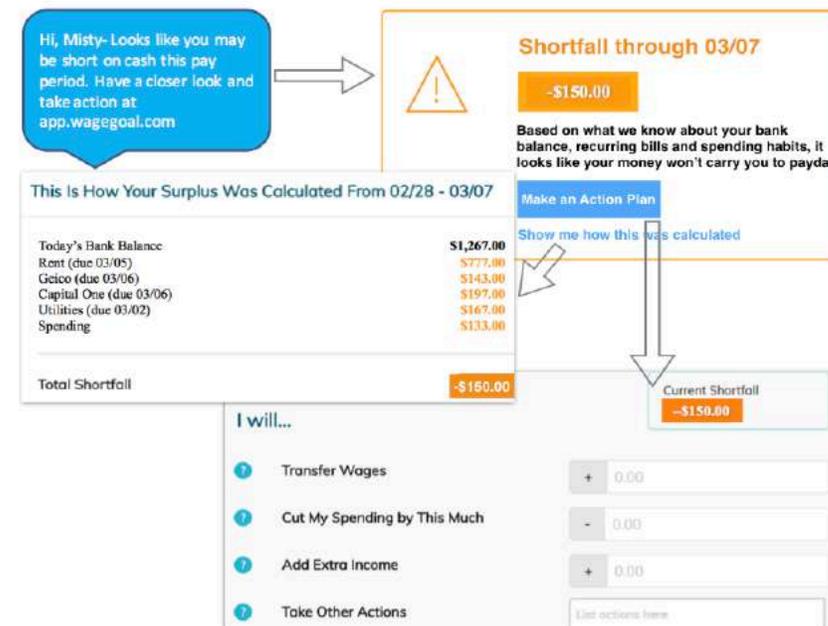
Приложение - финансовый советник, предоставляющий персонализированные финансовые рекомендации, помогающее пользователям скорректировать свое финансовое поведение и избегать долгов.

## Финансовые консультации для низкооплачиваемых работников

**Реализация:** Система, которая умеет точно оценивать траты каждого человека:

- Прогнозировать фиксируемые, обязательные траты, которые должны быть совершены в выбранный период
- Прогнозировать траты по стандартным категориям
- Оценивать вероятности/склонности к непредвиденным расходам

=> Предсказывать остатки на счетах физических лиц и давать рекомендации, чтобы дожить до зарплаты



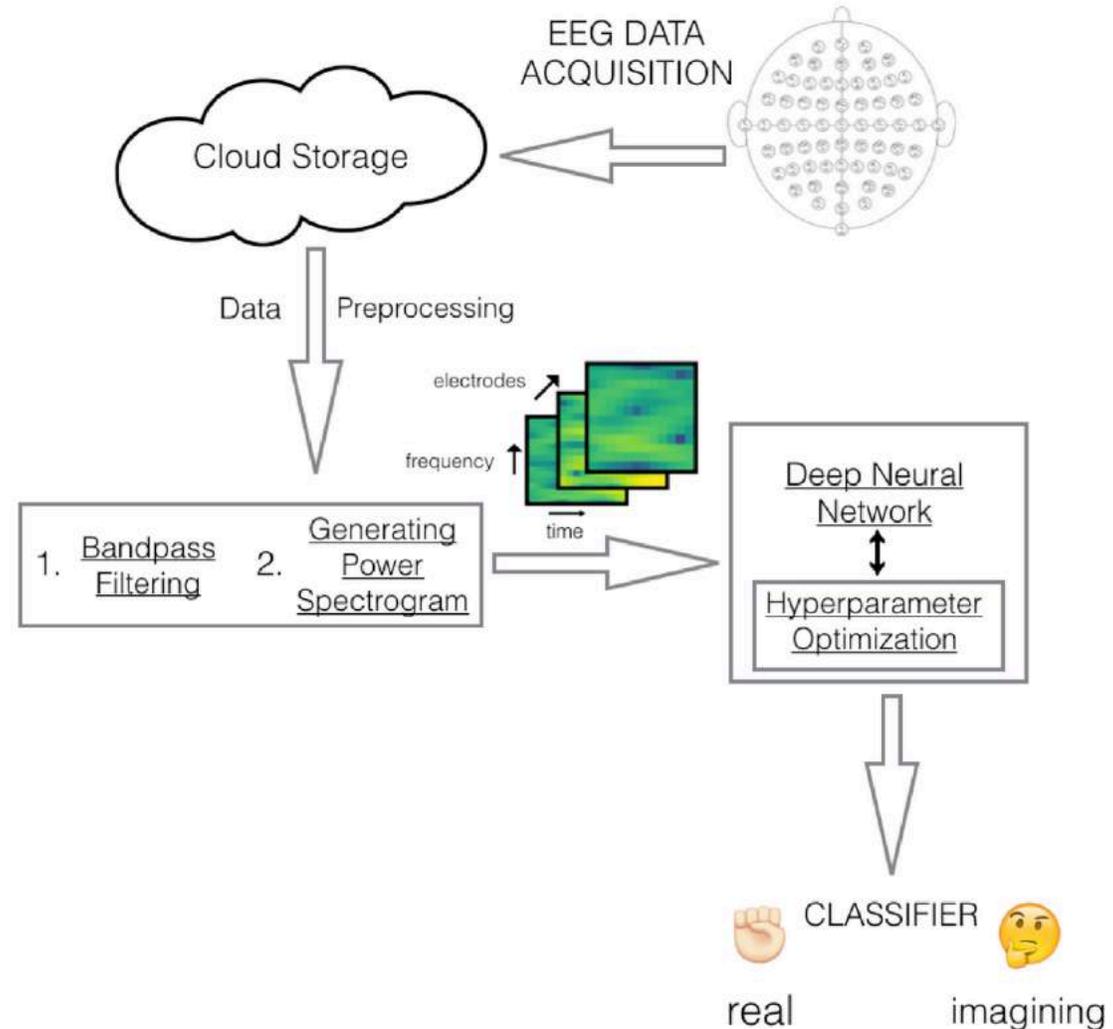
## Проблема

- Треть всех болезней в развивающихся странах приходится на такие расстройства мозга, как эпилепсия, шизофрения и депрессия.
- Есть нехватка квалифицированных специалистов для диагностики такого рода заболеваний

## Решение

Создано прототип - система обеспечения поддержки принятия решений для врачей-неспециалистов в области анализа и интерпретации электроэнцефалографии, определяющее случаи патологических изменений в ЭЭГ

## Использование машинного обучения для интерпретации ЭЭГ



# Echoing Green\* & IBM Research

## Проблема

- Echoing Green - получает около 3,000 заявок в год на ограниченное число стипендии, которые включают в себя финансовую и техническую поддержку.

- выявить тенденции и характеристики успешных заявок, чтобы помочь перспективным социальным предпринимателям лучше направлять свои усилия

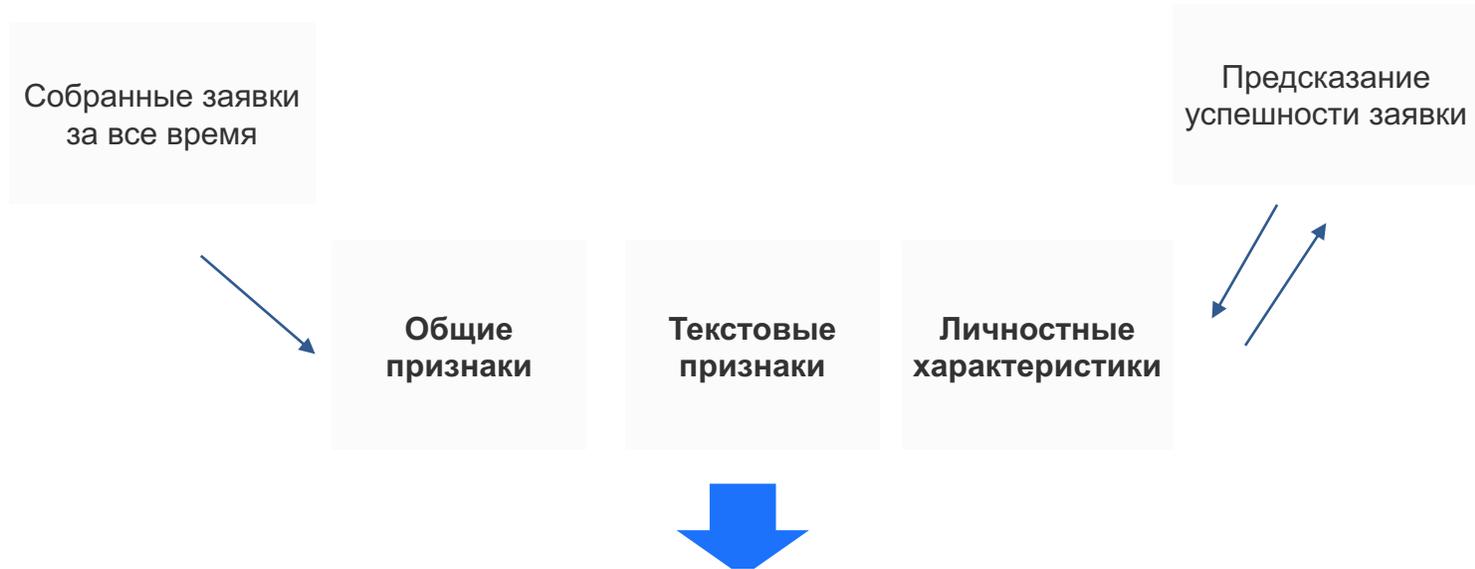
## Решение

Была проанализирована коллекция заявок для выявления тенденций и характеристик, приводящих к одобрению заявки.

\* Международная некоммерческая организация, занимающаяся выдачей стипендий социальным предпринимателям

# Изучение социальных тенденций, основываясь на данных

## Реализация:



Feature Type (Count)	Feature Name
Non-Textual (2)	Total Funds Raised, Re-applicant
Syntactic (8)	Number of Words – Applicant, Idea (2); Number of Uppercase Words - Applicant; Text Uniqueness – Applicant, Organization, Idea, Problem (4); Number of Sentences - Idea
Semantic (8)	Social Business and Team Experience, Online Platform / Data, Legal Organization/Rights, Water/Clean Energy, Women/Children, Local/Community Work, Business Venture/Project, Youth/Society
Lexical (1)	Helping People category
Personality (6)	Self-transcendence, Curiosity, Immoderation, Openness to Change, Self-Discipline, Neuroticism



# Коалиция грамотности Центрального Техаса & IBM Research

## Проблема

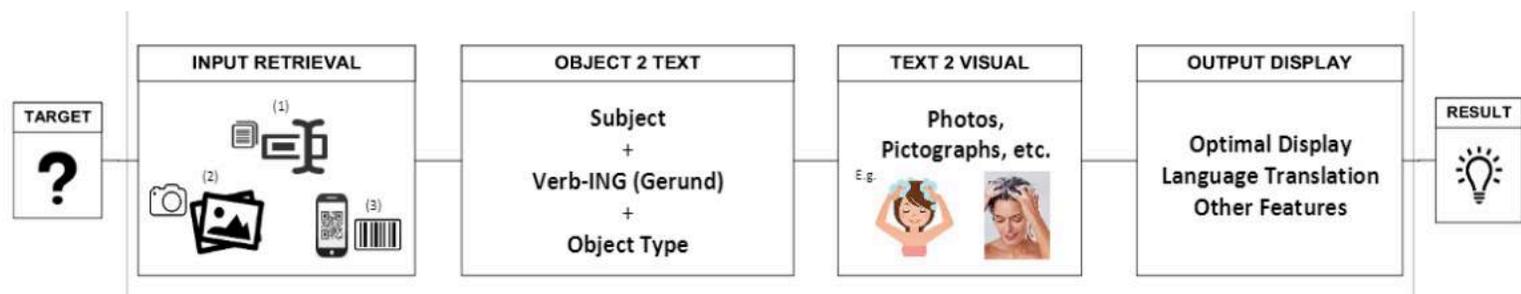
- Примерно 757 миллионов взрослых по всему миру не умеют читать
- У многих людей возникают трудности с пониманием сложных текстов

## Решение

- Создано приложение, помогающее малограмотным взрослым ориентироваться в информационном мире
- извлекает ключевые сообщения из текста и представляют их пользователю в виде изображений, сопровождаемых короткими и простыми устными сообщениями.

# Simpler Voice: Генератор визуального описания текста для малограмотных людей

## Реализация:



## Пример результата:

The screenshot shows the application's output for a scanned UPC code (3700098084) for Cascade Complete dish detergent. The interface includes:

- Scanned UPC code: 3700098084
- Category path: home-and-kitchen -> laundry-and-household -> dish-care -> auto-dish-care
- Search string: woman washing with Dishwasher Detergent
- Simplified output: A grid of images showing a woman washing dishes, a woman washing her hands, and a woman washing her hair.

<b>Введение</b>
<b>Основы машинного обучения</b>
<b>IBM Science for social good</b>
<b>Конкурс проектных идей Smart Social</b>
<b>Кейсы клиентского центра</b>

### AppealBot

- Интеллектуальный помощник.
- помощь людям с редкими заболеваниями в получении необходимых жизненно-необходимых лекарственных препаратов

### Веб-трекинг

- Анализ поведения посетителей сайта благотворительного фонда
- Выработка решений по управлению контентом

### DBetter

- Поиск слепоглухих людей по косвенным признакам

### Агрегатор событий НКО

- Автоматизированный сбор информации с сайтов НКО
- Классификация мероприятий

### Технологии Чудес

- Анализ аудитории портала о благотворительности
- Повышение конверсии портала
- Интеллектуальный ассистент для посетителей

### Автоматизация социального проектирования

- Прогнозирование результата социального проекта на основе анализа больших объемов данных

<b>Введение</b>
<b>Основы машинного обучения</b>
<b>IBM Science for social good</b>
<b>Конкурс проектных идей Smart Social</b>
<b>Кейсы клиентского центра</b>

# Рекомендательные системы

2.5 нониллионов\*  
байтов новых данных  
создается ежедневно

Профессиональный опыт сегодня важен,  
как никогда, но даже ведущие эксперты  
не могут справиться с таким потоком  
новой информации

Вы готовы к новым вызовам и  
возможностям?

# IBM Watson Explorer -

это когнитивный инструмент по обработке любой **текстовой информации.**

**Основа** для работы с естественным языком

### Доступ к информации

Данные, приложения и сервисы, локально или в облаке – сотрудником приходится прилагать усилия, чтобы получить полную картину



### Исследование

Возможность видеть информацию из **всех** систем и источников на одном экране, настраивать кластеризацию и фильтры

### Неструктурированный контент

**80% данных** – не структурированы. Только небольшая их часть используется для формирования гипотез



### Анализ

Анализ структурированного, а также неструктурированного контента для понимания сути предметной области

### Поиск

Эксперты проводят **88% времени** в поисках скрытых ответов, вместо того, чтобы действовать



### Обнаружение

Значительное увеличение скорости и качества работы сотрудника при предоставлении ему контекстного анализа запрашиваемой информации

### Быстрое эффективное принятие решений

Необходимо постоянно увеличивать производительность и инновационность меньшими ресурсами



### Расширение знаний

Получение новых знаний и эффективное принятие решений при возможности быстро изучить содержание исследуемой области

# Сценарий: Система анализа мнений клиента

### Задачи:

- Автоматизация сбора данных отзывов
- Статистика
- Анализ мнений по продуктам компании
- Анализ контекста упоминаний продуктов
- Анализ эмоций клиентов

### Источники информации:

- Веб-ресурсы: twitter, блоги, сайт заказчика, новости
- Текстовые файлы: xml, txt, html, .....



### Описание

Анализ эмоциональной окраски социальных медиа вокруг определенных событий

### Пример

[https://vk.com/championat?w=wall-1331201\\_5050345](https://vk.com/championat?w=wall-1331201_5050345)

### Продукты

- Watson Explorer



# Сценарий: Система рейтинга поставщиков

## Сценарий 1

### Система рейтинга поставщиков

#### Задачи:

- Автоматизация сбора статистических данных о деятельности подрядчика
- Формирование рейтинга подрядчиков на основании собранной статистики
- Рекомендации по формулированию тактики закупок
- Прогноз «благонадежности» поставщика

# Требования

## Возможности анализа неструктурированных данных:

- **Веб-ресурсы** (Новостные источники, Реестр госзакупок, Росстат, портал раскрытия информации и т.д.)
- **Текстовые файлы** Adobe Acrobat, MS Office (выгрузка из ЕГРЮЛ, поставления из картотеки арбитражных дел)
- **Корпоративная почта** (переписка сотрудников)

## Оценка поставщиков

Формирование оценки поставщиков по следующим областям:

- **Сроки поставок**
- **Финансовая стабильность**
- **Благонадежность**
- **Инновации**

## Формирование рекомендаций по работе с поставщиками (сделано сверх требований)

Рекомендации сформированы на основе правил обработки неструктурированных данных на примере компании N

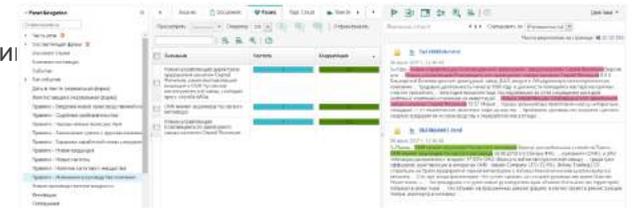
# Построение решения



1. Сбор данных из различных источников по ключевым критериям



Наименование поставщи  
Тип события  
Событие



2. Обработка неструктурированных данных с помощью правил работы с естественным языком



Сфера влияния

3. Классификация событий, определение сфер влияния

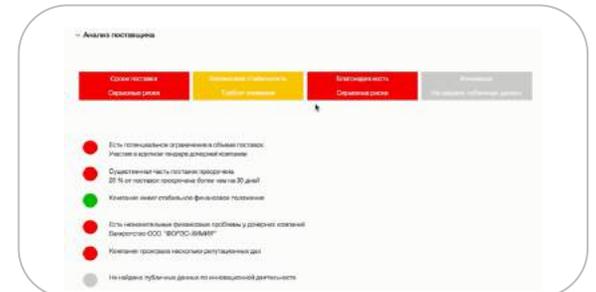


Формирование рекомендаций

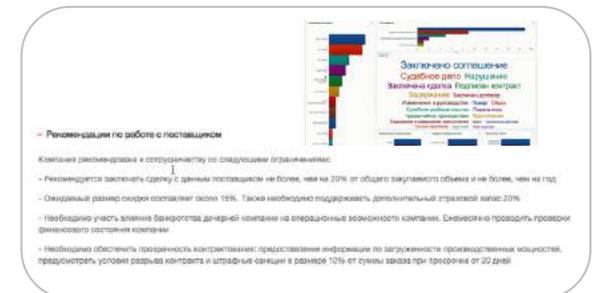
# Результаты



Карточка поставщика



Оценка поставщика по сферам влияния



Формирование рекомендаций



## Досье поставщика на примере компании N

### Выделены важные события

Сроки поставки  
Серьезные риски

- Есть потенциальное ограничение в объеме поставок  
Участие в крупном тендере дочерней компании
- Существенная часть поставок просрочена  
20 % от поставок просрочена более чем на 30 дней

Финансовая стабильность  
Требует внимания

- Компания имеет стабильное финансовое положение
- Есть незначительные финансовые проблемы у дочерних компаний  
Банкротство ООО "ФОРЭС-ХИМИЯ"

Благонадежность  
Серьезные риски

- Компания проиграла несколько репутационных дел

Иновации  
Не найдено публичных данных

- Не найдено публичных данных по инновационной деятельности

Аккумулированная оценка поставщика складывается из оценок по каждой рискованной области

### Сформированные рекомендации

#### ▼ Рекомендации по работе с поставщиком

Компания рекомендована к сотрудничеству со следующими ограничениями:

- Рекомендуется заключать сделку с данным поставщиком не более, чем на 20% от общего закупаемого объема и не более, чем на год
- Ожидаемый размер скидки составляет около 15%. Также необходимо поддерживать дополнительный страховой запас 20%
- Необходимо учесть влияние банкротства дочерней компании на операционные возможности компании. Ежемесячно проводить проверки финансового состояния компании.
- Необходимо обеспечить прозрачность контрактования: предоставление информации по загрузке производственных мощностей, предусмотреть условия разрыва контракта и штрафные санкции в размере 10% от суммы заказа при просрочке от 20 дней

### Компании присвоен высокий риск

#### ▼ Карточка поставщика

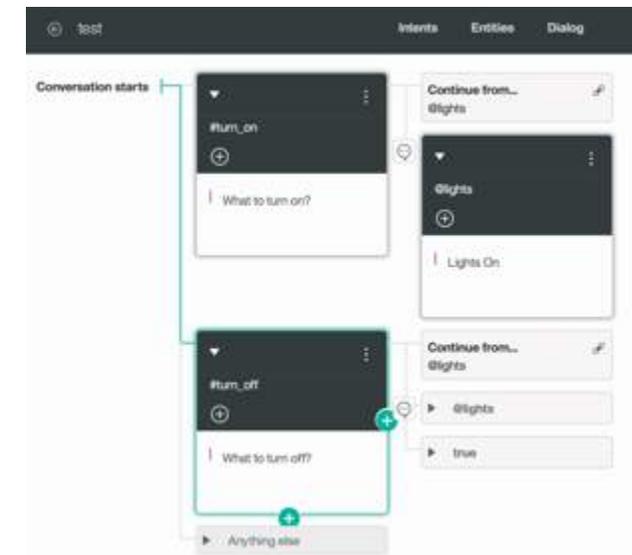
Краткое наименование компании: ООО "ФОРЭС"  
Генеральный директор: ЛОВКОВ МИХАИЛ АЛЕКСАНДРОВИЧ  
Официальный сайт: [www.forestld.com](http://www.forestld.com)  
[Показать все данные](#)



Итоговая оценка:  
Высокие риски

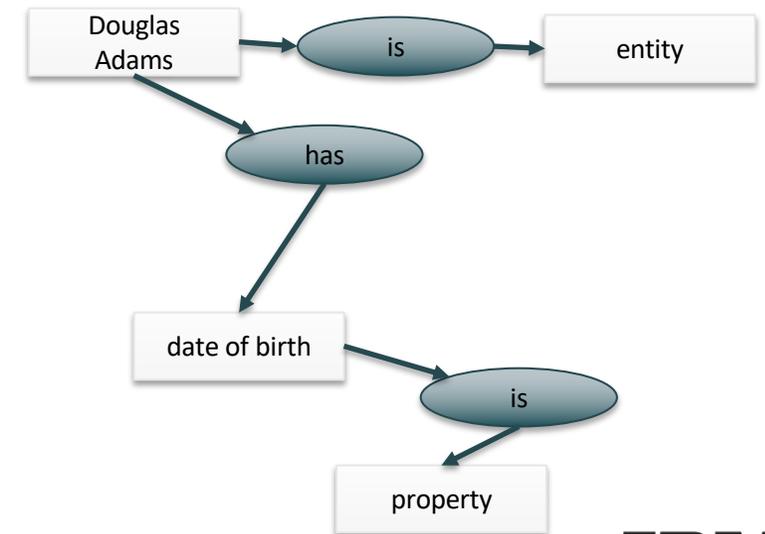
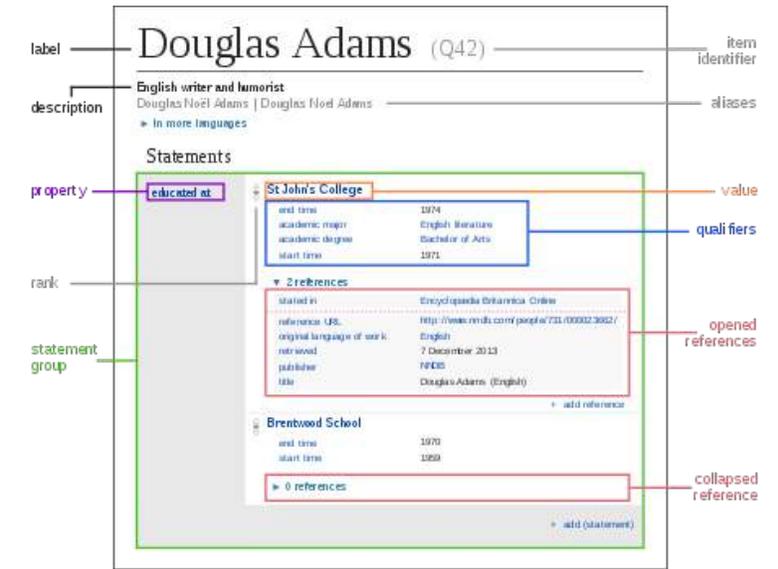
# Интеллектуальные ассистенты и чат-боты

- Платформа для создания и управления ассистентами
- Построен вокруг механизма разбора естественного языка, позволяющего распознавать намерение говорящего и уточняющие сущности.
- Также предоставляет богатый набор функций для более удобного создания, отладки, обучения и интеграции ассистентов



## Шаги по построению ассистента:

- ✓ Формулировка задачи
- ✓ Выявление классов обращений клиентов на основании имеющихся диалогов
- ✓ Понимание «языка» клиентов по каждому из классов обращений
- ✓ Построение дерева диалога
- ✓ Наполнение/обучение
- ✓ Тестирование



# Сценарий: Чат-бот и интеллектуальный помощник сотрудника службы поддержки

## Сценарий

Чат-бот и интеллектуальный помощник сотрудника службы поддержки

### Цель:

- Оптимизация нагрузки на call-центр путем перенаправления части рутинных вопросов в чат-бот
- Оптимизация времени специалиста call-центра

### Задачи:

- Обучение системы автоматической классификации и ответу на входящие вопросы на основании исторических данных
- Реализация вопросно-ответной системы в виде чат-бота

## Требования

### Возможности анализа неструктурированных данных

- Массив исторической переписки с клиентами
- Текст текущего обращения клиента, полученный через интерфейс чат-бота

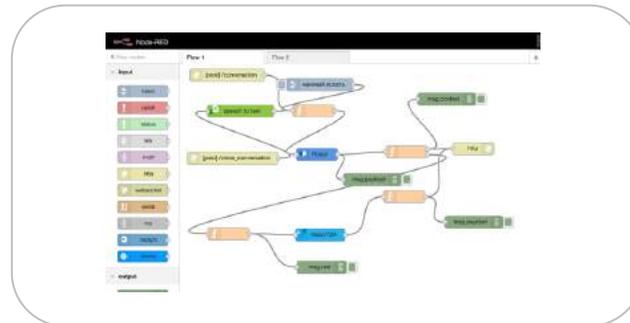
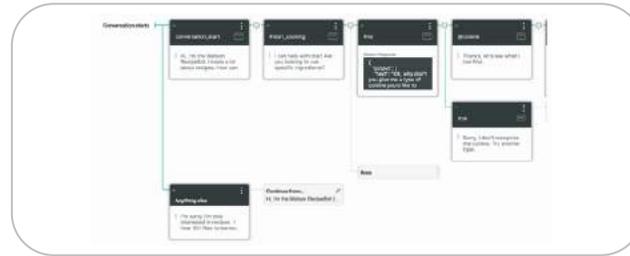
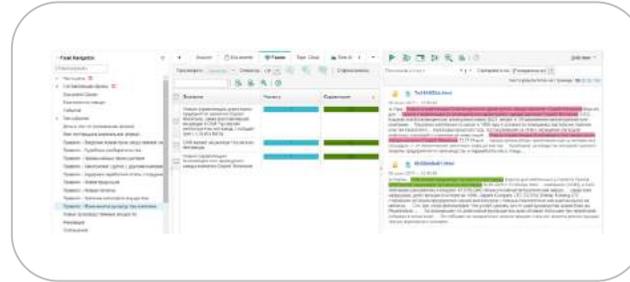
### Работа с историческими данными

- **Кластеризация** – объединение в одну сущность различных в написании, но несущих один смысл формулировок вопросов и ответов
- **Формирование графа диалога** – при необходимости добавление ветвления - разный ответ на один вопрос в зависимости от дополнительных параметров, и проч.

### Реализация чат-бота

- Автоматический ответ на вопросы клиентов
- Режим «подсказок» оператору
- Возможность подключения к системам заказчика для получения дополнительной информации

## Построение решения



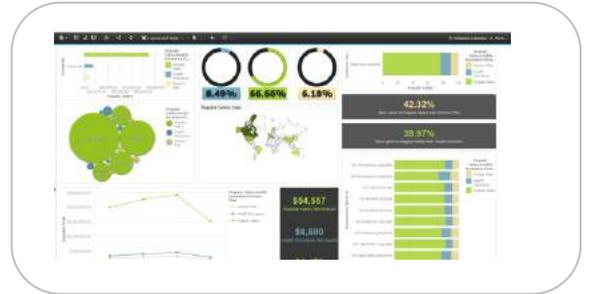
1. Загрузка исторических данных

2. Выявление классов вопросов и релевантных языковых конструкций для каждого из них

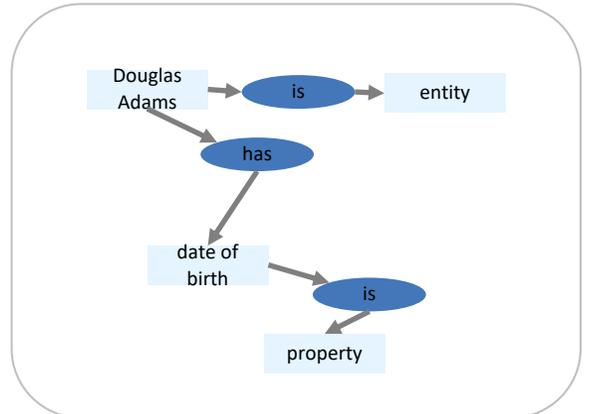
3. Формирование графа диалогов

4. Наполнение и обучение чат-бота

## Результаты



Аналитический отчет по истории переписки с клиентами



Чат-бот, позволяющий автоматически отвечать на часть запросов клиентов

## Сценарий

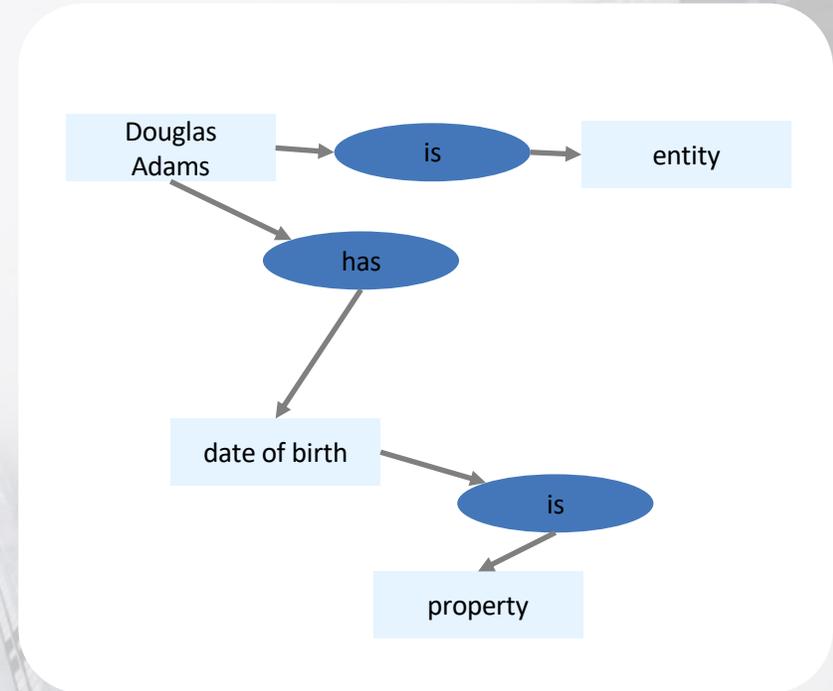
Чат-бот и интеллектуальный помощник сотрудника службы поддержки

### Достигнутые результаты:

- Выявление классов обращений клиентов на основании имеющихся диалогов
- Понимание языковых конструкций, релевантных каждому типу вопроса
- Построение дерева диалога
- Наполнение/обучение чат-бота
- Тестирование

### Целевые показатели (\*):

- Оптимизация нагрузки на call-центр - 40%
- Оптимизация времени специалиста call-центра - 25%



*\* Показатели рассчитаны исходя из опыта предыдущих типовых проектов. Показатели приведены для информации и могут отличаться после внедрения.*

**Сценарий:** Интеллектуальный помощник сотрудника торгового зала в ритейле

## Сценарий

Интеллектуальный помощник  
сотрудника торгового зала в  
ритейле

### Цель:

- Оптимизация навигации в инструкциях и документах для различных запросов в торговом зале
- Снижение времени на первичное изучение документов
- Оптимизация скорости и качества работы сотрудника торгового зала

### Задачи:

- Обучение системы автоматической классификации и ответу на входящие запросы на естественном языке на основании исторических данных
- Выдача релевантных инструкций и документов
- Реализация интерфейса в виде поисковой системы и чат-бота

## Требования

Неструктурированные данные

- **Документы** – инструкции по работе с клиентами для разных потенциальных ситуаций

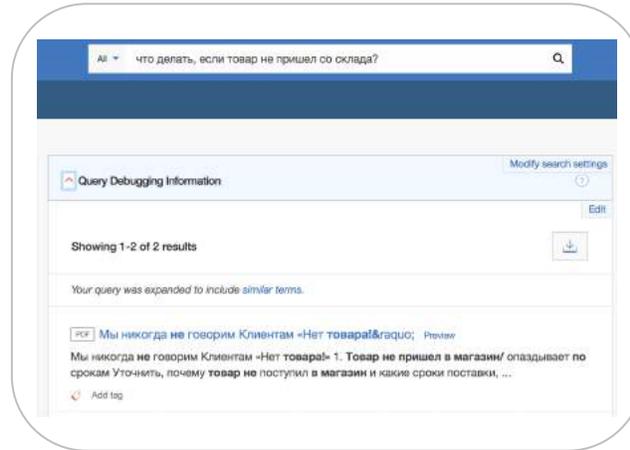
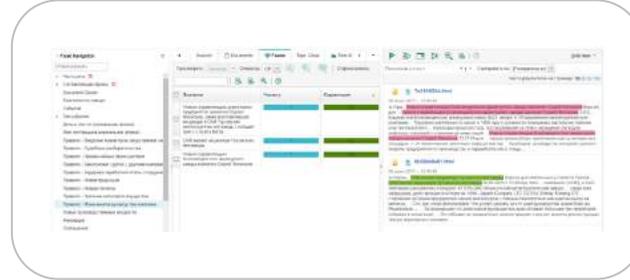
Анализ данных

- **Кластеризация** – объединение в одну сущность различных в написании, но несущих один смысл формулировок вопросов и ответов
- **Выделение сущностей**, релевантных и значимых для различных типов документов и ситуаций

Интеллектуальный помощник сотрудника

- Выдача необходимых документов по запросу в свободной форме
- Реализация интерфейса в виде чат-бота

## Построение решения

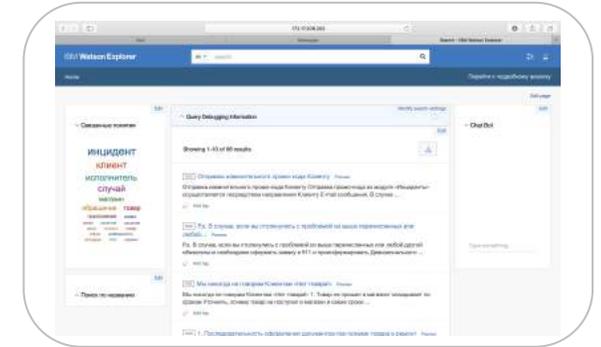


1. Загрузка документов

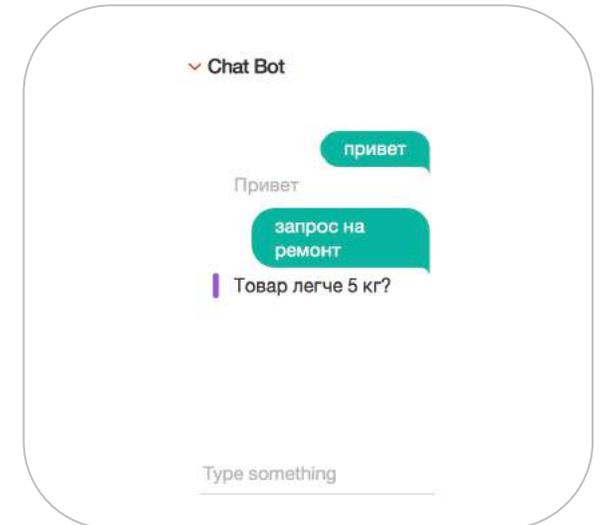
2. Выявление релевантных языковых конструкций для различных ситуаций и запросов

3. Настройка поиска документов по запросу в свободной форме

## Результаты



Поисковый интерфейс



Интерфейс в виде чат-бота

## Сценарий

Интеллектуальный помощник  
сотрудника торгового зала в  
ритейле

### Достигнутые результаты:

- Текстовый анализ инструкций-документов, понимание языковых конструкций, релевантных каждому типу документов
- Автоматическая классификация входящих запросов
- Реализация поисковой системы, поддерживающей запросы на естественном языке, и интерфейса в виде чат-бота

# Сценарий: Помощник консьержа

## Сценарий

Помощник консьержа

## Hilton & IBM

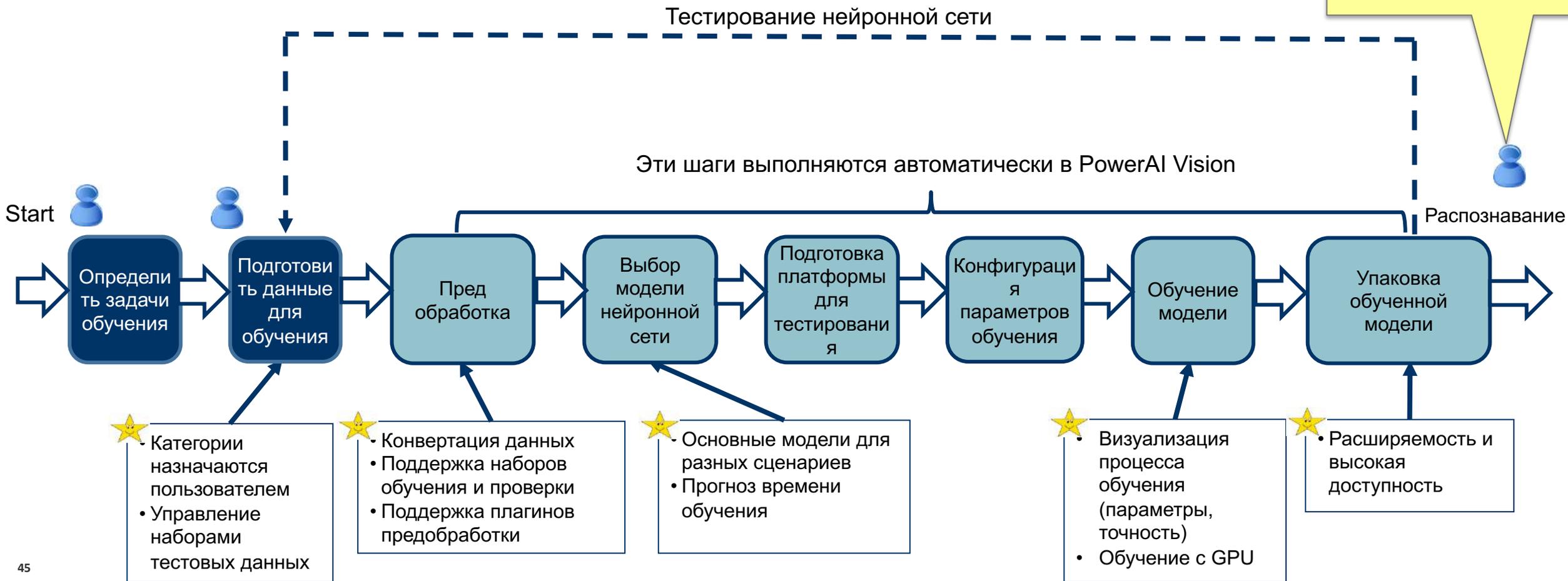


# Компьютерное зрение

# В чем ценность PowerAI Vision?

- **PowerAI Vision** автоматизирует циклы машинного обучения для разработчиков.
- Технологии машинного/глубокого обучения уже встроены в продукт.

Можно пользоваться API для распознавания!

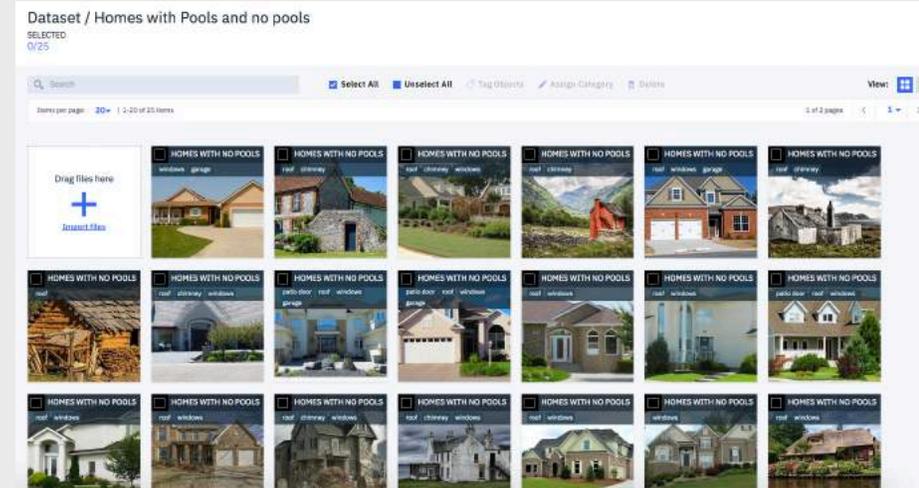
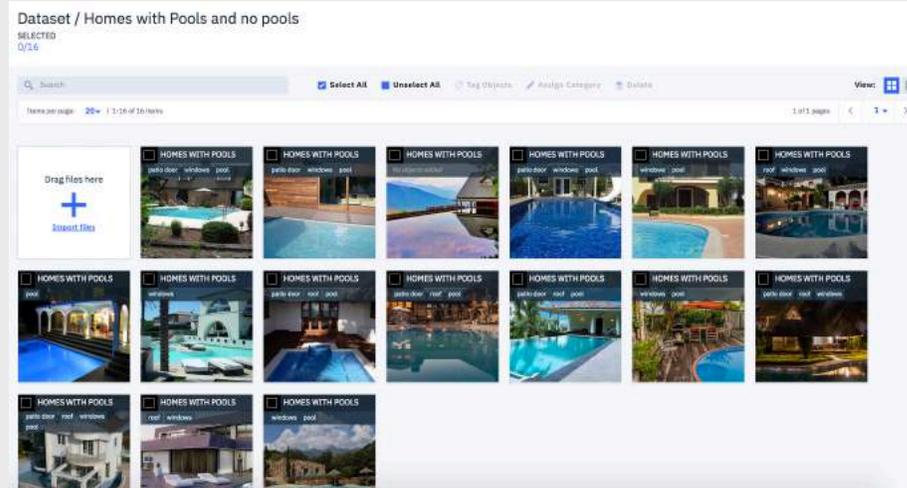


# Классификация изображений

[DEMO LINK](#)

## Разметка

1. Соберите изображения
2. Создайте категории
3. Добавьте изображения в категории
4. Обучите
5. Тестируйте

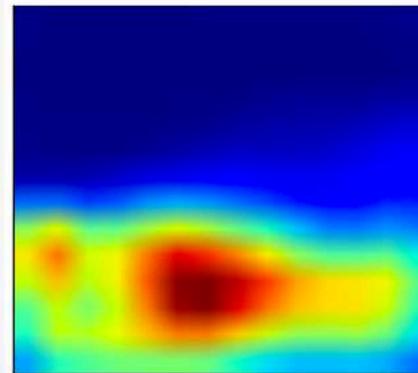


## Вывод



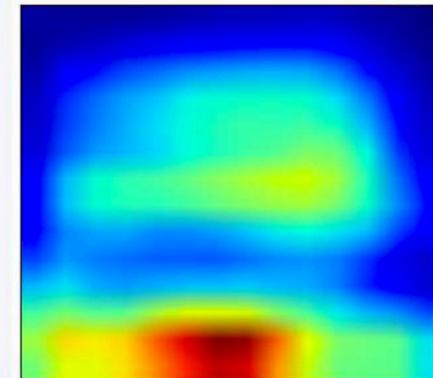
### Results

Homes with Pools: 100% Accuracy



### Results

Homes with no Pools: 100% Accuracy



# Детекция объектов на изображениях

[DEMO LINK](#)

1. Соберите изображения
2. Обведите интересующие объекты
3. Обучите
4. Тестируйте



Разметка



Распознавание



Обучающие наборы данных



## Аналитика в Ритейле

Отслеживание перемещения покупателей по магазину, идентификация мошеннических действий, идентификация отсутствия продуктов на полках

Генерация отчетов, пометка отсутствия объектов, индикация подозрительных действий



## Соблюдение требований безопасности

Использование камер наблюдения для обеспечения требований безопасности труда

Мониторинг зон, тепловые карты, обнаружение праздношатания



## Удаленный контроль и управление активами

Определение неисправного или изношенного оборудования в удаленных и труднодоступных местах

Оповещение о планировании работ по техническому обслуживанию и контроль безопасности инфраструктуры



## Охрана и Безопасность

Определение и отслеживание подозрительных объектов в целях повышения безопасности.

Распознавание лиц, оставленных без присмотра объектов, отслеживание траектории

## Удаленное инспектирование и мониторинг

Удаленная оценка потенциальных угроз и текущего состояния критически важных активов.

Распознавание объектов, анализ состояния активов, отслеживание вандализма и краж, обеспечение безопасности инфраструктуры

## Мониторинг зон отслеживания

Мониторинг периметров и запрещенных зон.

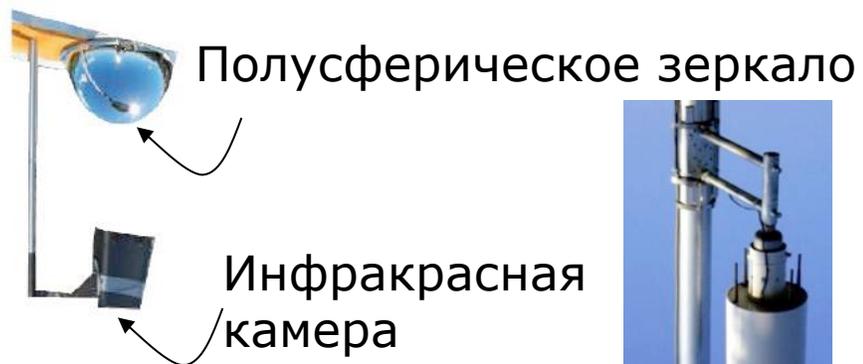
Отслеживание нарушений зоны, обнаружение вторжений, тепловые карты, праздношатание

## Распознавание дефектов

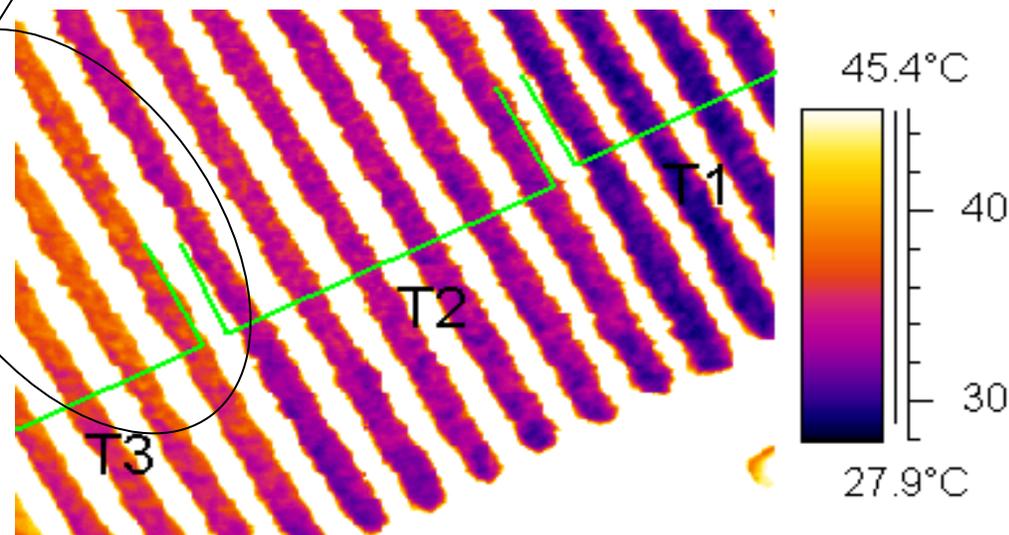


Дефекты солнечных панелей

## Анализ сельскохозяйственных посевов



Менее влажные участки



- Дроны могут использоваться для мониторинга посевов:
  - (1) Анализ почвы и полей;
  - (2) Мониторинг урожая;
  - (3) Анализ полива;
  - (4) Оценка состояния урожая.

## Анализ видео с дронов. Определение огня на видео.

- Дроны могут использоваться для мониторинга областей данных, но это требует ручного анализа видео.
- Разработчик может легко использовать PowerAI Vision, научить модель на поиск видеоучастков с огнем. Через API можно искать видеокadres, где был замечен огонь.



- Проект в Китае – Определение огня на видео с дронов
- Решение работает на PowerAI Vision

## Распознавание действий: драка, ходьба, бег и т.д.

№	Изображение с дракой	Визуализация работы нейронной сети	Результат
1			Fight: 1.00000
2			Fight: 1.00000
3			Fight: 1.00000

№	Изображение без драки	Визуализация работы нейронной сети	Результат
1			Not_Fight: 0.95200
2			Not_Fight: 1.00000
3			Not_Fight: 1.00000

Демонстрация распознавания активностей. Слева - тестирование изображений с драками. Справа – изображения без драк. В колонке «результат» представлены результаты распознавания.

**Спасибо!**

**IBM Client Center**

