

# Демократизация искусственного интеллекта и цифровая трансформация бизнеса

**Дмитрий Сошников**

Senior Technical Evangelist  
Microsoft

к.ф.-м.н., доцент  
МАИ, МФТИ, НИУ ВШЭ

# \$ whoami



## Технический евангелист

- Код
- Статьи
- Выступления



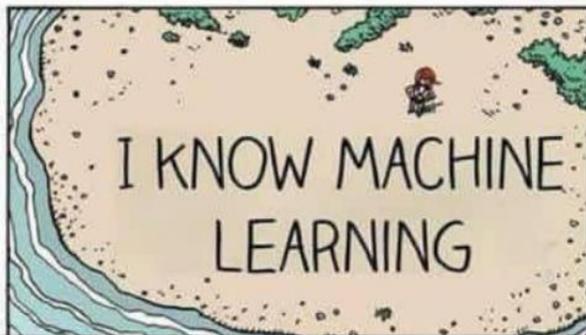
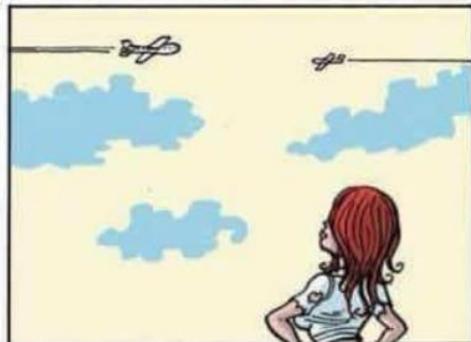
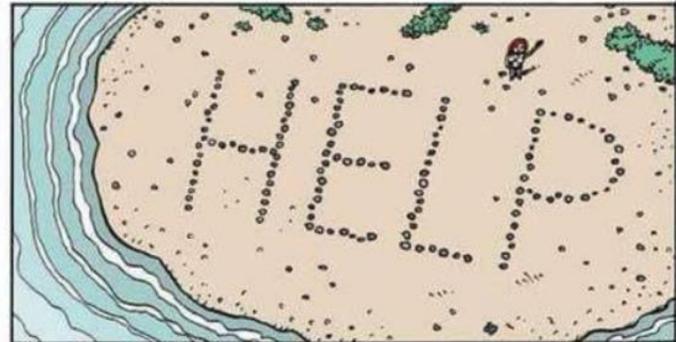
## Доцент МФТИ, ВШЭ, МАИ

- Функциональное программирование
- Искусственный интеллект



## Отец

- IT в образовании
- Как научить ребенка программировать



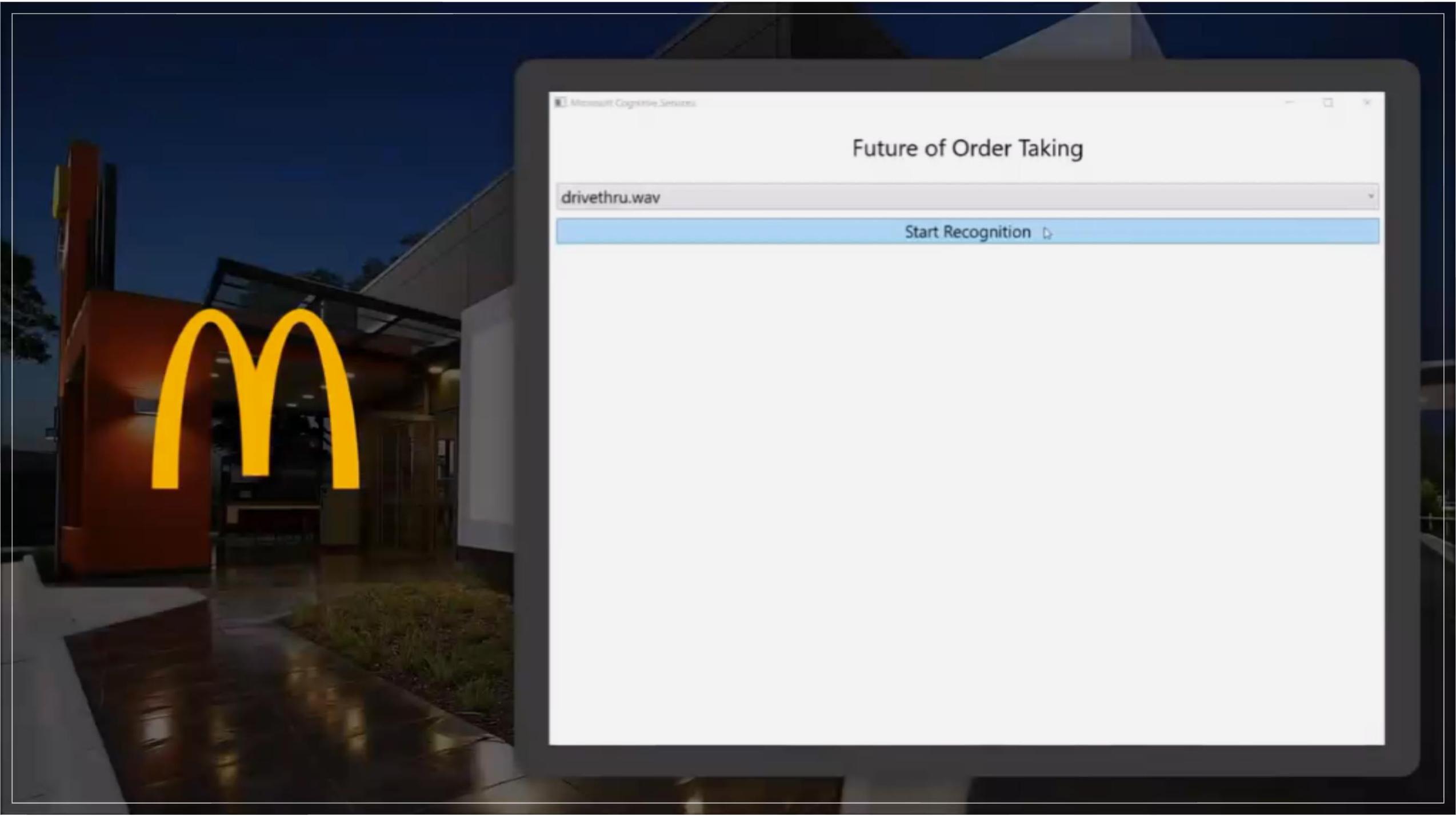




## Future of Order Taking

drivethru.wav

Start Recognition ▶

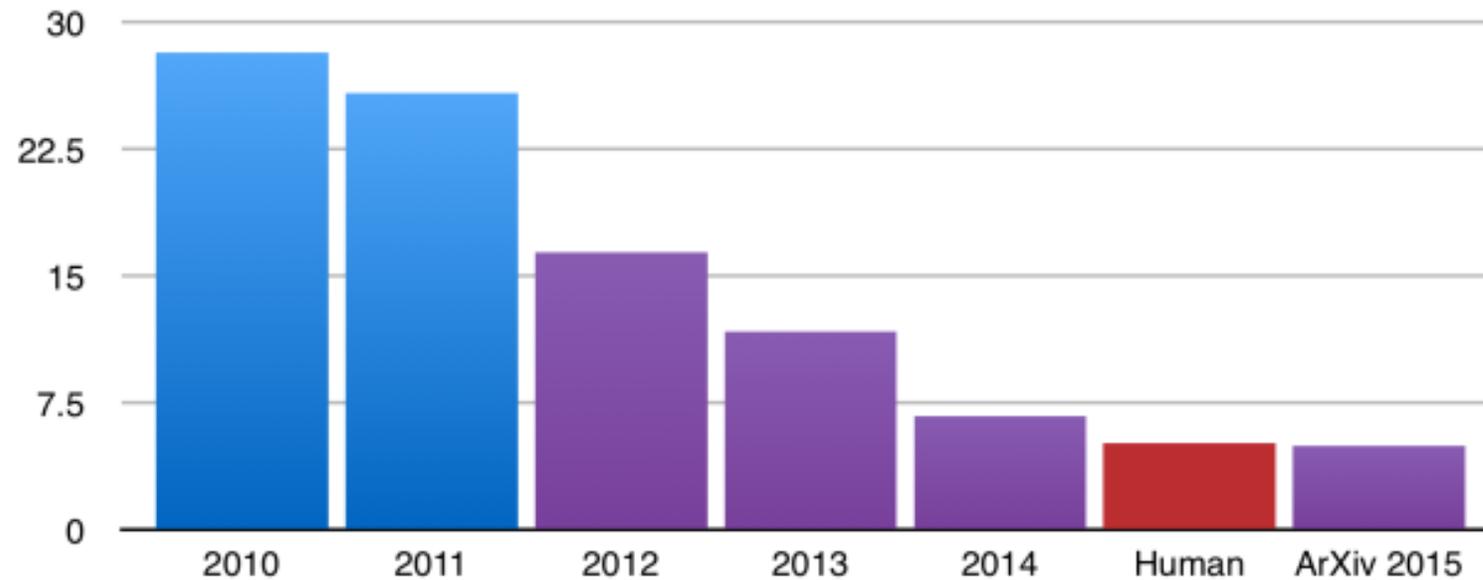


# Microsoft и AI

- Achieving human parity for Speech Recognition
  - <https://arxiv.org/abs/1610.05256>
- Winners in ImageNet 2015 Challenge
  - <http://image-net.org/challenges/LSVRC/2015/results>
  - Proposal of deep residual networks => human parity in image recognition
- Project Brainwave
  - Running AI on FPGA chips in the cloud
  - <https://www.microsoft.com/en-us/research/blog/microsoft-unveils-project-brainwave/>

# Распознавание изображений

ImageNet Large Scale Visual Recognition Challenge



Глубокое обучение + свёрточные сети



# YOLO v2

<http://pureddie.com/yolo>

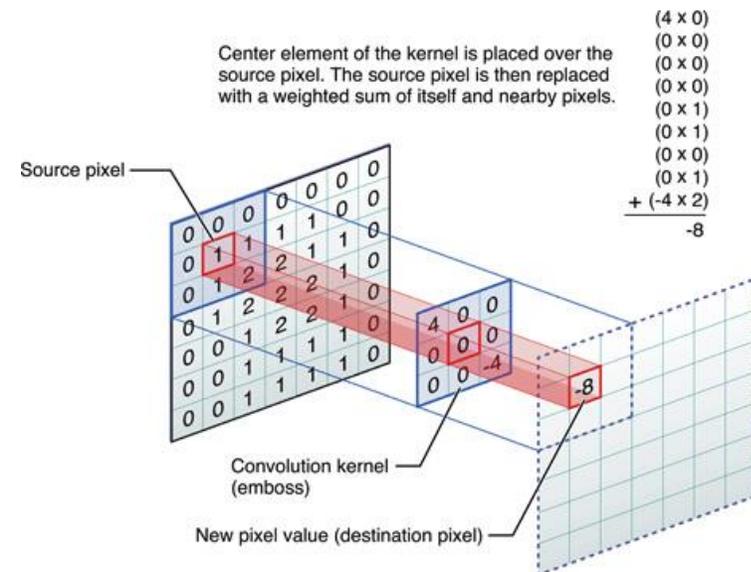
# Революция в области ИИ



Доступные  
вычислительные  
мощности



Открытые  
массивы  
данных



Глубокое  
обучение, новые  
архитектуры сетей

# Стратегия Microsoft в области ИИ



Демократизация  
ИИ



Внедрение ИИ  
в продукты



Исследования  
(FPGA в облаке, ..)

# Демонстрация

- В каких продуктах Microsoft есть ИИ?
  - Skype Translator
  - Powerpoint
  - Excel



# Демократизация искусственного интеллекта



Cognitive Services



Bot Framework

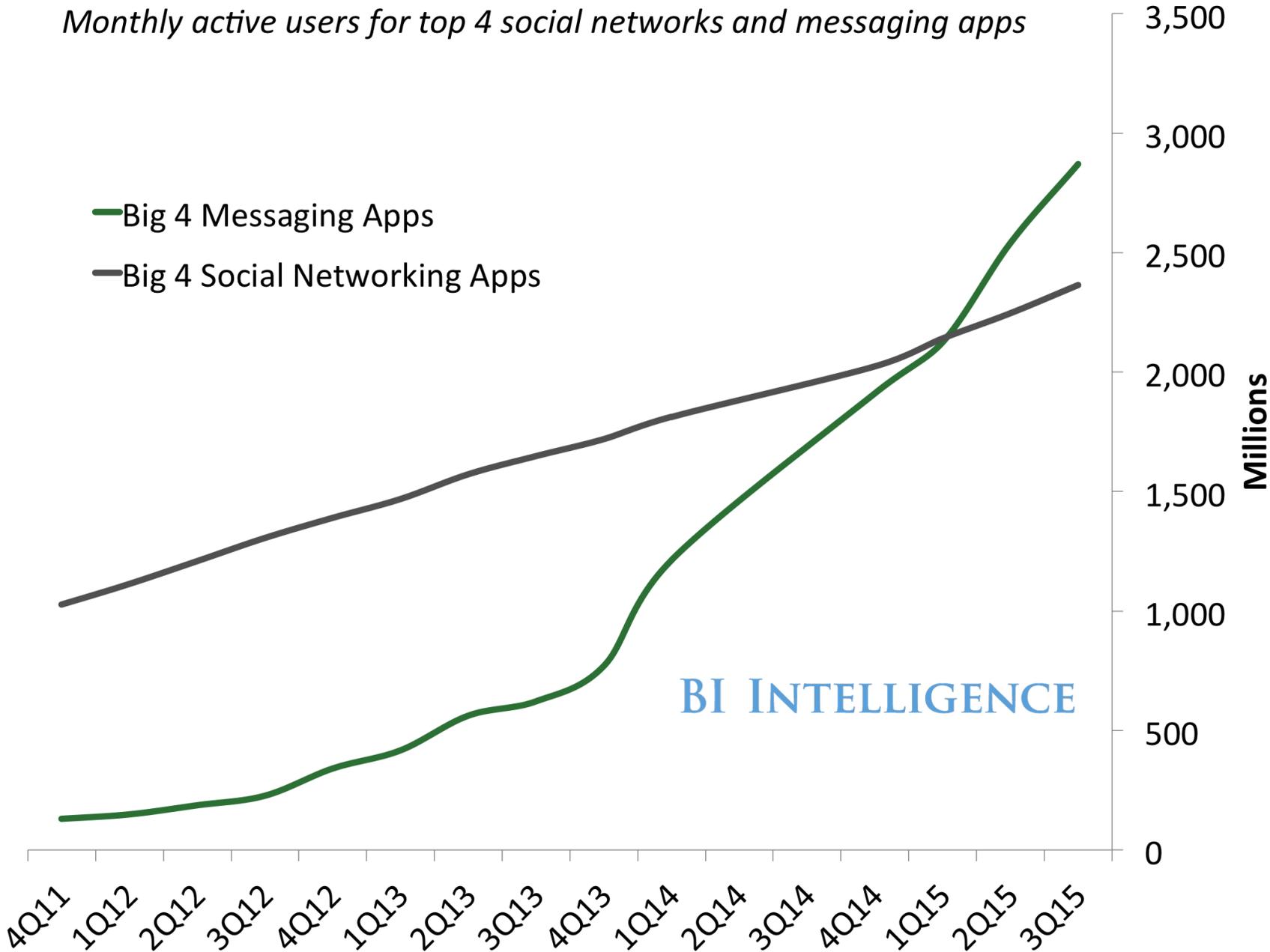


# Часть 1

## Боты и Microsoft Bot Framework

# Messaging Apps Have Surpassed Social Networks

Monthly active users for top 4 social networks and messaging apps



BI INTELLIGENCE

Source: Companies, BI Intelligence

**Бот:** Здравствуйте! Вы можете заказать пиццу!

**Я:** Прекрасно, хочу!

**Бот:** Скажите «заказать» чтобы приступить к заказу.

**Я:** Заказать.

**Бот:** Спасибо! Вы хотите закрытую или открытую пиццу?

**Я:** Открытую.

**Бот:** Я вас не понял. Скажите «открытую» (или 1) или «закрытую» (2)

**Я:** 1

**Бот:** Отлично! На тонком или толстом тесте?

**Я:** На тонком.

**Бот:** Вы хотите пиццу пепперони, капрезе, кваттро стаджони? Напишите «состав пепперони», чтобы узнать подробнее состав пиццы.

**Я:** пепперони.

**Бот:** А не хотите ли вместо обычного теста сделать слоёное?

...

 САЛЬМОН ДИ ГАРДЕН Цена за блюдо <b>1 260</b> руб. <a href="#">Заказать в 1 клик</a> ♥ <b>Заказать</b>	 СУПЕР ПЕППЕРОНИ Цена за блюдо <b>950</b> руб. <a href="#">Заказать в 1 клик</a> ♥ <b>Заказать</b>	 JACK'S СПЕШЛ Цена за блюдо <b>1 050</b> руб. <a href="#">Заказать в 1 клик</a> ♥ <b>Заказать</b>
 КАПРЕЗЕ	 КВАТТРО СТАДЖОНИ	 ПЕППЕРОНИ

# Боты как замена приложениям

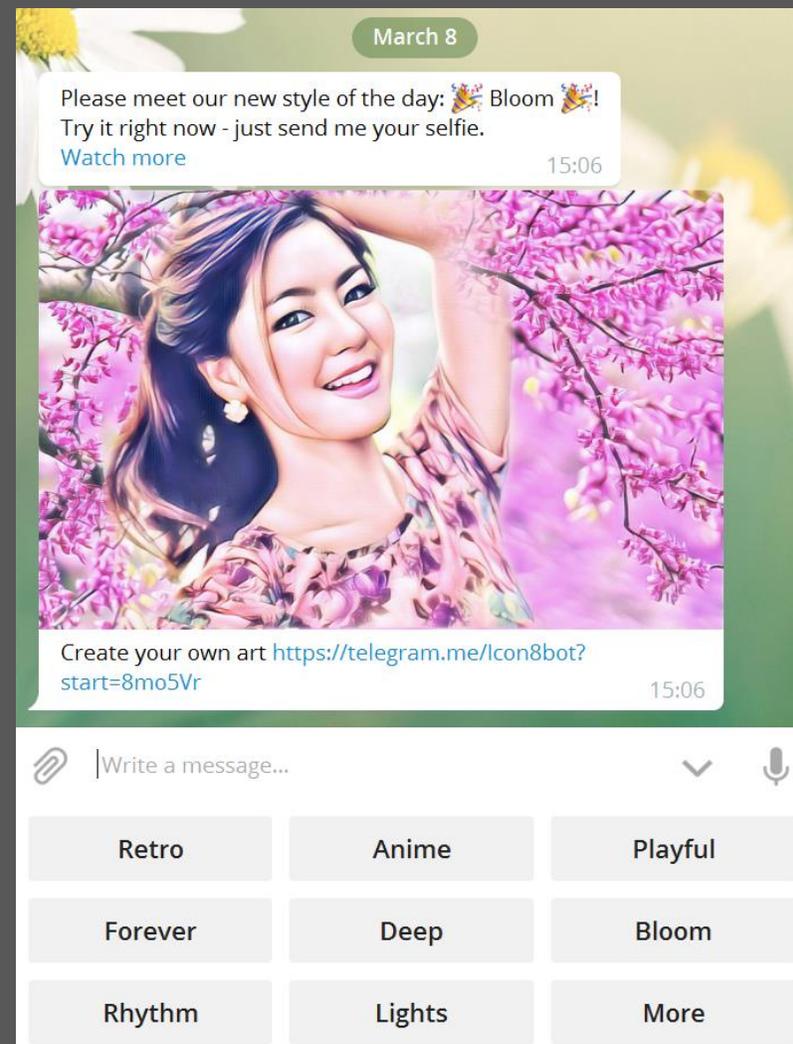
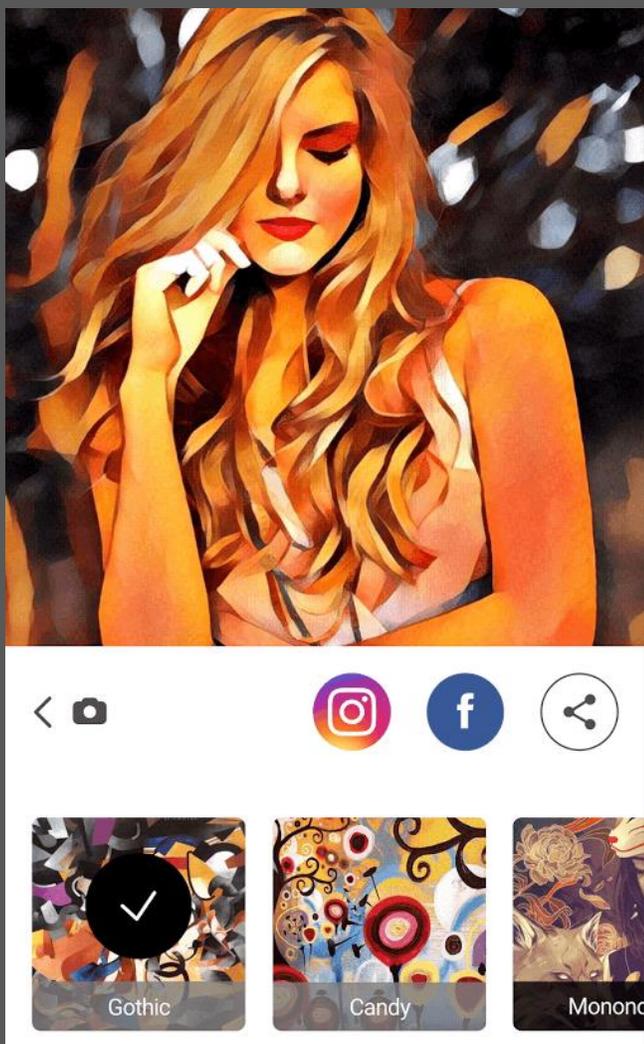


Проще разрабатывать  
(нет дизайна)

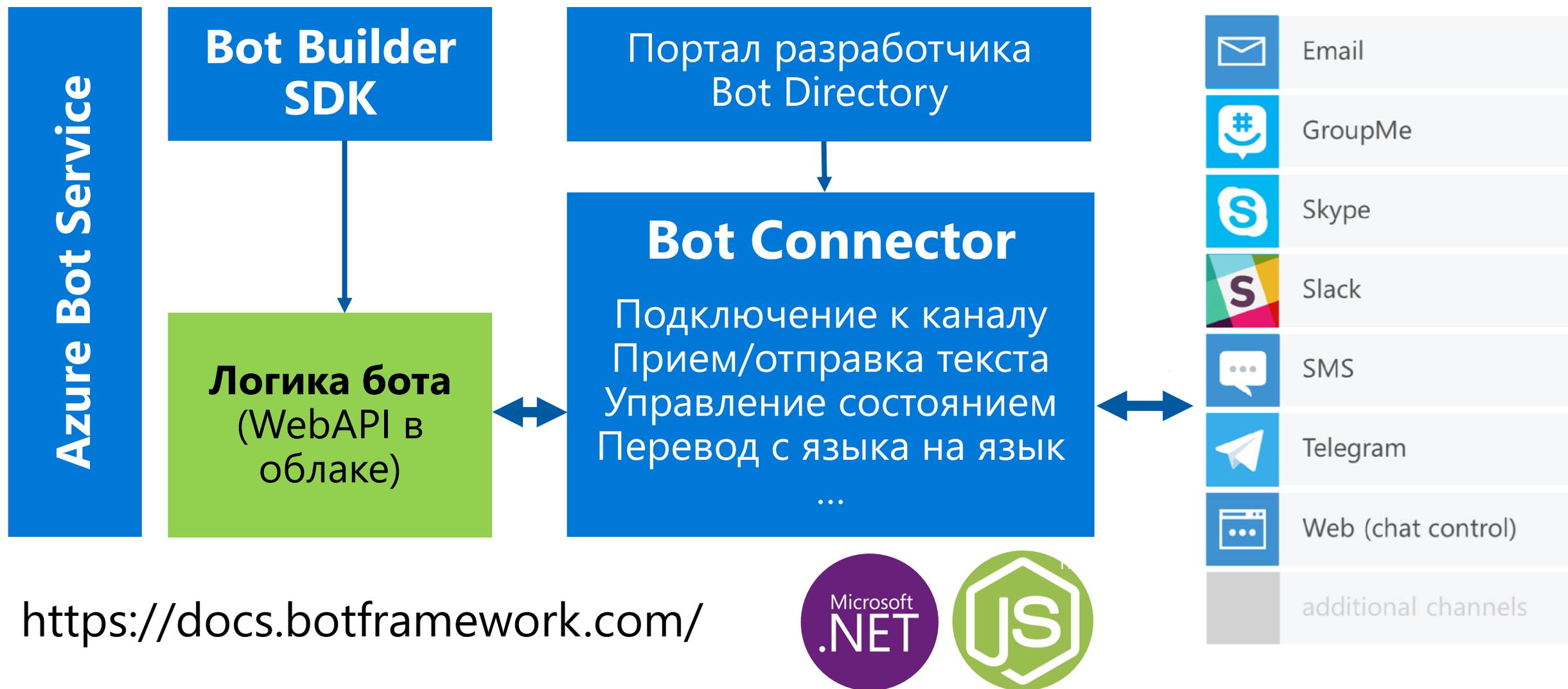


Проще использовать  
(нет инсталляции)

# Пример: Prisma vs. Icon8



# Microsoft Bot Framework



<https://docs.botframework.com/>

# ОСНОВНЫЕ ПОНЯТИЯ

Предложения (utterance)

Намерения (intent)

Hello!

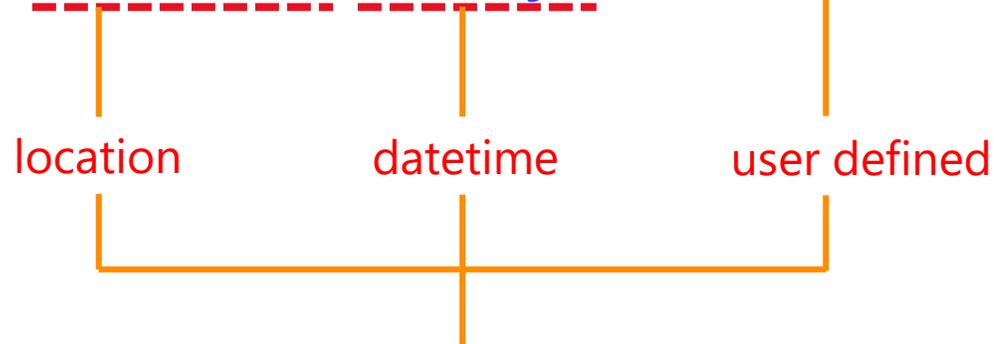


Greeting

What is the temperature  
in Moscow today?



Weather



Сущность (entity)

# Вавилонский бот

## ЗАДАЧА

Добавить в бот, основанный на распознавании намерений (intents), поддержку нескольких языков.

## БОЛЬ

Промежуточные сервисы (например, LUIS) поддерживает ограниченное число языков.

Прямая подстановка перевода дает плохие результаты.



# Вавилонский бот

## РЕШЕНИЕ

1. Использовать сервис для перевода, чтобы получить текст на поддерживаемом языке (английский).
2. Обучать сервис на кривом (!) английском.

// Оригинальный текст

אני רוצה לקבוע פגישה

אני רוצה לקבוע נסיעת מבחן

אני רוצה לקבוע נסיעת מבחן למחר

אפשר לקבוע נסיעת מבחן למחר?

// Корректный перевод

I want to schedule a meeting.

I want to schedule a test drive.

I want to schedule a test drive for tomorrow.

Can I schedule a test drive tomorrow?

// Кривой перевод

I want to schedule an appointment.

I want to schedule a test drive.

I want to make a test tomorrow.

Can set a test tomorrow?

**Часть 2:**

Когнитивные сервисы

# Vision API



a group of people sitting at a table

table, indoor, person, sitting, child, food, group, eating, laptop, computer, woman, people, front, small, boy, young, little, cake, man, desk, baby, room, plate, white

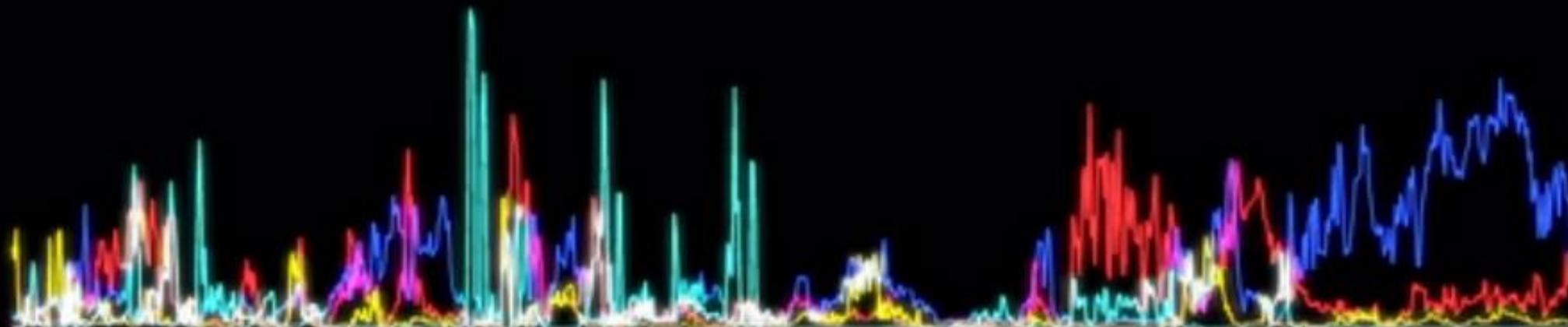
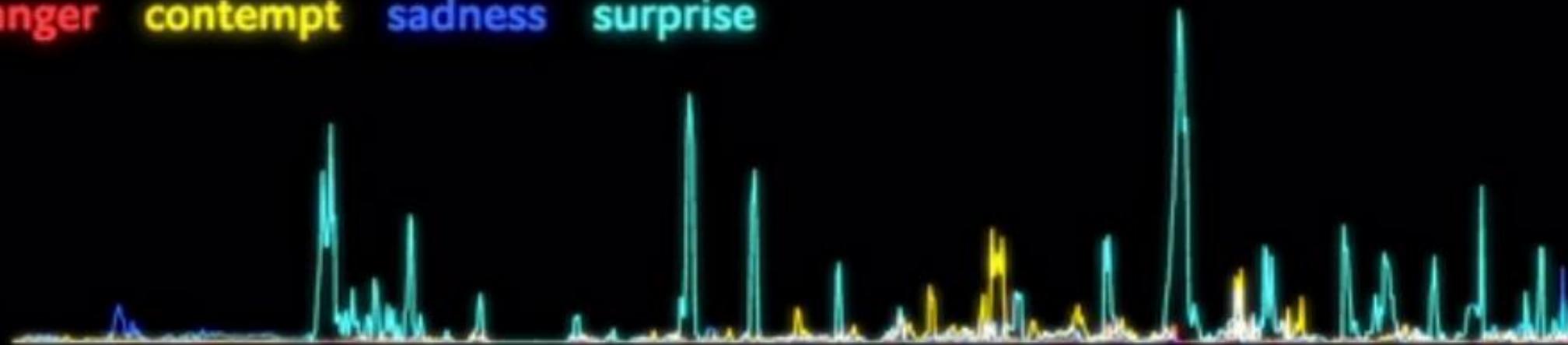


a man in a suit and tie

person, building, outdoor, man, cake, holding, carrying, standing, street, wearing, walking, young, food, woman, table, large, suit, city, people, riding



anger contempt sadness surprise





Intelligent Photo Booth

**Happiness**



**Sadness**



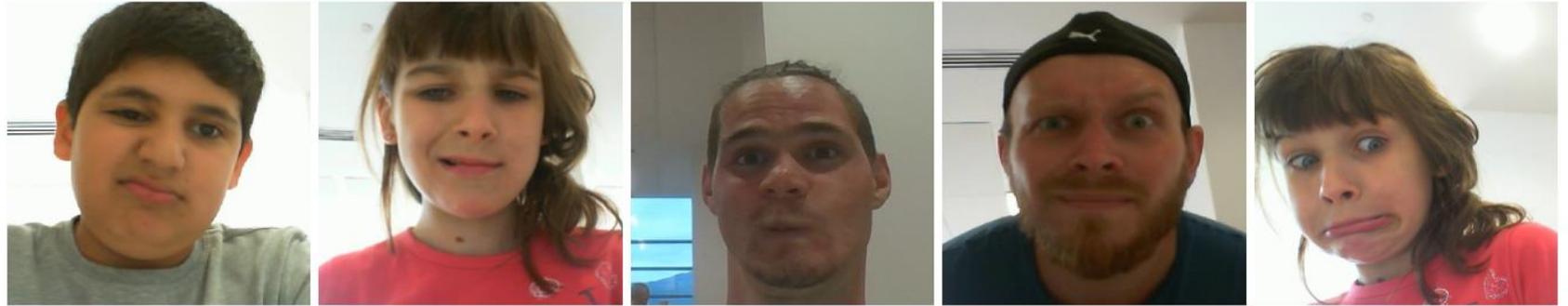
**Surprise**



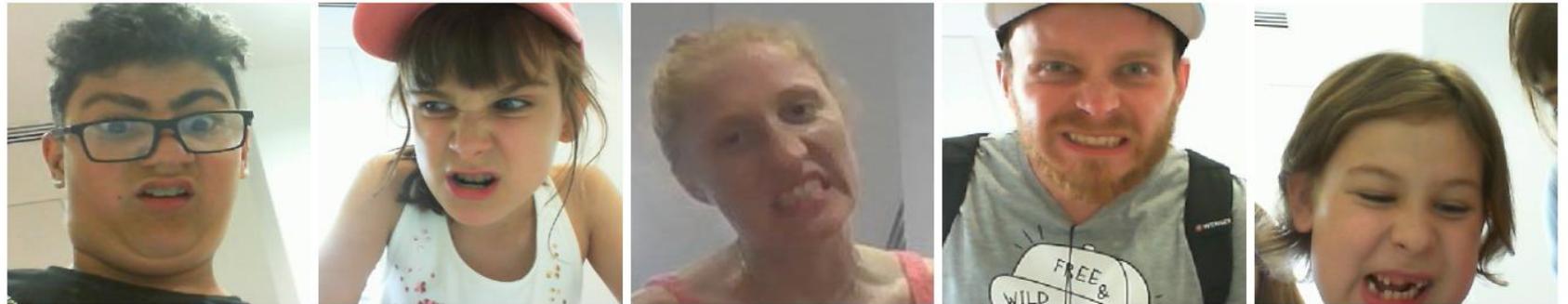
**Anger**



# Contempt



# Disgust

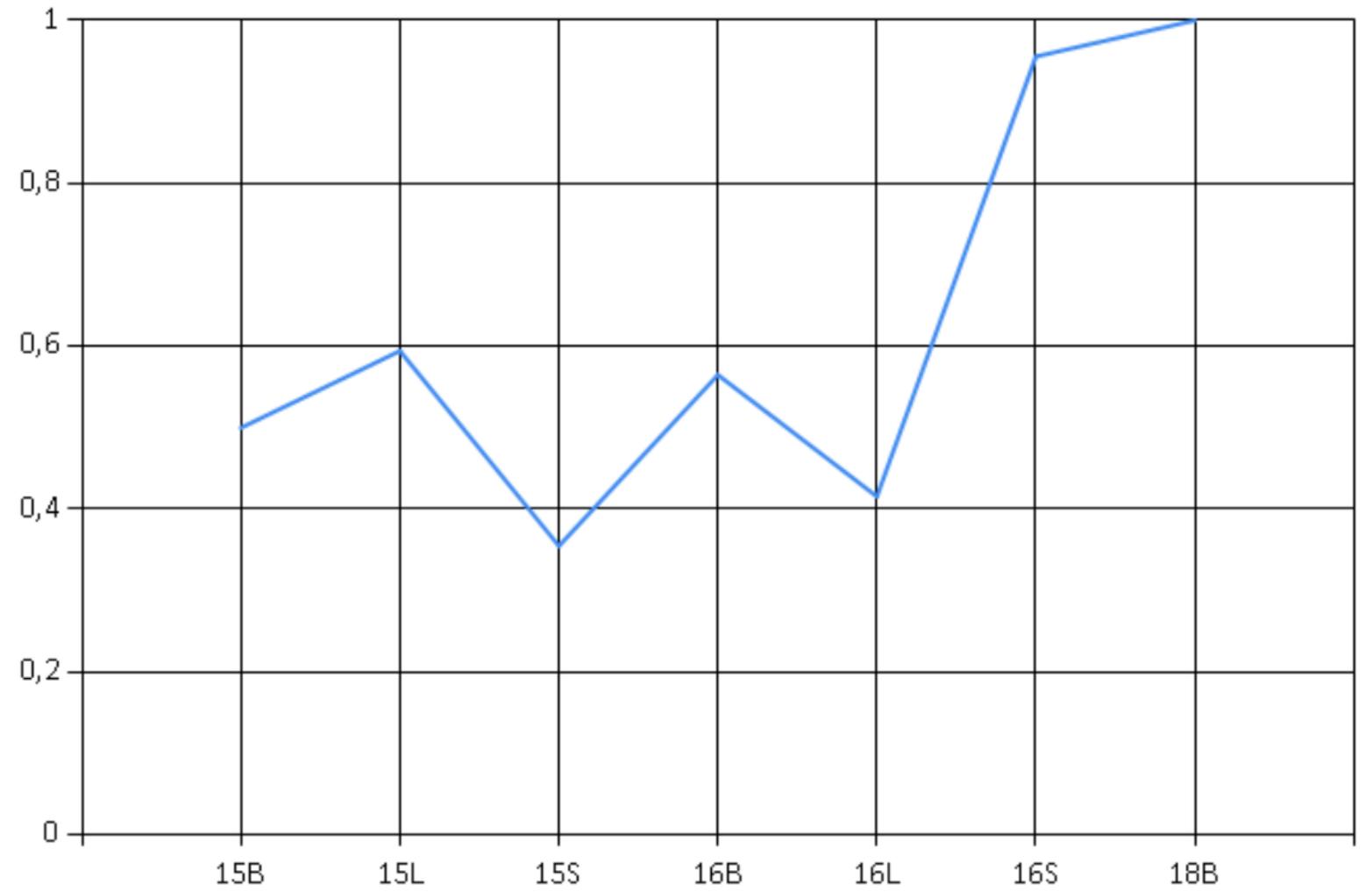


# Fear





# Happiness of Alex



# Success Story: Heedbook

## 1. Промо-экран со стерео микрофоном



### Высококачественная аудио и видео съемка

Мы работаем над выбором лучших поставщиков комплектующих, чтобы обеспечить лучшее качество видео и аудио потока. Устройство также оснащено стерео микрофоном (один - на столе сотрудника, второй - вмонтирован в корпус устройства) для четкого разделения речи фронт-менеджера и клиента.

### Демонстрация таргетированного медиаконтента

Сервис Heedbook позволяет выводить на дисплее устройства рекламу ваших товаров и услуг (или ваших партнеров) в соответствии с профилем сидящего клиента (пол и возраст). При интеграции Heedbook с вашим CRM вы также можете выводить на медиа экран целевое персональное предложение для клиента.

## 2. Мобильное приложение



### Возможности сервиса Heedbook + мобильность

С помощью мобильного приложения вы можете измерять и повышать качество сервиса и кросс-продаж у мобильных служб вашей компании - DSA, службы доставки, продавцов в торговом зале, службы охраны

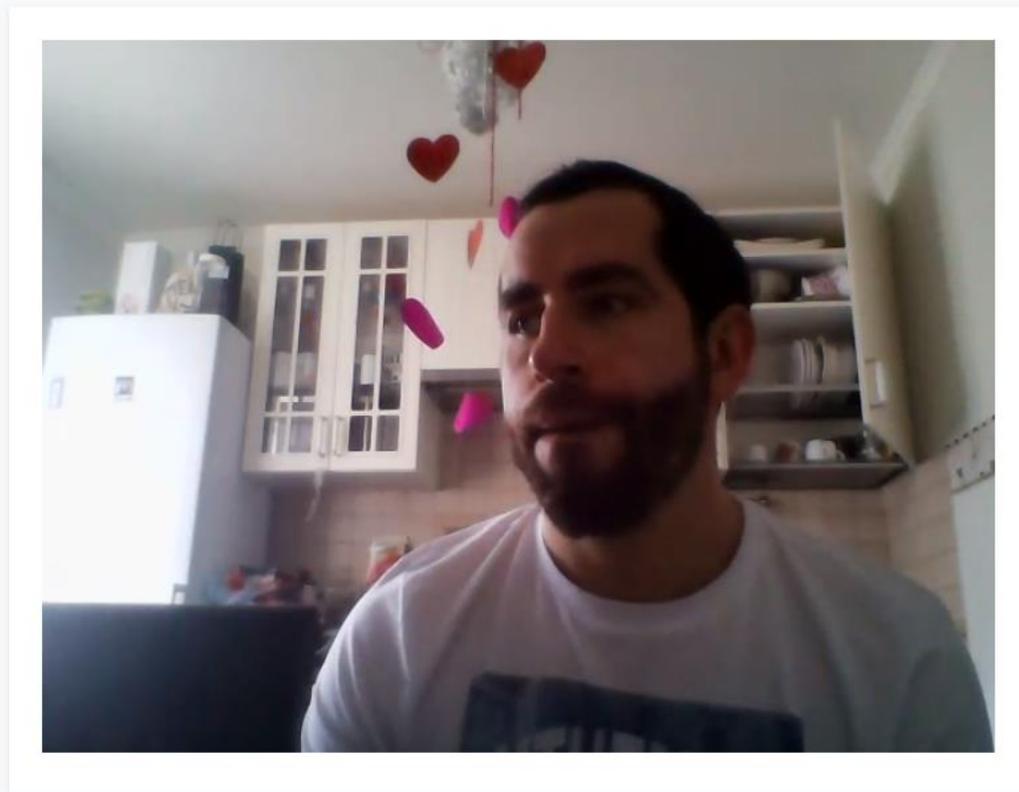
### Работает в фоновом режиме

Как и десктоп сервис, мобильное приложение работает в фоновом режиме, не мешая вашему сотруднику использовать смартфон для работы с другими программами.

Функционирует при минимальной скорости интернет (от 1Мб/с). Работает с большим числом внешних камер. Установка и использование мобильного приложения бесплатны, вы оплачиваете только время работы системы.

- Online
- Staff
- Dialogues
- Ratings
- Script
- FAQ

Dialogue video



Dialogue details

Front manager  
**Maslov Alexey**  
Start time  
**4/18/2017 3:58:44 PM**  
Duration  
**00:01:06**

Client profile



Gender  
**male**  
Age  
**34**

Dialogue indicators



Dialogue text

Привет она мне во всём Привет она мне во всём

Script usage indicators



- Online
- Staff
- Dialogues
- Ratings
- Script
- FAQ

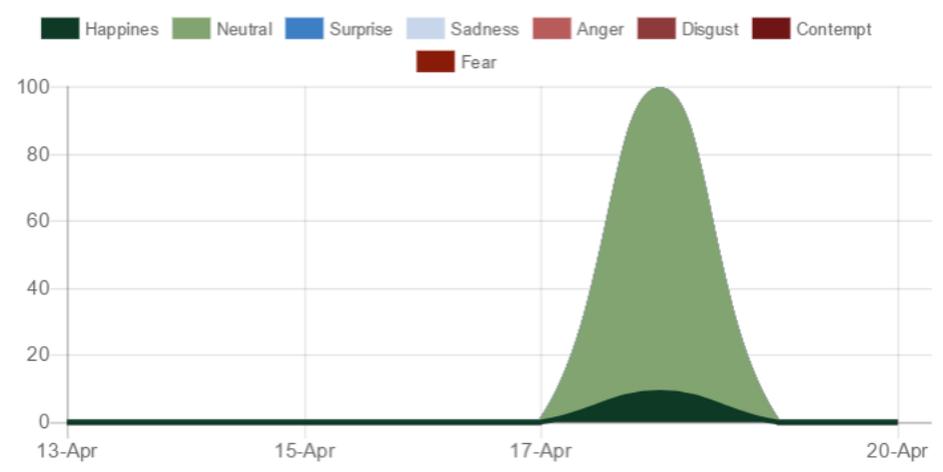
## Key indicators

Period: 04/13/2017 - 04/20/2017

### Client emotions

**9%**  
client happiness

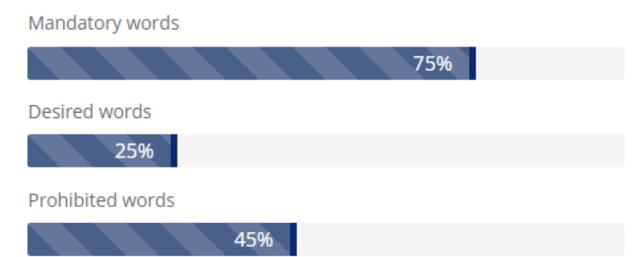
**1%**  
client frustration



### Client attention



### Script usage



### Dialogue duration

**1.1**  
minutes average dialogue duration



### Front office rating



# Microsoft Cognitive Services

<http://microsoft.com/cognitive>

REST, C#, Node.js, ...

## Cognitive Services API



### Vision

Computer Vision | Emotion | Face | Video | Custom Vision



### Speech

Bing Speech | Custom Recognition | Speaker Recognition



### Language

Bing Spell Check | Language Understanding  
Linguistic Analysis | Text Analytics | Translator |  
Web Language Model



### Knowledge

Recommendations | Academic Knowledge | Entity  
Linking | Knowledge Exploration | QnA Maker |  
Custom Decision



### Search

Bing Autosuggest | Bing Image Search | Bing  
News Search | Bing Video Search | Bing Web  
Search

# Распознавание изображений





DEMO: Custom Vision

# Полезные фоточки еды

## ЗАДАЧА

Добавить в приложение для трекинга питания возможность определения еды.

## БОЛЬ

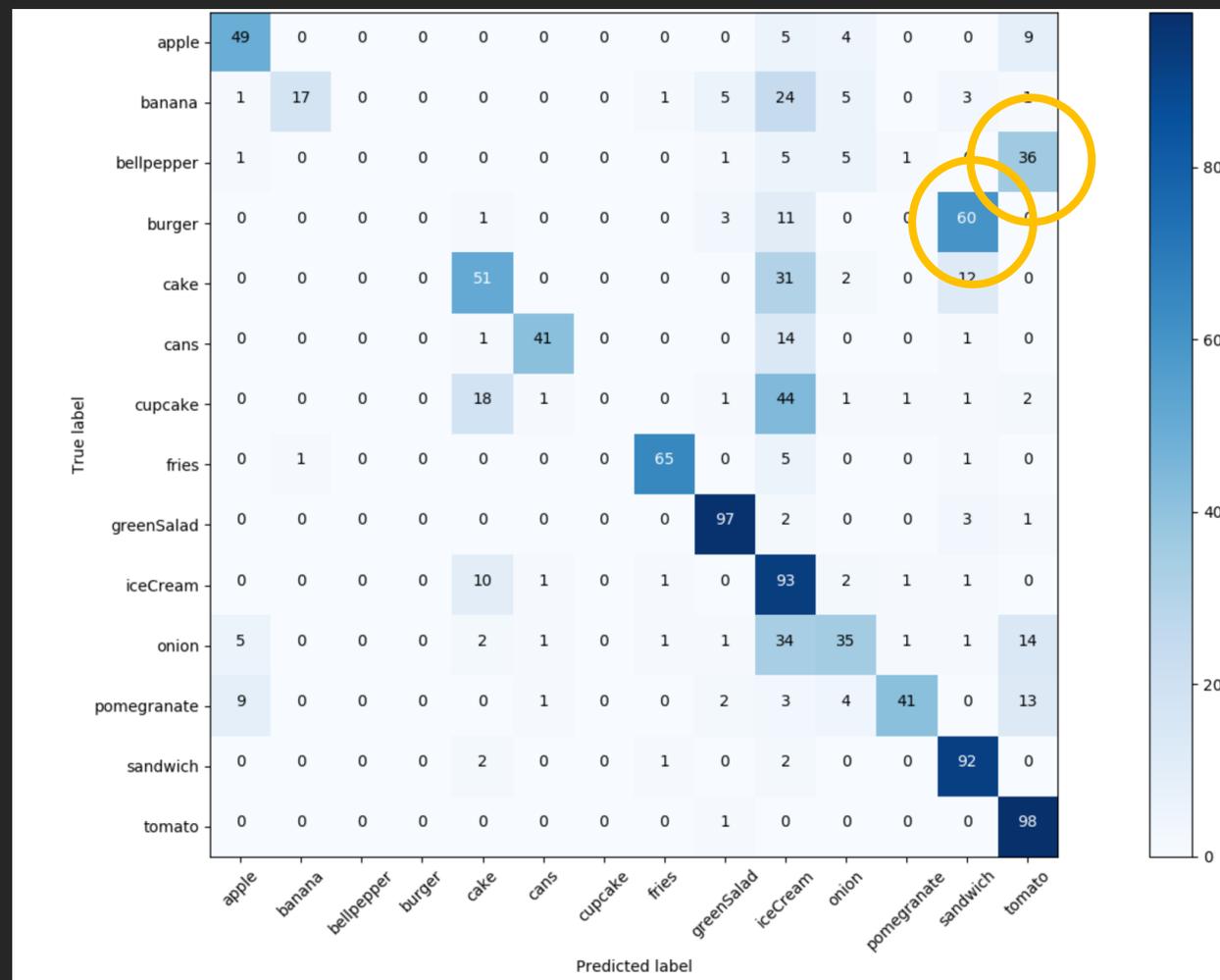
Разработчик – эксперт в мобильных приложениях, а не в обучении нейронных сетей.



# Полезные фоточки еды

## РЕШЕНИЕ

1. Собрать фоточки из Flickr или Instagram по релевантным тегам.
2. Использовать готовый дообучаемый сервис распознавания изображений (например, *Custom Vision* из наших Cognitive Services).



# Полезные фоточки еды

## РЕШЕНИЕ

1. Собрать фоточки из Flickr или Instagram по релевантным тегам.
2. Использовать готовый дообучаемый сервис распознавания изображений (например, *Custom Vision* из наших Cognitive Services).
3. Использовать микро-модели для сложных случаев.



**Часть 3:**

Машинное обучение

# Какие задачи решаются машинным обучением

- Есть много обучающих примеров
  - 10-100 x количество входных признаков – для классических моделей
  - 10-100К – глубокие нейросетевые модели
- Данные структурированы
  - можно выделить входные / выходные признаки
- Имеется зависимость выходных данных от ВХОДНЫХ
  - Пример: как имя влияет на одиночество

# Инструменты Microsoft



Azure ML Studio



Azure Notebooks

Cognitive Toolkit (CNTK)  
LightGBM  
MML Spark



Azure ML Service  
Azure ML Workbench

Beginner

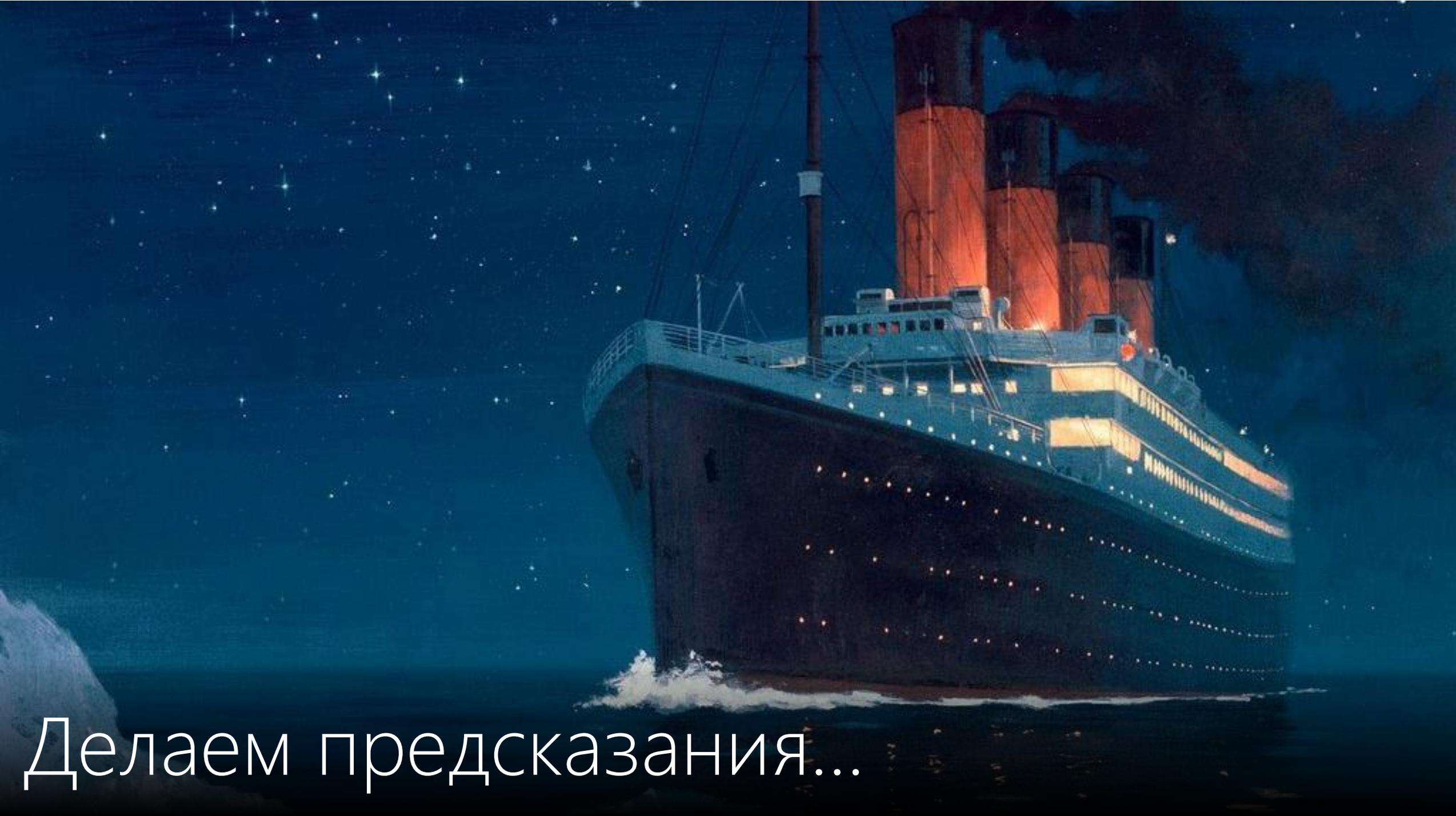
Professional



Azure Data Science VM



Azure Batch AI



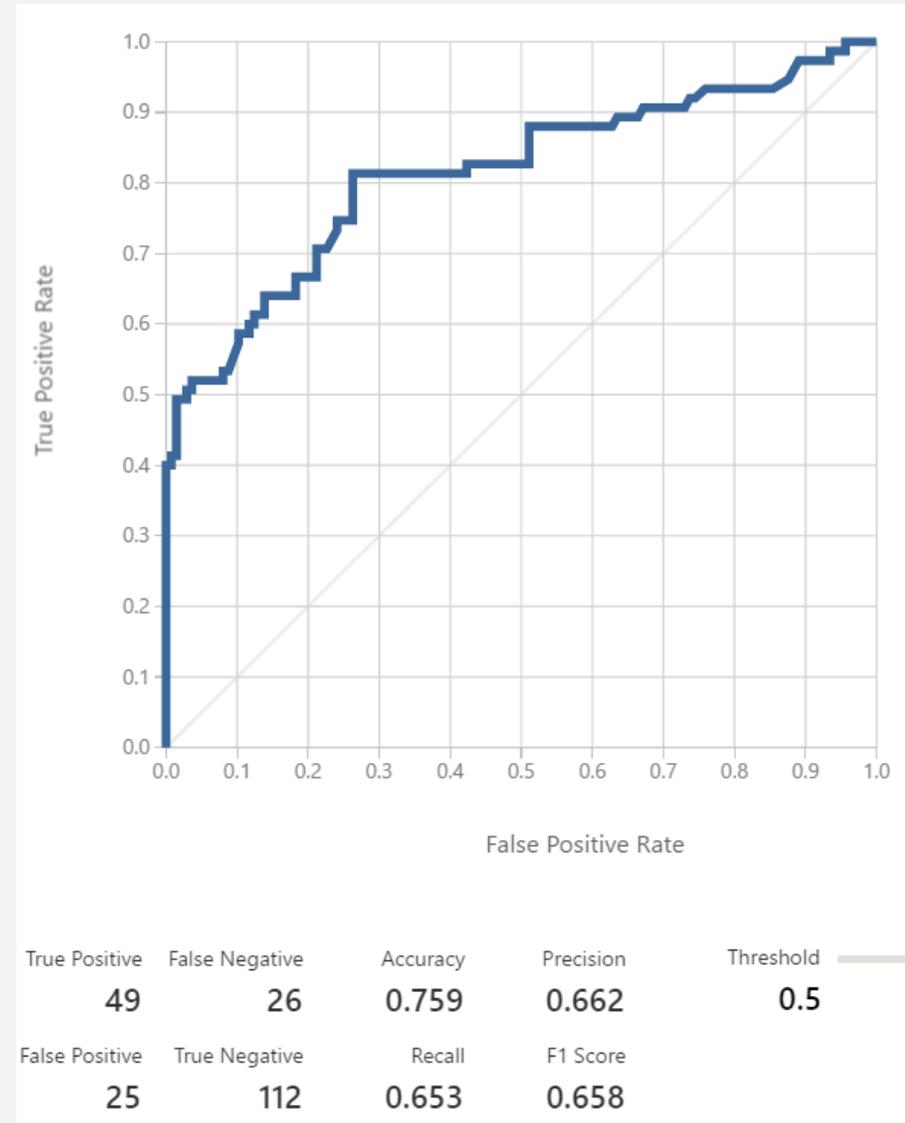
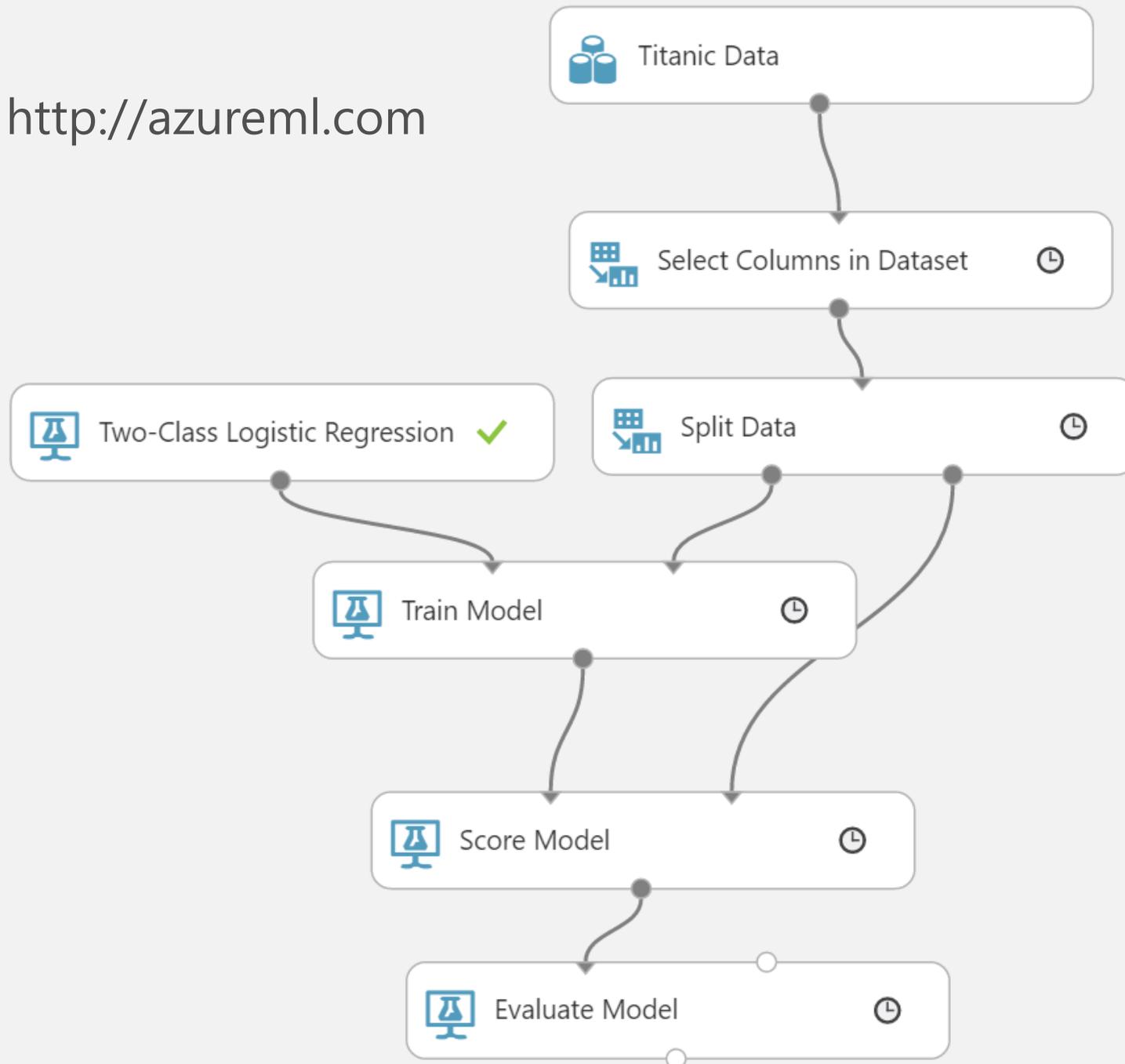
Делаем предсказания...

# Предсказываем вероятность гибели на титанике

<http://www.soshnikov.com/temp/titanic.csv>

pclass	survived	name	gender	age	sibsp	parch
1	1	Allen, Miss. Elisabeth Walton	female	29	0	0
1	1	Allison, Master. Hudson Trevor	male	0,9167	1	2
1	0	Allison, Miss. Helen Loraine	female	2	1	2
1	0	Allison, Mr. Hudson Joshua Creighton	male	30	1	2
1	0	Allison, Mrs. Hudson J C (Bessie Waldo Daniels)	female	25	1	2

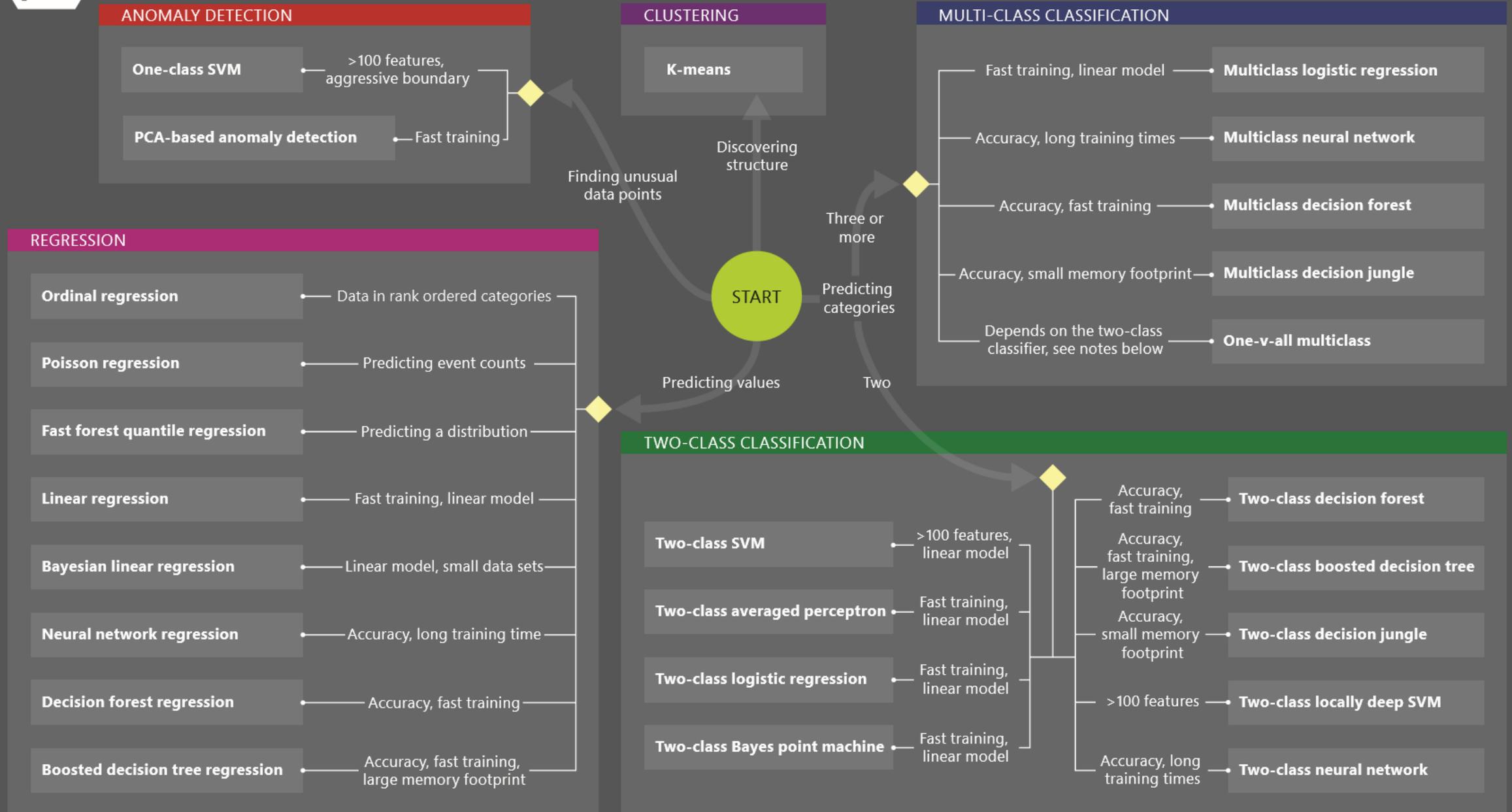
<http://azureml.com>





# Microsoft Azure Machine Learning: Algorithm Cheat Sheet

This cheat sheet helps you choose the best Azure Machine Learning Studio algorithm for your predictive analytics solution. Your decision is driven by both the nature of your data and the question you're trying to answer.



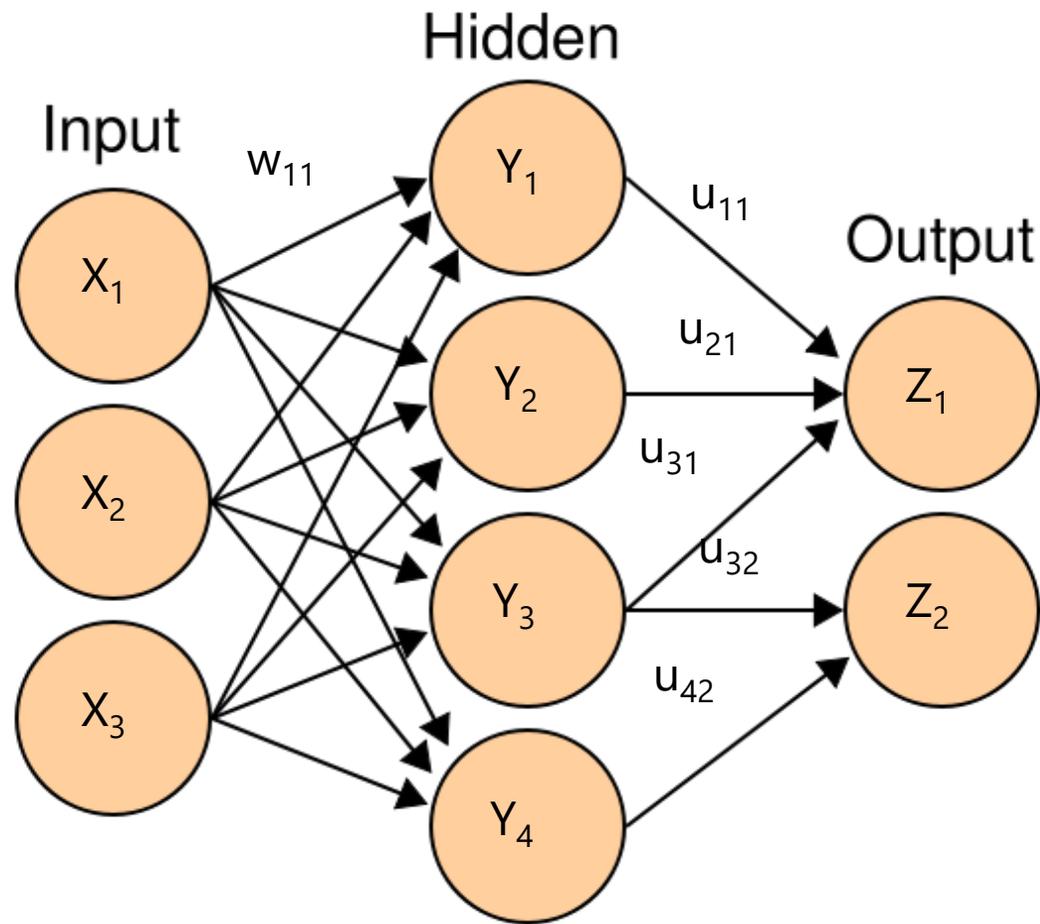
**Часть 5:**

Нейронные сети

# Самое важное про нейронные сети

- В последние годы нейросети считают синонимом ИИ
- Всё, что вы видели в когнитивных сервисах – нейронные сети
- Нейросеть – это не магия, а всего лишь способ оптимизации функций
- Часто работа нейросети выглядит как магия
- Для серьезных экспериментов нужны большие данные и вычислительные ресурсы, которые есть в облаке!

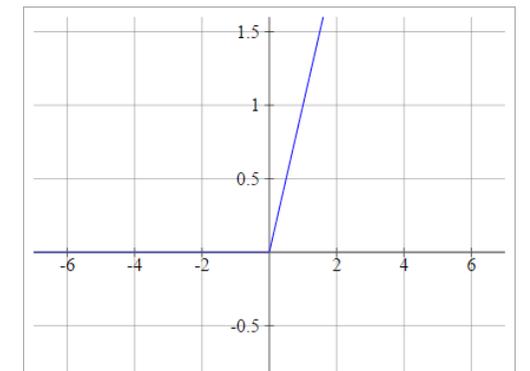
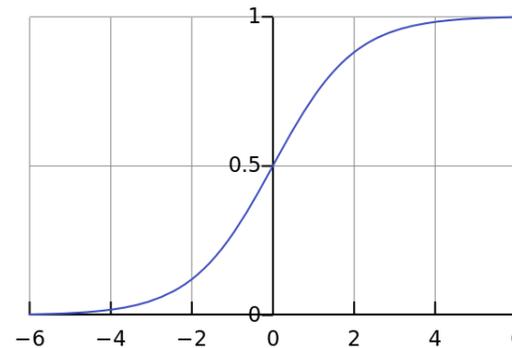
# Нейронные сети



$$Y_k = f \left( \sum_{i=1}^N w_{ik} X_i + b_k \right)$$

$$Y = f(WX + B)$$

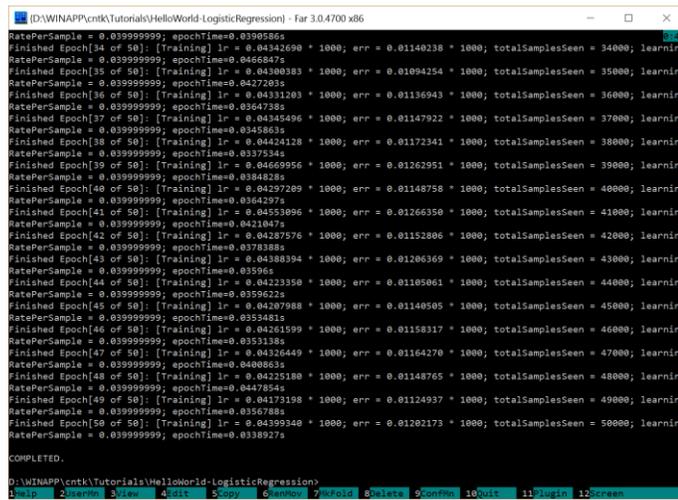
$$Z = f(UY + C)$$



# Основные задачи и архитектуры сетей

Задача	Класс сетей
Обычная классификация / регрессия	Feed-forward / Fully Connected (DNN)
Распознавание изображений	Convolutional / Свёрточная (CNN)
Выделение объектов на изображении	Fast R-CNN (Region CNN)
Анализ последовательностей (речь, текст)	Рекуррентная (RNN)
Машинный перевод	Sequence-to-Sequence, RNN+RNN
Описание изображения	CNN + RNN
Стилизация изображения	Style Transfer

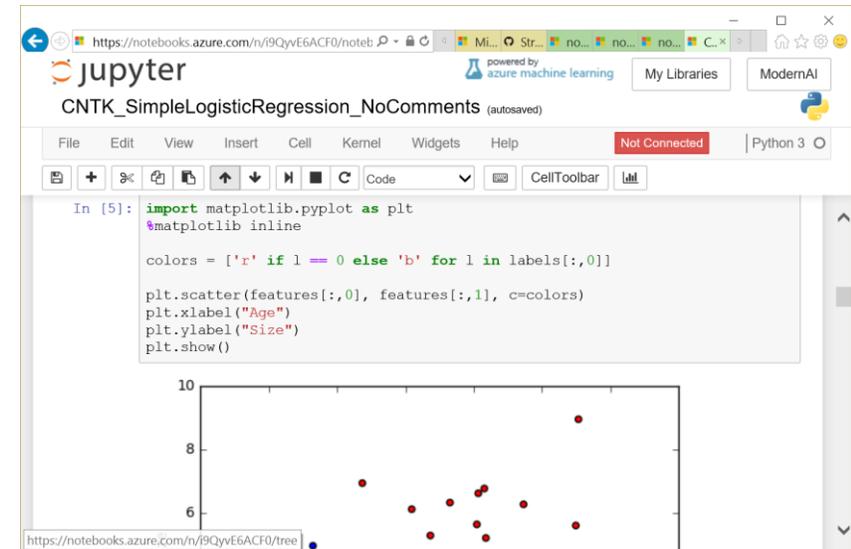
# Microsoft Cognitive Toolkit CNTK



```
[D:\WINAPP\cntk\tutorials\HelloWorld-LogisticRegression] - Far 3.0.4700 x86
RatePerSample = 0.039999999; epochTime=0.0390586s
Finished Epoch[34 of 50]: [Training] lr = 0.04342690 * 1000; err = 0.01140238 * 1000; totalSamplesSeen = 34000; learning
RatePerSample = 0.039999999; epochTime=0.0466047s
Finished Epoch[35 of 50]: [Training] lr = 0.04300383 * 1000; err = 0.01094254 * 1000; totalSamplesSeen = 35000; learning
RatePerSample = 0.039999999; epochTime=0.0427203s
Finished Epoch[36 of 50]: [Training] lr = 0.04311203 * 1000; err = 0.01136943 * 1000; totalSamplesSeen = 36000; learning
RatePerSample = 0.039999999; epochTime=0.0364728s
Finished Epoch[37 of 50]: [Training] lr = 0.04345496 * 1000; err = 0.01147922 * 1000; totalSamplesSeen = 37000; learning
RatePerSample = 0.039999999; epochTime=0.0345863s
Finished Epoch[38 of 50]: [Training] lr = 0.04242128 * 1000; err = 0.01172341 * 1000; totalSamplesSeen = 38000; learning
RatePerSample = 0.039999999; epochTime=0.0337534s
Finished Epoch[39 of 50]: [Training] lr = 0.04669956 * 1000; err = 0.01262951 * 1000; totalSamplesSeen = 39000; learning
RatePerSample = 0.039999999; epochTime=0.0384828s
Finished Epoch[40 of 50]: [Training] lr = 0.04297209 * 1000; err = 0.01148758 * 1000; totalSamplesSeen = 40000; learning
RatePerSample = 0.039999999; epochTime=0.0364297s
Finished Epoch[41 of 50]: [Training] lr = 0.04553096 * 1000; err = 0.01266350 * 1000; totalSamplesSeen = 41000; learning
RatePerSample = 0.039999999; epochTime=0.0412047s
Finished Epoch[42 of 50]: [Training] lr = 0.04297576 * 1000; err = 0.01152806 * 1000; totalSamplesSeen = 42000; learning
RatePerSample = 0.039999999; epochTime=0.0378388s
Finished Epoch[43 of 50]: [Training] lr = 0.04388394 * 1000; err = 0.01206369 * 1000; totalSamplesSeen = 43000; learning
RatePerSample = 0.039999999; epochTime=0.03596s
Finished Epoch[44 of 50]: [Training] lr = 0.04223350 * 1000; err = 0.01105061 * 1000; totalSamplesSeen = 44000; learning
RatePerSample = 0.039999999; epochTime=0.0359622s
Finished Epoch[45 of 50]: [Training] lr = 0.04297988 * 1000; err = 0.01140505 * 1000; totalSamplesSeen = 45000; learning
RatePerSample = 0.039999999; epochTime=0.0354818s
Finished Epoch[46 of 50]: [Training] lr = 0.04261199 * 1000; err = 0.01158317 * 1000; totalSamplesSeen = 46000; learning
RatePerSample = 0.039999999; epochTime=0.0353138s
Finished Epoch[47 of 50]: [Training] lr = 0.04326449 * 1000; err = 0.01164270 * 1000; totalSamplesSeen = 47000; learning
RatePerSample = 0.039999999; epochTime=0.0400863s
Finished Epoch[48 of 50]: [Training] lr = 0.04251100 * 1000; err = 0.01148765 * 1000; totalSamplesSeen = 48000; learning
RatePerSample = 0.039999999; epochTime=0.0447854s
Finished Epoch[49 of 50]: [Training] lr = 0.04173198 * 1000; err = 0.01124937 * 1000; totalSamplesSeen = 49000; learning
RatePerSample = 0.039999999; epochTime=0.0367804s
Finished Epoch[50 of 50]: [Training] lr = 0.04390240 * 1000; err = 0.01202173 * 1000; totalSamplesSeen = 50000; learning
RatePerSample = 0.039999999; epochTime=0.0338927s
COMPLETED.
```

- Утилита для тренировки сетей **cntk.exe**
- Библиотека для использования (C#, C++)
- Язык описания сетей BrainScript
- Надо устанавливать CNTK на ПК (с GPU или без)
- Спец. входной формат для данных

- Интерфейс с Python для обучения и использования сетей
- Конфигурация сети описывается программой на Python
- Можно использовать Azure Notebook или Jupyter Notebook на машине с GPU



# Подсчет моевок на Аляске

## ЗАДАЧА

Наблюдение за красноногими моевками (они же чайки) на побережье Аляски – оценка размера популяции по видео.

## БОЛЬ

Нет готовых датасетов и моделей.



# Подсчет моевок на Аляске

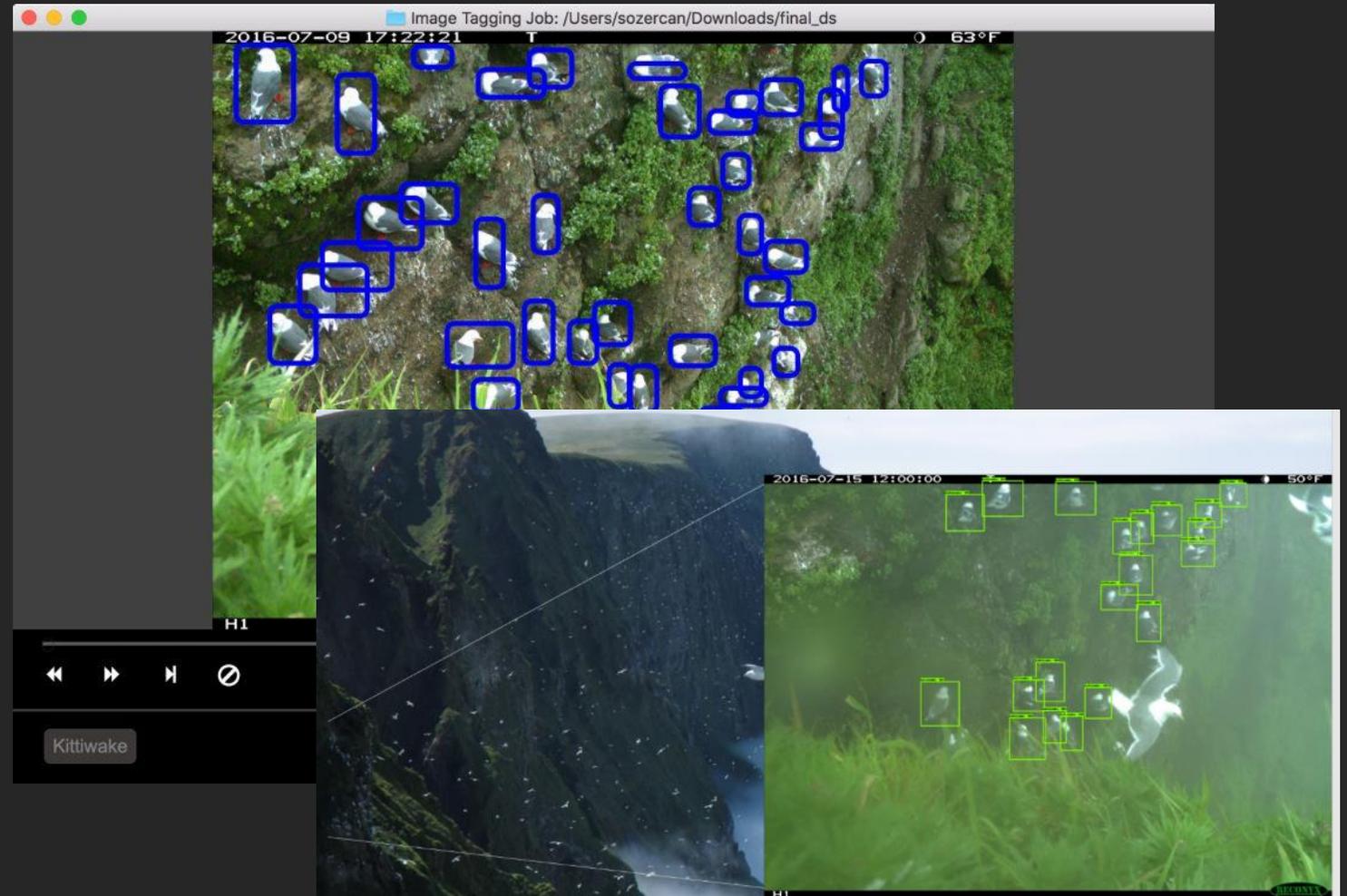
## РЕШЕНИЕ

Использовать специальный инструмент для разметки фото/видео по классам.  
(20 часов ручной работы)

Экспорт данных для Faster R-CNN (CNTK или TensorFlow).

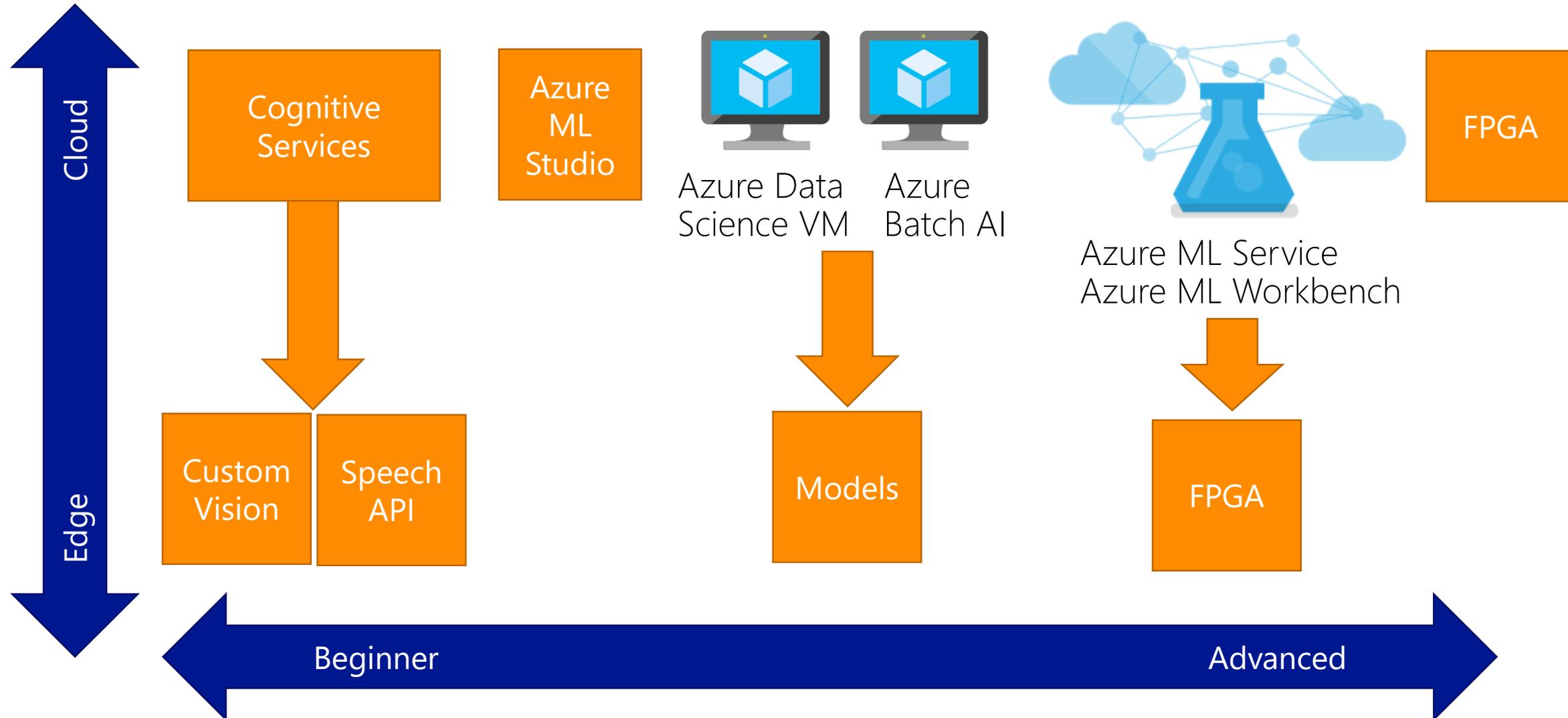
VOTT:

<https://github.com/CatalystCode/VoTT/releases>



# Microsoft AI Technologies

LightGBM

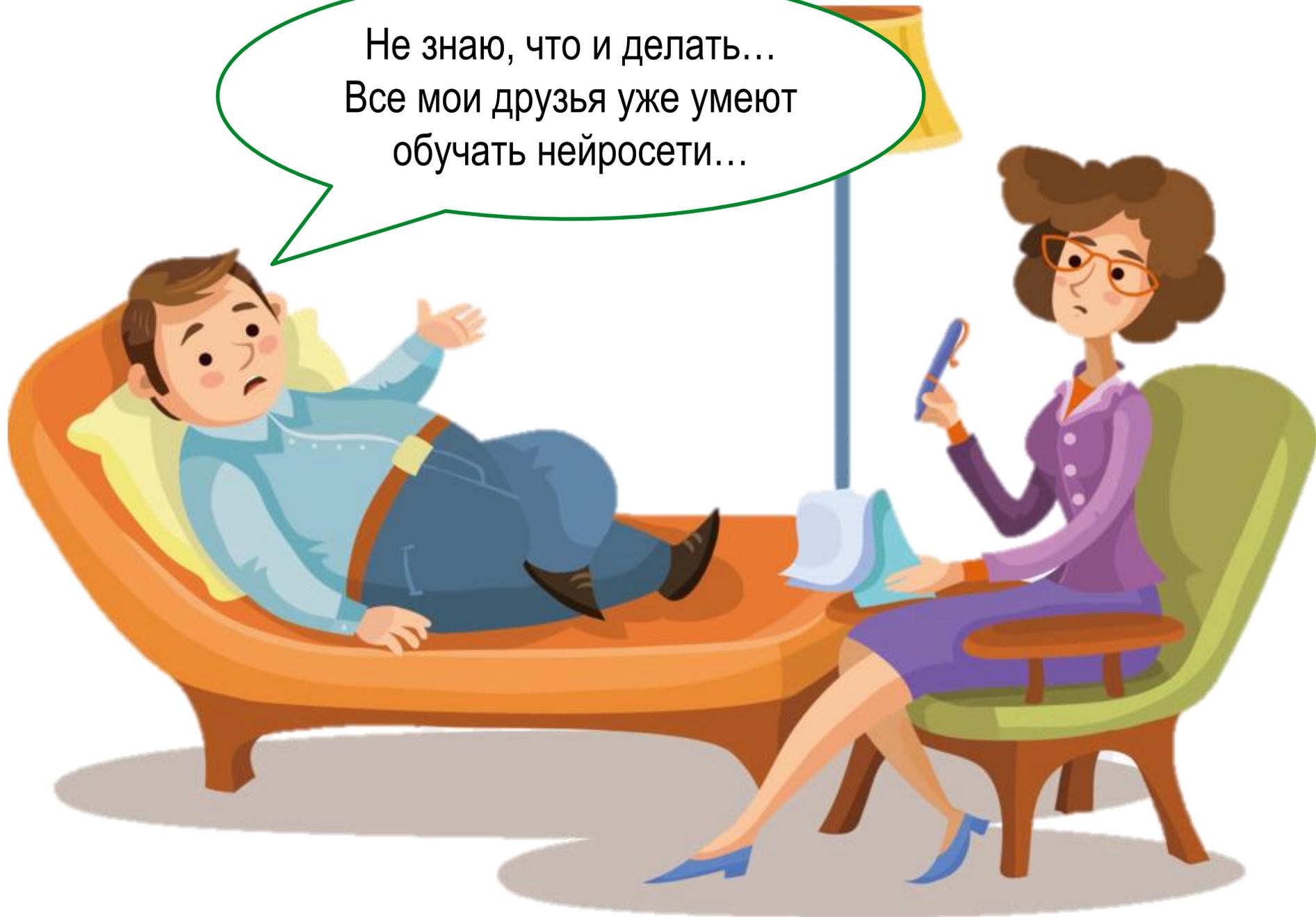


**1** Мы живем в период стремительных изменений. ИИ – один из основных драйверов.

**2** Происходит демократизация ИИ – начать использовать ИИ в проектах совсем не сложно

**3** Посмотрите на Azure Machine Learning и машинное обучение

Не знаю, что и делать...  
Все мои друзья уже умеют  
обучать нейросети...



🔗 Q&A

# Microsoft и искусственный интеллект

Дмитрий Сошников

<http://vk.com/shwars>

