

# Artificial intelligence for humans...





All resources:  
[aka.ms/human-ai](https://aka.ms/human-ai)



Let's talk about  
"Artificial Intelligence"



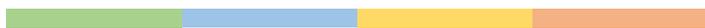
# Artificial Intelligence

---

- Is **nothing new** – the concepts go back to the 50ies
- Is quite the **hype** and very often **misattributed**
- Is an **umbrella term** for a lot of math and science around repetition, pattern recognition and machine learning
- Got a **huge boost** because of availability of **hardware**



CGP Grey:  
"How machines  
learn"

A screenshot of a YouTube video player. The video thumbnail shows a cartoon robot character standing next to a real bee. There are two blue folders labeled '3's' and 'bees' on the left. A subtitle at the bottom of the video reads "only the kinds of questions it's been taught to." Below the video, the title "How Machines Learn" is visible, along with "2,331,981 views", "217K" likes, "1.5K" comments, and "SHARE" options.

How Machines Learn  
2,331,981 views  
217K 1.5K SHARE

[youtube.com/watch?v=R9OHn5ZF4Uo](https://youtube.com/watch?v=R9OHn5ZF4Uo)



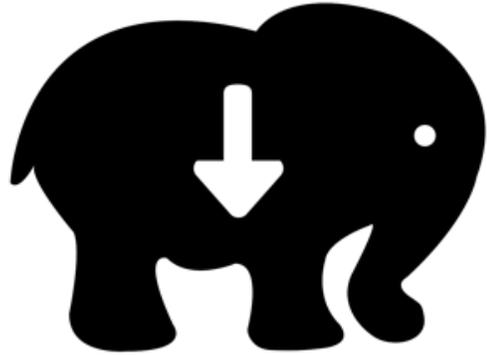
The machines  
are watching...



Big brother is  
redundant...



- Everything we do online is **monitored** and **recorded**
- We often don't realise that our **data** is how we **pay** for "free" services
- We're happy to use systems that **record** all the time **in exchange** for convenience
- Often people don't realise just how **dangerous** this can be in the **wrong** hands.



Everything counts  
in large amounts



- We create a **massive** amount of information – actively and without our knowledge.
- It is tough to make that amount of information **consumable** again.
- That's why we have **computers**
- With cloud computing, on demand processing and advances in hardware **we're faster than ever.**



## Leaving invisible marks...



- By using **other people's machines and infrastructure**, we leave traces
- This allows companies to **recognise** us, and accumulates a **usage history**
- This leads to **better results**, but can leak data
- We should have more **transparency** about what digital legacy we left behind.



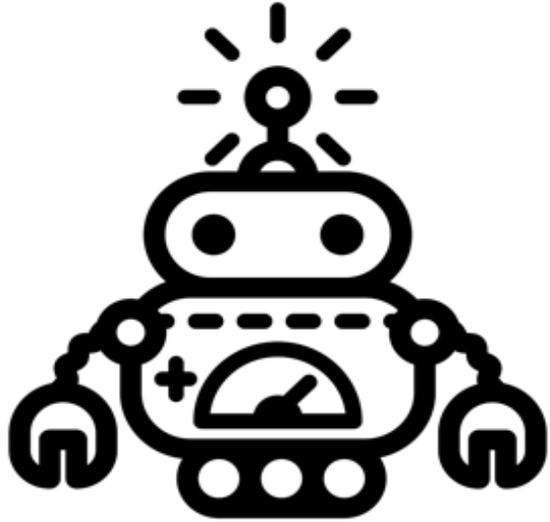
Are machines  
friend or foe?



# Artificial Intelligence Myths



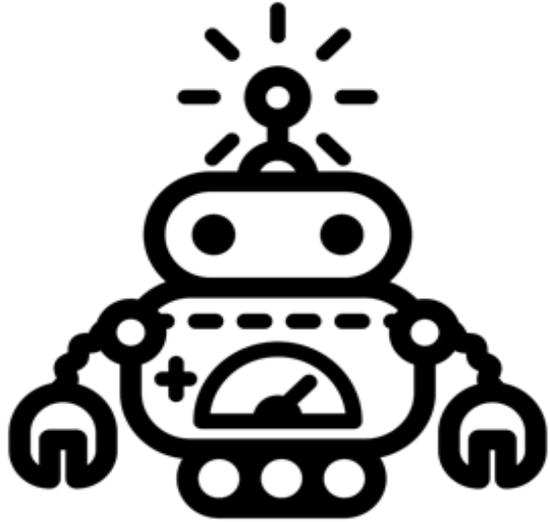
- AI can't replace a thinking, creative human
- AI can not magically fill gaps with perfect information – it can only compare and assume
- AI doesn't learn in a creative fashion. It makes **no assumptions**
- AI has **no morals and ethics**, but – used wrongly – it can **amplify our biases**



Machines can be  
great tools or  
weapons...

---

- Machine Learning is all about returning assumptions
- We don't get any definitive **truth** from **algorithms**, we get answers to our questions
- AI can answer questions, but it is up to you to **ask good questions** – generic questions yield assumed results.



Machines can be  
great tools or  
weapons...

---

- Untrained and limited data leads to terrible and biased AI results
- It is very easy to get either **wrong deductions** or **false positives**
- AI is as **intelligent** and **good** as the **people** who **apply** it

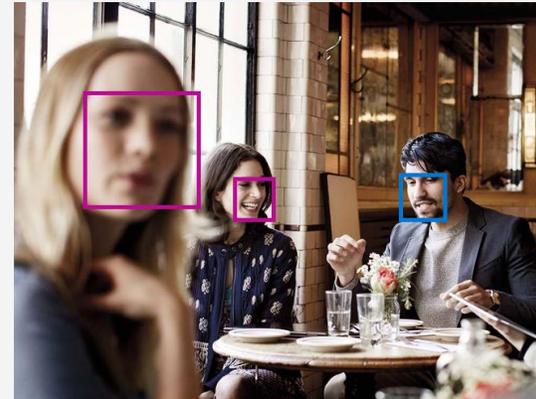


About face...



## Face detection

Detect one or more human faces in an image and get back face rectangles for where in the image the faces are, along with face attributes which contain machine learning-based predictions of facial features. The face attribute features available are: Age, Emotion, Gender, Pose, Smile, and Facial Hair along with 27 landmarks for each face in the image.



```
Detection result:
JSON:
[
  {
    "faceId": "0e96b668-a0d0-46ec-a5c9-ad26b16a1ca9",
    "faceRectangle": {
      "top": 166,
      "left": 128,
      "width": 218,
      "height": 218
    },
    "faceAttributes": {
      "hair": {
        "bald": 0.0,
        "invisible": false,
        "hairColor": [
          {
            "color": "blond",
            "confidence": 1.0
          },
          {
            "color": "other".
```

Image URL

Submit

Browse





## About face...



- Face rectangle / Landmarks
- Pose (pitch/roll/yaw)
- Smile
- Gender/Age
- Type of glasses
- Makeup (lips/eye)
- Emotion (anger, contempt, disgust, fear, happiness, neutral, sadness, surprise)
- Occlusion (forehead/eye/mouth)
- Facial hair (moustache/beard/sideburns)
- Attributes: Hair (invisible, bald, colour)

# Is this you? Are those also you?

## Face verification

Check the likelihood that two faces belong to the same person. The API will return a confidence score about how likely it is that the two faces belong to one person.

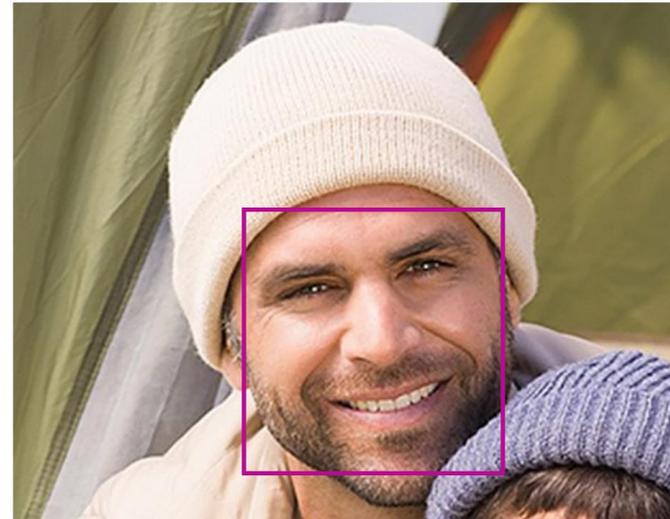


Image URL

Submit

Browse

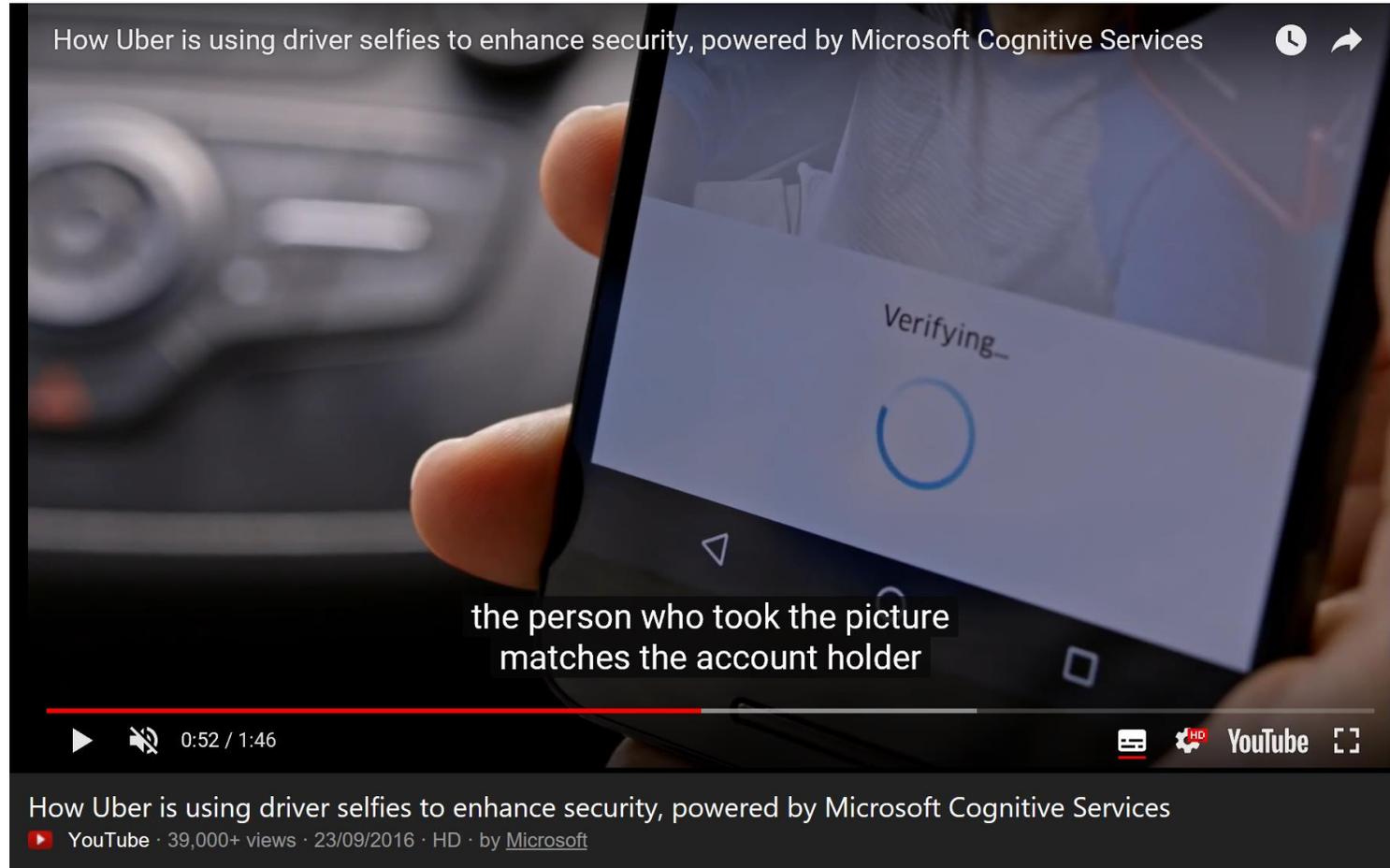
Image URL

Submit

Browse

Verification result: The two faces belong to the same person. **Confidence is 0.7349.**

# Is this your driver?



# Taking it too far?

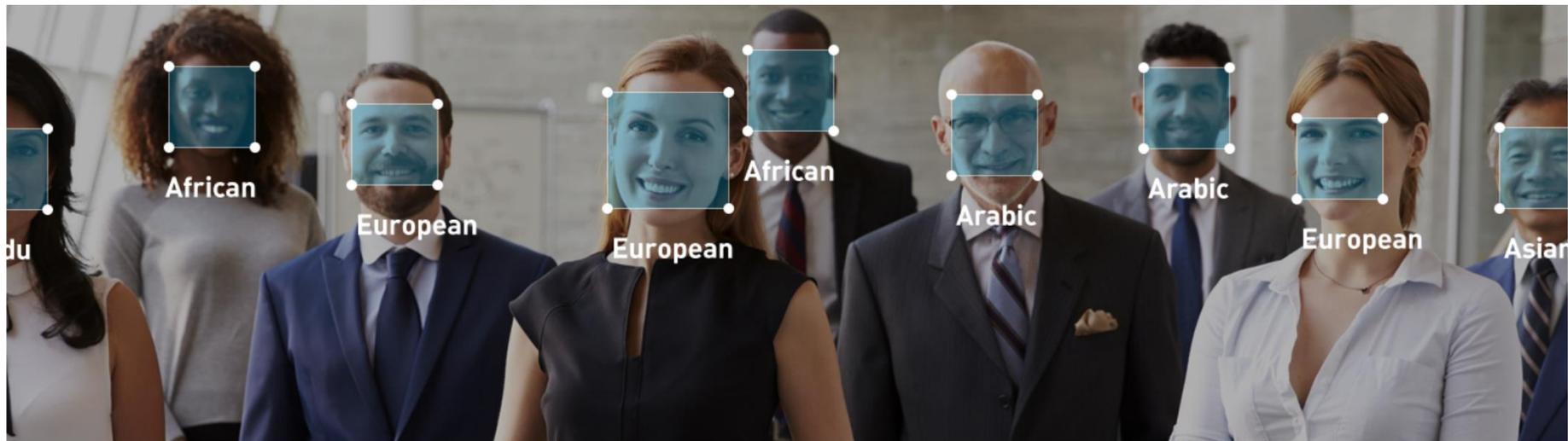
<https://ntechlab.com/>



## Coming soon

PATH TRACKING | ETHNICITY RECOGNITION

Recognizes a person's ethnicity





# AI for humans



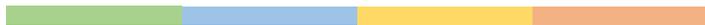
# How AI can help humans...



- Automation
- Error prevention
- Data reduction / Muffling the noise
- Prediction based on historical data
- Ploughing through massive amounts of data
- Creating more human interfaces



# How AI can help humans...

A screenshot of the Microsoft AI for Good website. The top navigation bar includes the Microsoft logo and search and shopping icons. The main image shows a man with glasses and a beard smiling while interacting with a group of diverse children around a laptop. Below the image, the text reads: "AI for good" followed by a paragraph: "The intersection of AI with people and society presents us with an incredible opportunity to leave a lasting, positive impact on the world. Our work with organizations on the front lines of education, environmental advocacy, accessibility, and healthcare, is creating inclusive solutions designed to help every person and organization on the planet achieve more." At the bottom, there are four menu items: "AI for Accessibility", "AI for Earth", "Education", and "Healthcare".

Microsoft

## AI for good

The intersection of AI with people and society presents us with an incredible opportunity to leave a lasting, positive impact on the world. Our work with organizations on the front lines of education, environmental advocacy, accessibility, and healthcare, is creating inclusive solutions designed to help every person and organization on the planet achieve more.

AI for Accessibility   AI for Earth   Education   Healthcare



How AI can help humans...



Jennifer Marsman

@jennifermarsman

Following



Team selfie at the end of a successful #AIforEarth Summit!



2:50 AM - 19 May 2018

4 Retweets 22 Likes





# How AI can help humans...



- Automation
- Error prevention
- Data reduction / Muffling the noise
- Prediction based on historical data
- Ploughing through massive amounts of data
- **Creating more human interfaces**

# Humans

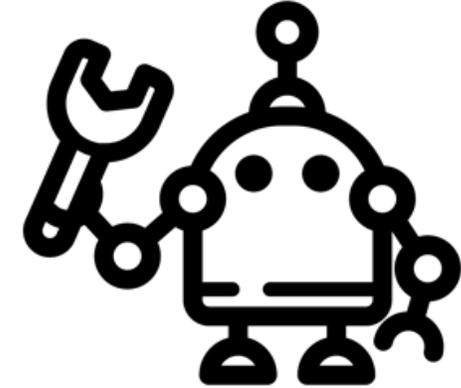


- Messy and prone to mistakes
- Forget things and filter them by their biases
- Bored when doing repetitive tasks
- When bored create more errors
- Non-optimised communication, lots of nuances and misunderstanding



Data  
Insights  
Patterns

# Bots and computers...



- Make no mistakes, other than physical fatigue
- Never forget, don't judge
- Great at tedious, boring tasks
- Repeat things with minor changes on iterations till a result is met
- Highly optimised, non-nuanced communication.



# Humans and Bots/Computers

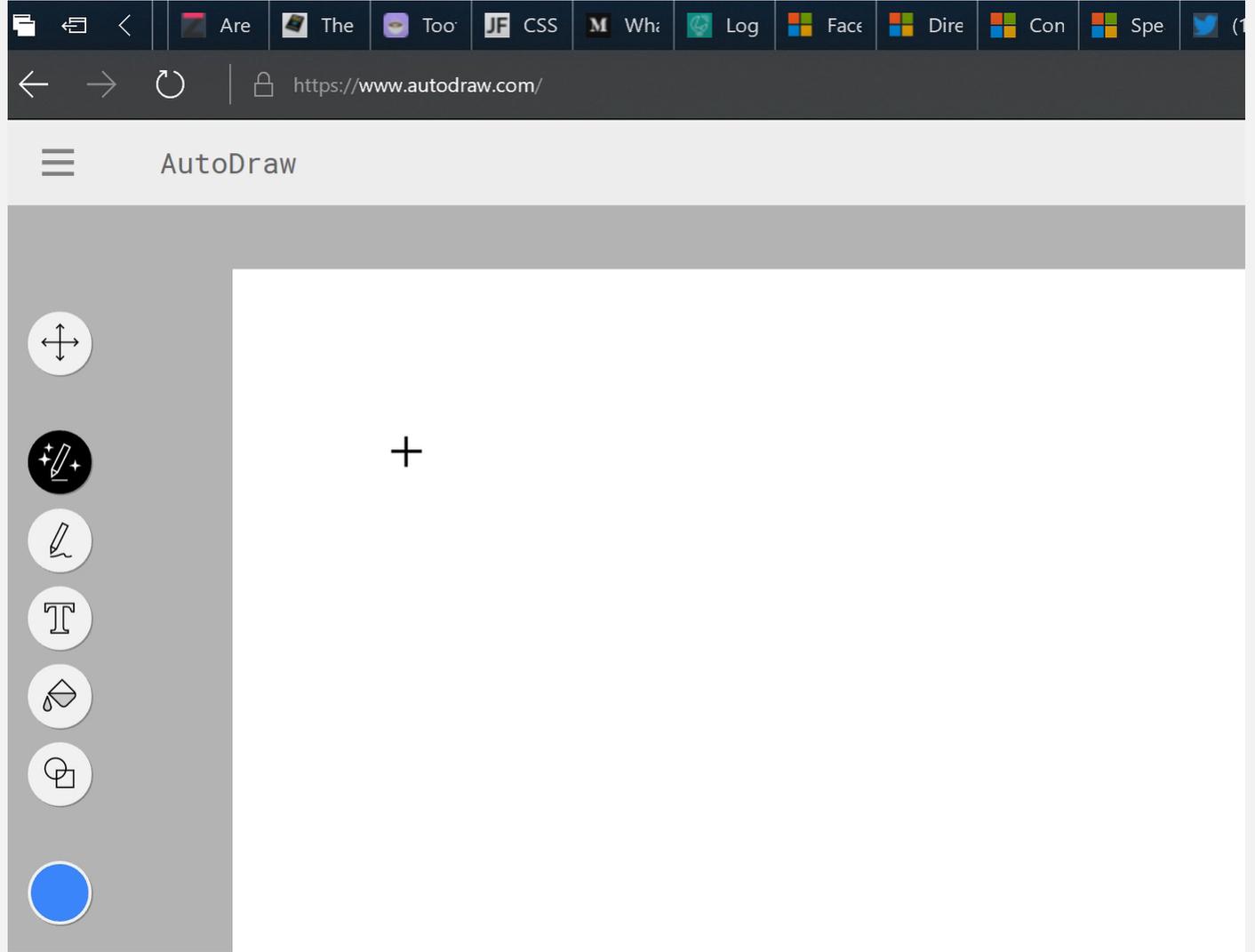


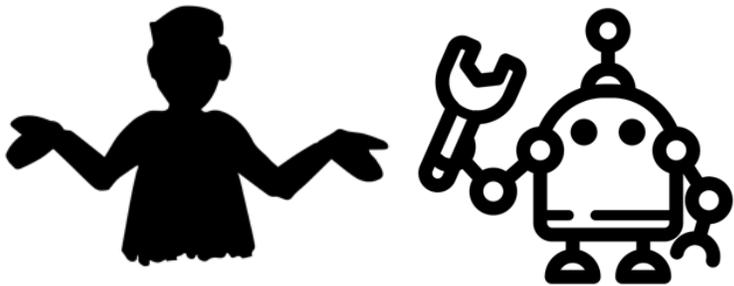
**AutoDraw**  
Fast drawing for everyone.

Start Drawing    Fast How-To\*

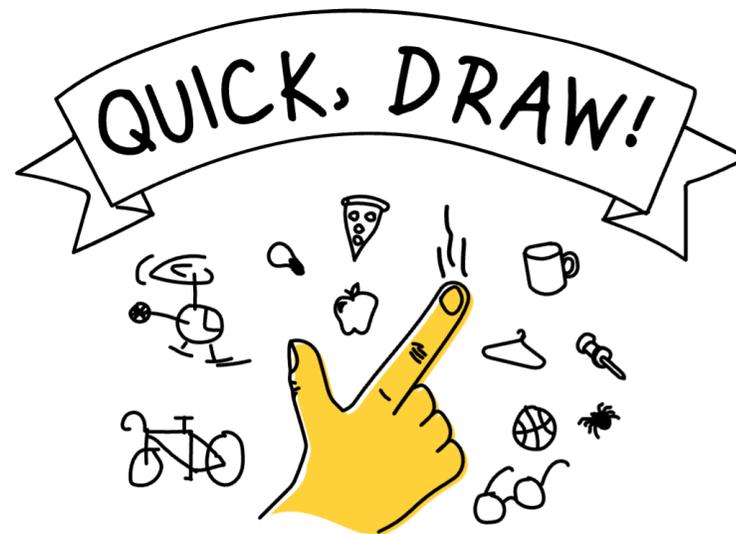


Humans  
and  
Bots/Computers





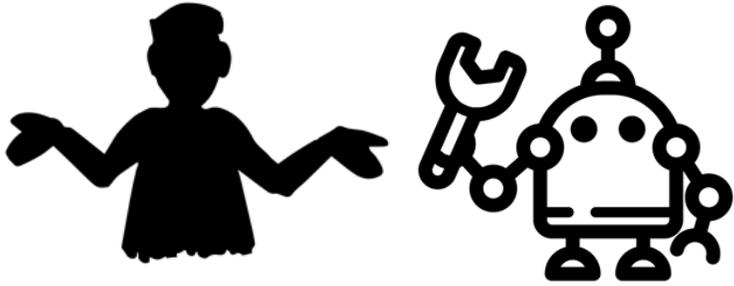
# Humans and Bots/Computers



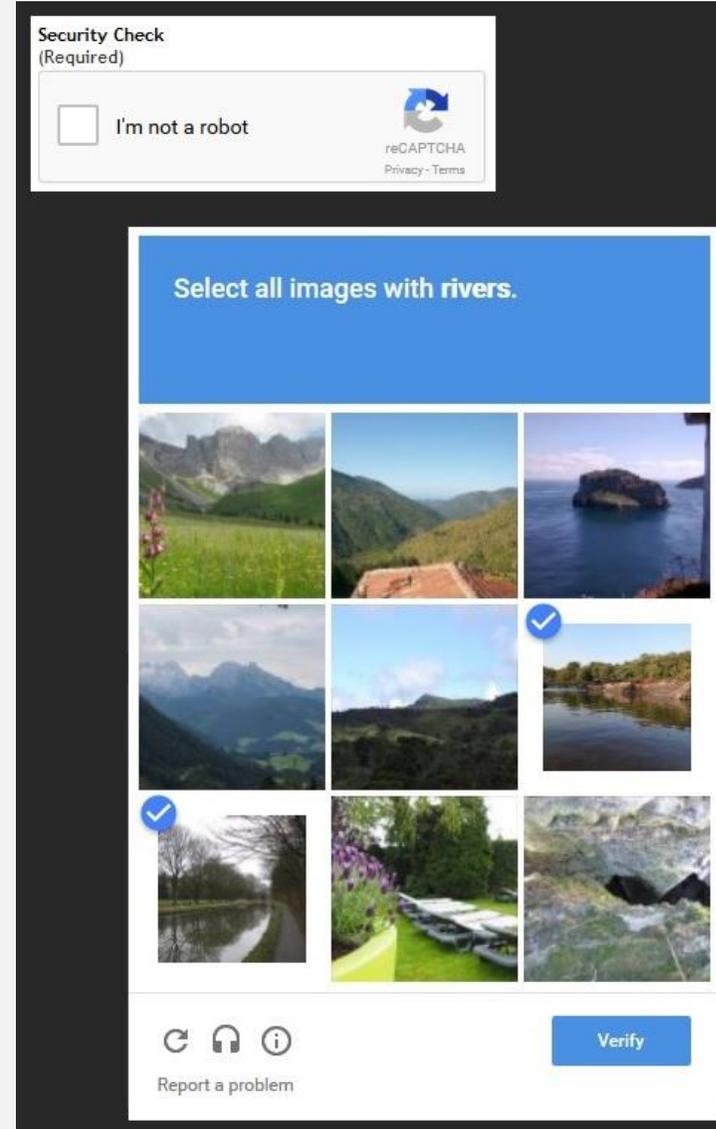
Can a neural network learn to recognize doodling?

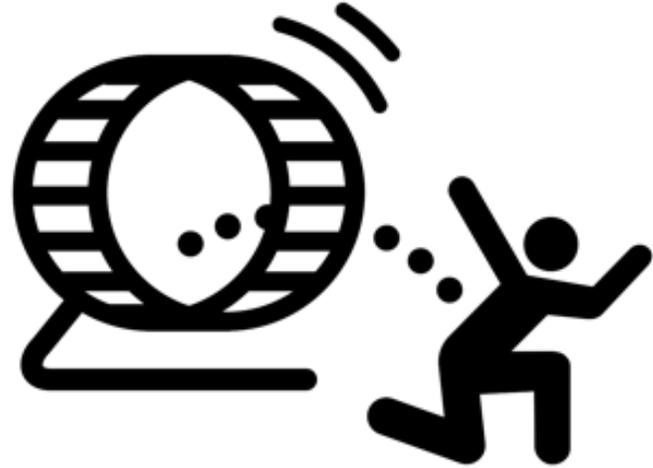
Help teach it by adding your drawings to the [world's largest doodling data set](#), shared publicly to help with machine learning research.

Let's Draw!



# Humans and Bots/Computers





# Intelligent, responsive systems



- AI services offer us lots of data to compare our users' input with
- Thus our users don't need to speak computer but be human instead
- We can prevent them from making mistakes
- **We can help getting around physical barriers**

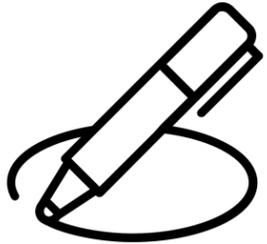
Google: [cloud.google.com/products/machine-learning](https://cloud.google.com/products/machine-learning)

Amazon: [aws.amazon.com/machine-learning](https://aws.amazon.com/machine-learning)

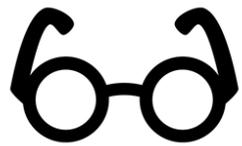
Microsoft: [azure.microsoft.com/en-us/services/cognitive-services](https://azure.microsoft.com/en-us/services/cognitive-services)



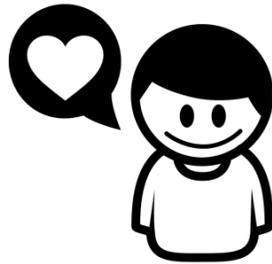
# Our toolkit for more human interfaces



Natural language processing



Computer Vision



Sentiment analysis



Speech conversion and analysis



Moderation



# Language and Writing

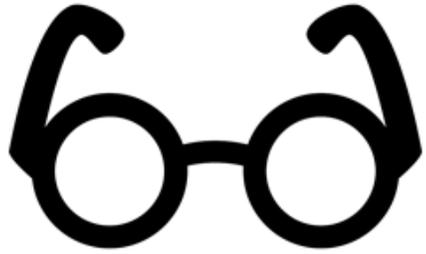


- Probably the oldest task on the web was **translation**
- This moved deeper into **Natural Language Processing** and **Language Detection**
- Using these, we can allow for **human commands** and finding out tasks by analyzing texts.

“How far am I from the capital of Denmark?”

“Where do I find a good restaurant around here?”

“Show me documents I wrote five days ago with more than 600 words”



# Computer Vision



- When text wasn't cool enough, we added **images** to our web media
- Often we forget that **not everyone can see them**, and we leave them without alternative text
- This is where machine learning steps in to help **turning an image into a dataset** we can work with.



# Vision and image analysis...



instagram: @larryandanke



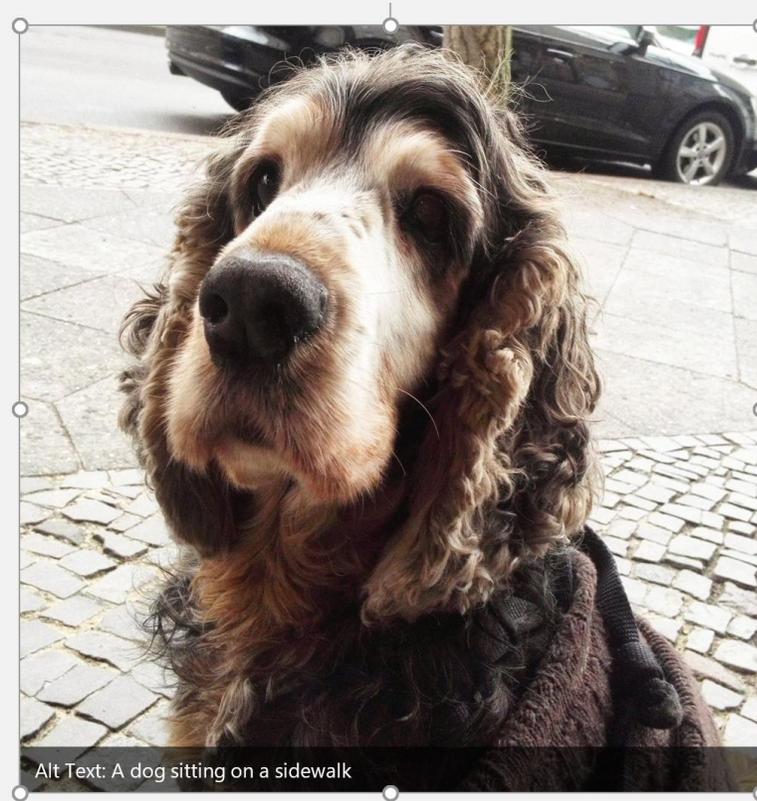
# Vision and image analysis...



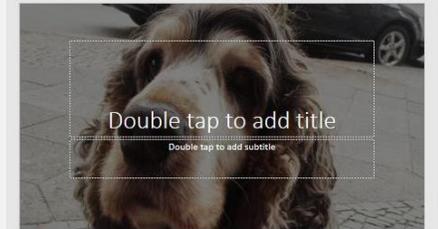
Alt Text: A dog sitting on a sidewalk



# Vision and image analysis...



## Design Ideas



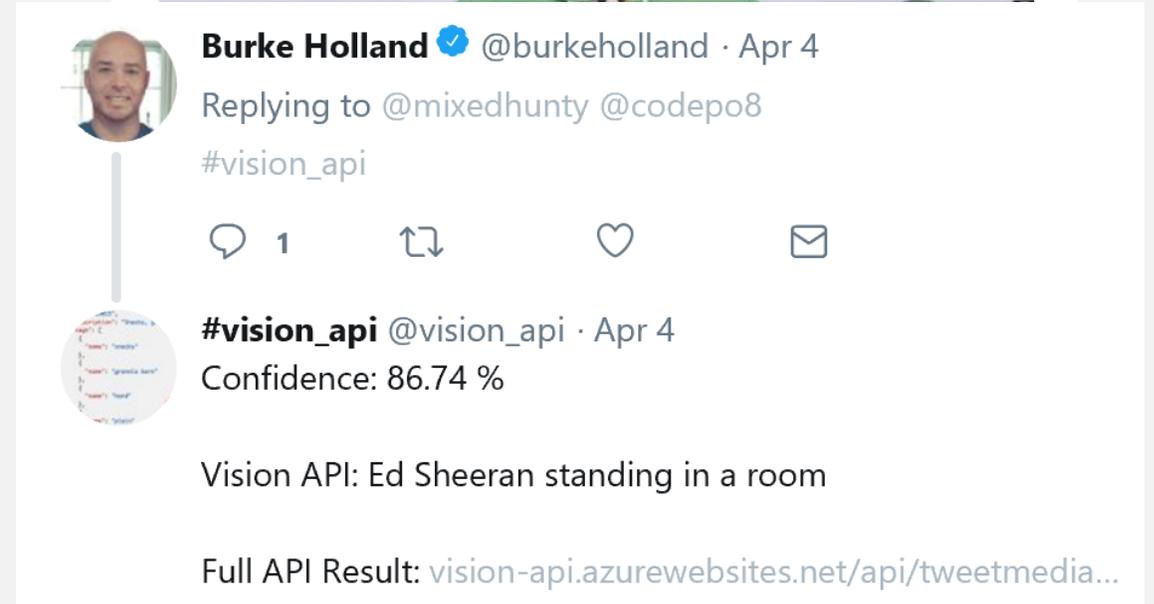


Vision and image  
analysis...





# Vision and image analysis...



# #vision\_api



# Vision and image analysis...



## Analyze an image

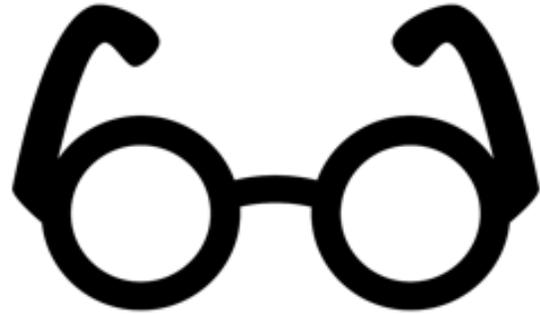
This feature returns information about visual content found in an image. Use tagging, descriptions, and domain-specific models to identify content and label it with confidence. Apply the adult/racy settings to enable automated restriction of adult content. Identify image types and color schemes in pictures.



FEATURE NAME:	VALUE
Description	{ "tags": [ "train", "platform", "station", "building", "indoor", "subway", "track", "walking", "waiting", "pulling", "board", "people", "man", "luggage", "standing", "holding", "large", "woman", "yellow", "suitcase" ], "captions": [ { "text": "people waiting at a train station", "confidence": 0.833099365 } ] }
Tags	[ { "name": "train", "confidence": 0.9975446 }, { "name": "platform", "confidence": 0.995543063 }, { "name": "station", "confidence": 0.9798007 }, { "name": "indoor",

Image URL





# Vision and image analysis...



## Read text in images

Optical character recognition (OCR) detects text in an image and extract the recognized words into a machine-readable character stream. Analyze images to detect embedded text, generate character streams, and enable searching. Take photos of text instead of copying to save time and effort.



Preview

JSON

IF WE DID  
ALL  
THE THINGS  
WE ARE  
CAPABLE\*  
OF DOING,  
WE WOULD  
LITERALLY  
ASTOUND  
QURSELV\*S.

Image URL

Submit

Browse





# Vision and image analysis...



## Preview: Read handwritten text from images

This technology (handwritten OCR) allows you to detect and extract handwritten text from notes, letters, essays, whiteboards, forms, etc. It works with different surfaces and backgrounds, such as white paper, yellow sticky notes, and whiteboards.

Handwritten text recognition saves time and effort and can make you more productive by allowing you to take images of text, rather than having to transcribe it. It makes it possible to digitize notes, which then allows you to implement quick and easy search. It also reduces paper clutter.

The screenshot shows a web interface for the vision API. At the top, there are two tabs: 'Preview' (selected) and 'JSON'. The main area displays a handwritten note on a pink background with the text: "Our greatest glory is not in never failing, but in rising every time we fall". Below this, there is a 'Preview' section showing the same text in a clean, digital font. At the bottom of the interface, there is an 'Image URL' input field, a blue 'Submit' button, and a 'Browse' button with a folder icon. Below the main preview, there are six small thumbnails of various handwritten notes on different backgrounds (pink, white, yellow, purple, white, and yellow).



# Vision and image analysis...



## Recognize celebrities and landmarks

The Celebrity and Landmark Models are examples of Domain Specific Models. Our celebrity recognition model recognizes 200K celebrities from business, politics, sports and entertainment. Our landmark recognition model recognizes 9000 natural and man-made landmarks from around the world. Domain Specific Models is a continuously evolving feature within Computer Vision API.



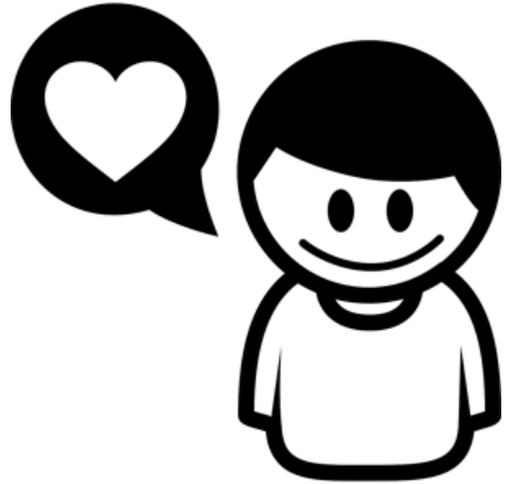
```
all"
"ions": [
'text': "a group of people in front of Colosseu",
'confidence': 0.84631330287730377

stId": "c7112a85-6b04-44fd-96d6-057e70fb8763",
ata": {
ch": 600,
ght": 399,
nat": "Jpeg"

": [],
': {
inantColorForeground": "Grey",
inantColorBackground": "White",
inantColors": [
rey",
hite"
```

Image URL

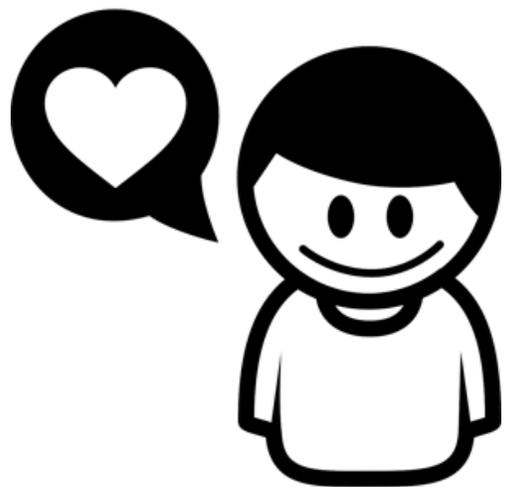




# Sentiment analysis

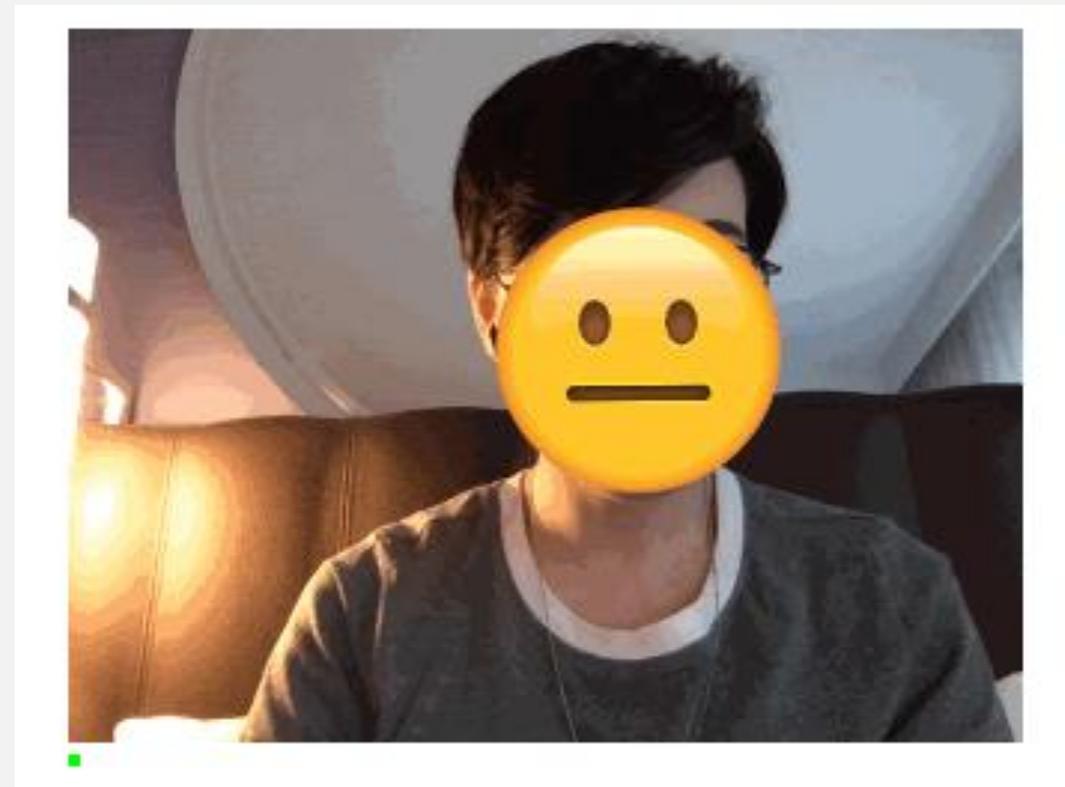
---

- Finding out the sentiment of a text, image or video can help with a lot of things
- You can navigate videos by only showing the happy parts
- You can detect which comment should be answered first by a help desk
- You can predict when drivers of cars get tired



# Sentiment analysis

---





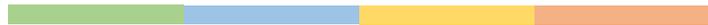
# Speech

---

- Audio interfaces are **all the rage**.
- You can allow **hands-free control** of devices
- You can have an “**always on**” system to help you out **without having to interface with it**
- It feels natural and has **a massive Sci-Fi feeling** – when it works.



# Text to speech



## Text to Speech

Convert text to spoken audio. When applications need to “talk” back to their users, this API can be used to convert text that is generated by the app into audio that can be played back to the user.

The Text-To-Speech API enables you to build smart apps that can speak. You can test it now, simply choose your target language, add your sentences then click on the play button to see how speech synthesis works. When you use this demo you consent to providing your voice input data to Microsoft for service improvement purposes.

English - US    ZiraRUS

Convert text to spoken audio. When applications need to “talk” back to their users, this API can be used to convert text that is generated by the app into audio that can be played back to the user.

00:12    View SSML    302 characters left    Play



# Speech recognition

---

## Speech Recognition

Convert spoken audio to text. The API can be directed to turn on and recognize audio coming from the microphone in real-time, recognize audio coming from a different real-time audio source, or to recognize audio from within a file. In all cases, real-time streaming is available, so as the audio is being sent to the server, partial recognition results are also being returned.

The Speech to Text API enables you to build smart apps that are voice triggered. To see how it works select your target language then click on the microphone and start speaking. Or simply click on one of the sample speech phrases to see how speech recognition works. When you use this demo you consent to providing your voice input data to Microsoft for service improvement purpose

 Start recording

English (US) 

Well hello there computer how are you today?

 Play sample 1

 Play sample 2



# Turning sentences into commands



LUIS: Intent Page

Secure | <https://www.luis.ai/applications/c4396135-ee3f-40a9-8b83-4704cddabf7a/vers...>

Language Understanding My apps Docs Pricing Support About Burke Holland

home-automation (v 0.1) DASHBOARD BUILD PUBLISH SETTINGS Train Test

App Assets

- Intents
- Entities

Improve app performance

- Review endpoint utterances
- Phrase lists

PREVIEW Prebuilt Domains

## Control Lights

Delete Intent

Type about 5 examples of what a user might say to trigger this task and hit Enter.

Search for an utterance Reassi... Delete utterance(s)

Filters:  Errors Entity  Entities view  Fuzzy search

Utterance Labeled intent ?

- turn **Power** the lights in the **Location** Control Lights -1 ...
- turn **Power** the **Location** lights Control Lights -1 ...
- turn **Power** the **Location** lights Control Lights -1 ...
- Location** lights **Power** Control Lights -1 ...

luis.ai  
aka.ms/luis-api



# Conversation as an interface



## The Rise Of Intelligent Conversational UI

# [UI](#) <sup>54</sup> # [Visual Design](#) <sup>59</sup> # [Interfaces](#) <sup>32</sup> # [User Interaction](#) <sup>49</sup>



### ABOUT THE AUTHOR

Burke Holland is a front-end developer living in Nashville, TN; the greatest city in the world. He enjoys JavaScript a lot because it's the only way he ... [More about Burke...](#)



**For a long time, we've thought of interfaces** strictly in a visual sense: buttons, dropdown lists, sliders, carousels (please no more carousels). But now we are staring into a future composed not just of visual interfaces, but of conversational ones as well. Microsoft alone reports that three thousand new bots are built every week on their [bot framework](#). Every.

Week.



# Speaker recognition

Text JSON

President George H W Bush  
is the one identified speaking in the selected audio.

▶ Stop ▶ Stop ▶ Audio 3

□ Stop ▶ Audio 5 ▶ Audio 6



# Speaker recognition

apple juice tastes funny after toothpaste

"apple juice tastes funny after toothpaste"

Your enrollment was not successful. We still need **two more** samples of your voice reading the above phrase.

 Start recording



```
{  
  "EnrollmentStatus": "Enrolling",  
  "EnrollmentsCount": 1,  
  "RemainingEnrollments": 2,  
  "Phrase": "i am going to make him an offer he cannot refuse"  
}
```



# Moderation

---

- Some things are not meant to be consumed by people
- Computers don't need counselling once they saw them – people should
- Known **illegal and terrible content** can be automatically removed



With great power  
comes great  
responsibility...



## Our responsibilities..

---

- AI can be an **amazing help** for humans
- It does need **transparency** – if you use people as data sources, they need to know what and where it goes
- When people get information **filtered by an algorithm**, it should be an **opt-in**
- People need to have a **chance to dispute** when an algorithm tagged or disallowed them access.



# Want to go deep?

---

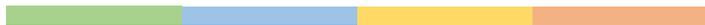
- The Math behind ML
- The ethics of AI
- Working with Data using Python
- Machine Learning Models
- Deep Learning Models
- Reinforcement Learning Models
- Microsoft Professional Program Certificate in Artificial Intelligence

[aka.ms/learn-ai](https://aka.ms/learn-ai)

10 courses, (8-16 hours each), 10 skills



# Want to go deep?



@codepo8

Demystifying Artificial Intelligence: Understanding Machine Learning | ORIGINAL

Christian Hellmann, Senior Developer at Microsoft [+ Follow](#)



12 Videos (58m) [View My Notes](#)

▶ 1. Introduction	1:44
🔒 2. What is Machine Learning	5:25
🔒 3. How We Teach Machines	5:48
🔒 4. Machine Learning to Help Humans	5:28
🔒 5. Tools for Machine Learning	3:44
🔒 6. Visual Uses	7:54
🔒 7. Speaking Human	6:07
🔒 8. Audio & Video	6:32

97 students are watching this class

[About](#) [Community 3](#) [Class Project](#) [All Projects](#) [Save](#) [Add to Calendar](#) [f](#) [t](#) [e](#) [Report class](#)

# skl.sh/christian

Free with trial sign-up

# Thanks!

Chris Heilmann

[Christianheilmann.com](http://Christianheilmann.com)

[Developer-evangelism.com](http://Developer-evangelism.com)

@codepo8

