

Building human interfaces powered by AI



Chris Heilmann (@codepo8) November 2018



All resources:
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**



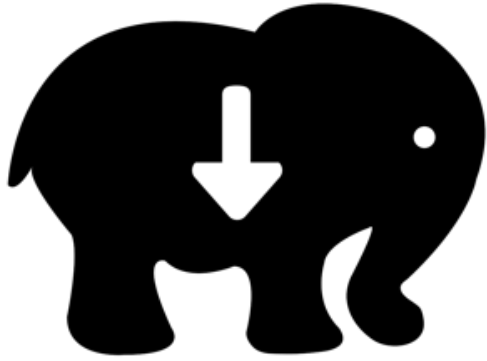
The machines
are watching...



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.



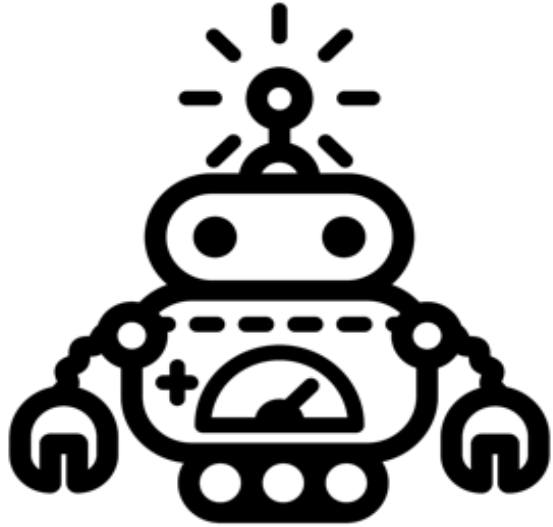
Are machines
friend or foe?



Artificial Intelligence Myths

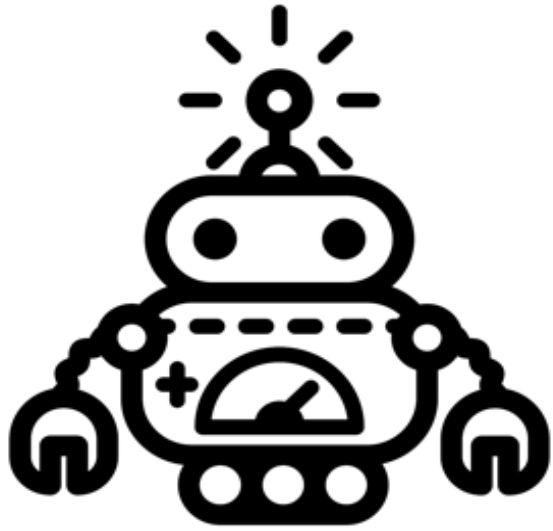


- AI can't replace a thinking, creative human
- AI can not **magically fill gaps with perfect information** – it can only compare and assume
- AI doesn't learn in a creative fashion. It makes **no assumptions**
- AI has **no morals and ethics**, but – used wrongly – it can **amplify our biases**



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

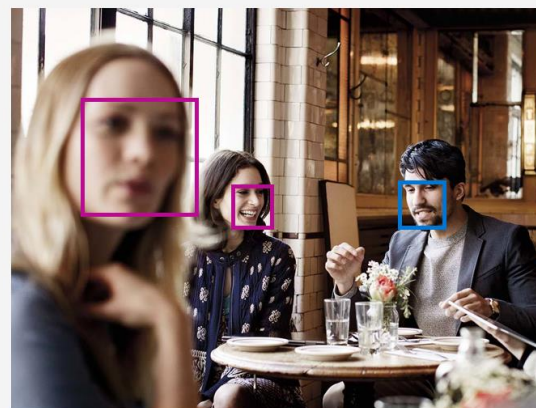


About face...

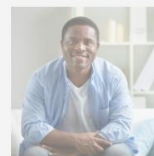


Face detection

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",
            "confidence": 0.0
          }
        ]
      }
    }
  }
]
```





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.

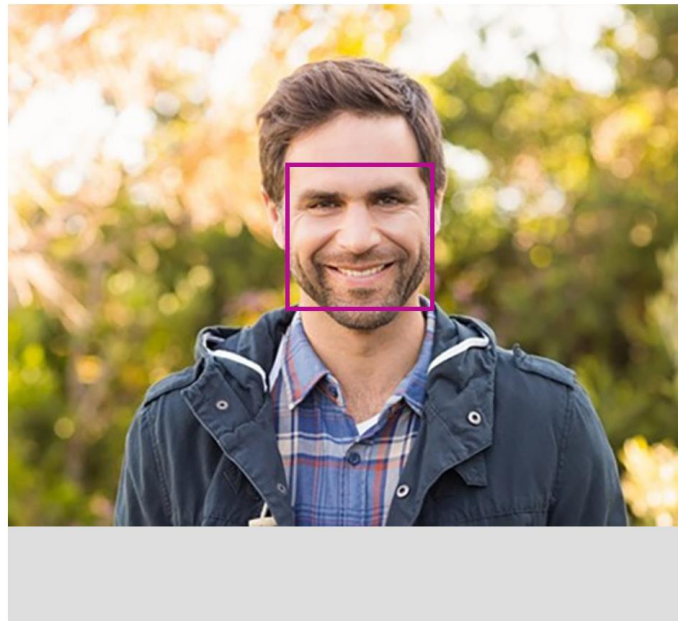


Image URL

Submit

 Browse

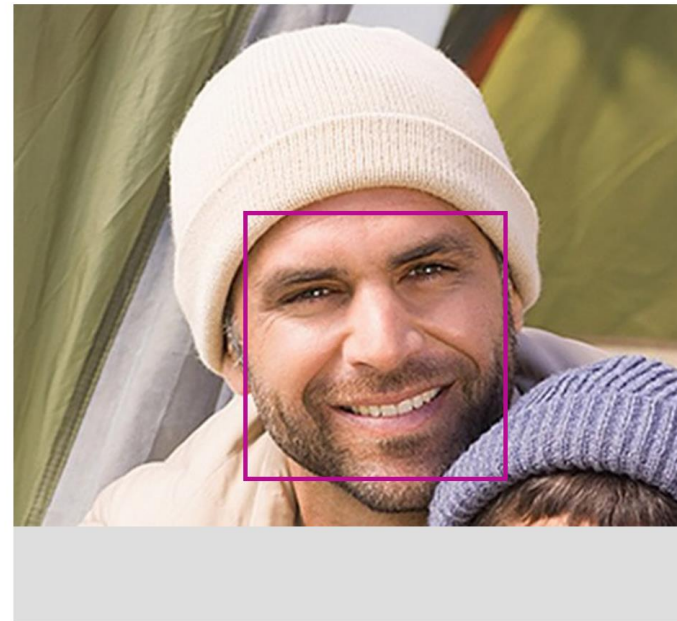


Image URL

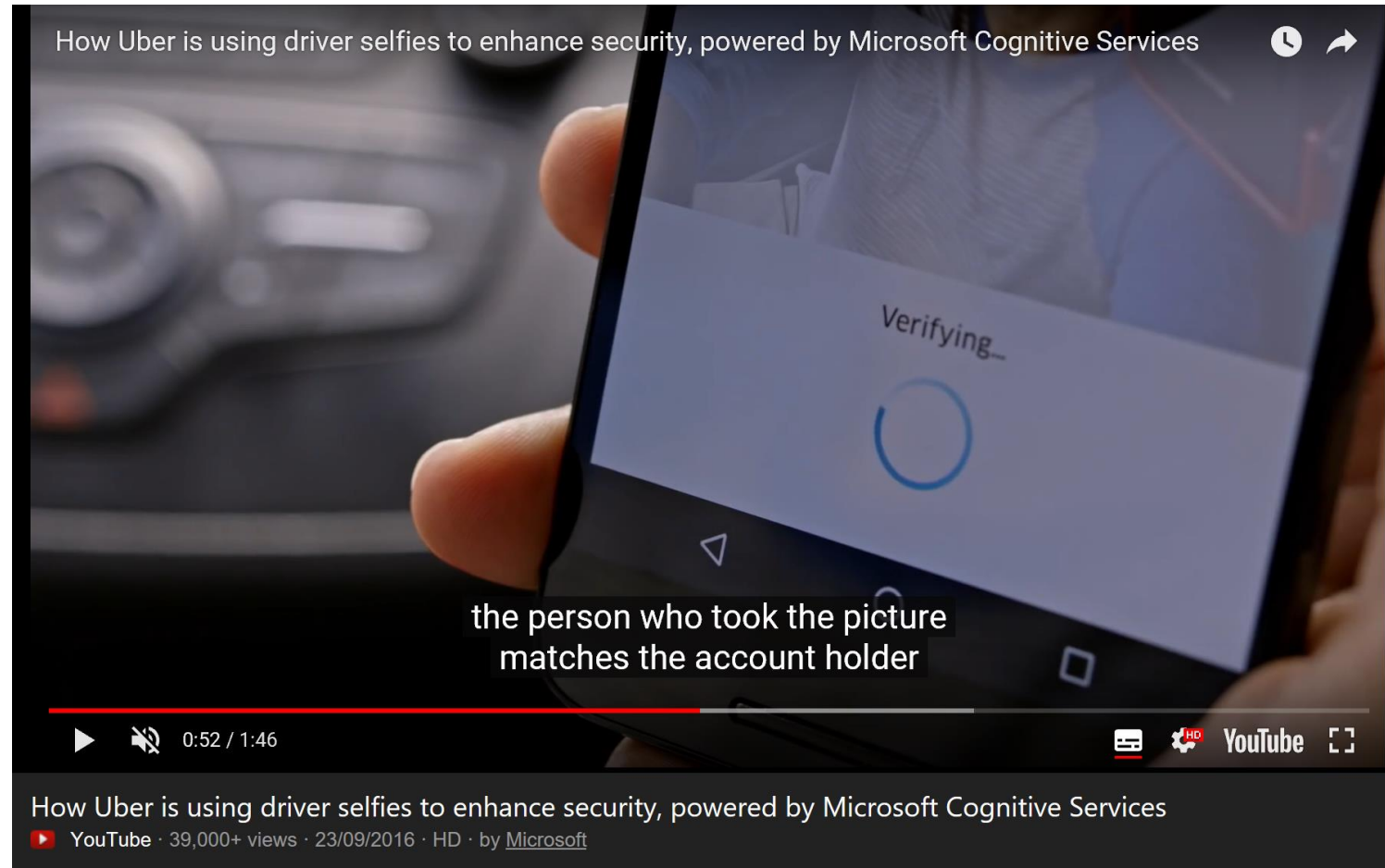
Submit

 Browse



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

Is this your driver?



Taking it too far?

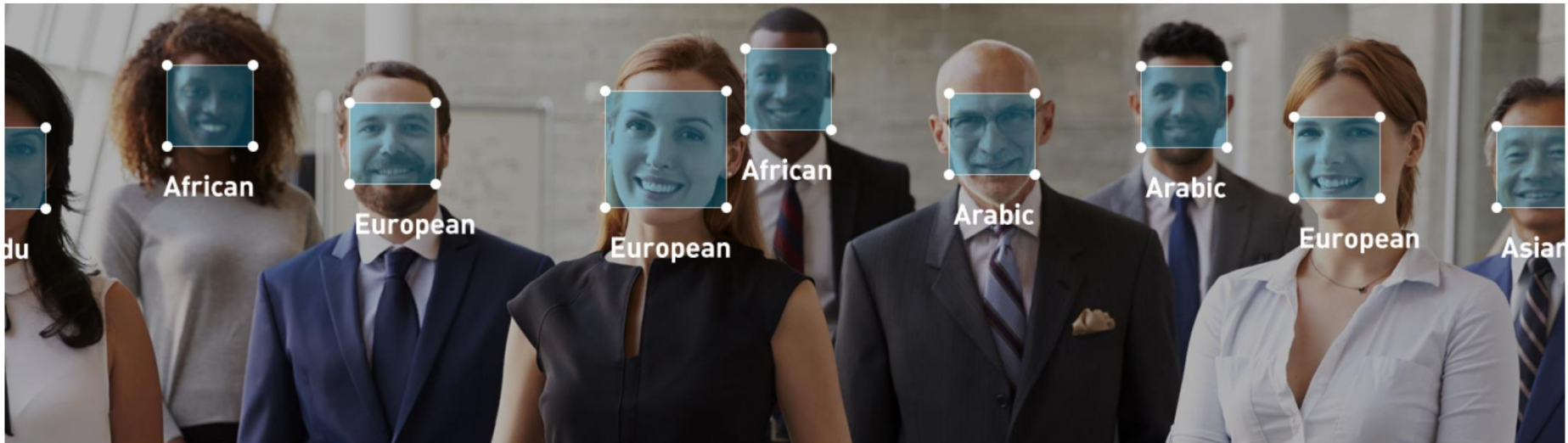
<https://ntechlab.com/>



Coming soon

PATH TRACKING | ETHNICITY RECOGNITION

Recognizes a person's ethnicity

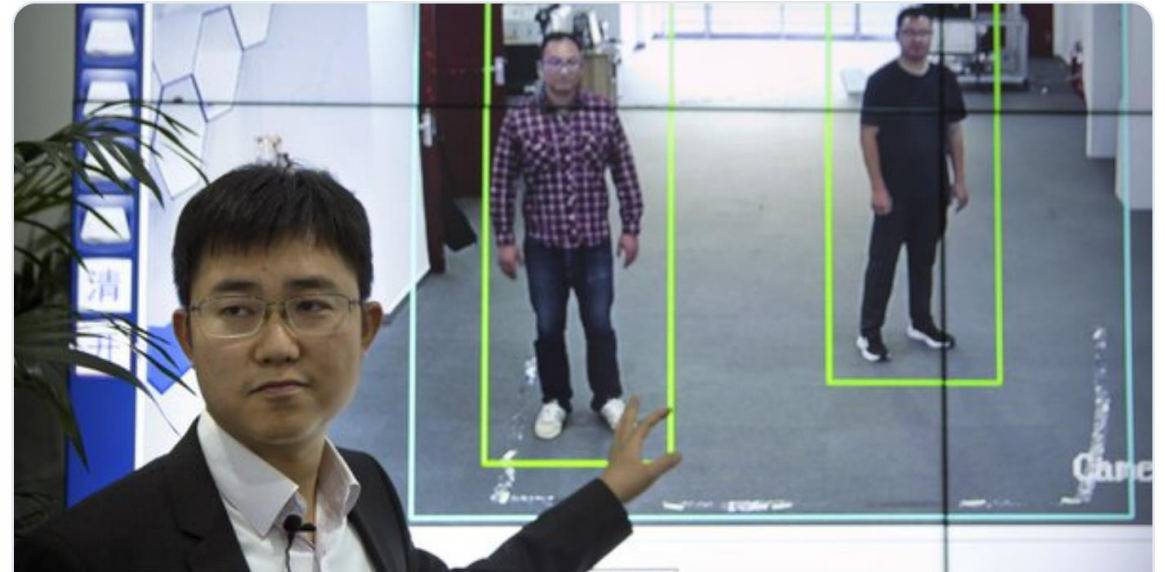




Detecting even
more...



Chinese authorities have begun deploying "gait recognition" AI software in Beijing and Shanghai that identifies people via their body shapes and how they walk (Dake Kang/Associated Press)



Chinese 'gait recognition' tech IDs people by how they walk

BEIJING (AP) — Chinese authorities have begun deploying a new surveillance tool: "gait recognition" software that uses people's body shapes and how they walk to

[apnews.com](https://apnews.com/bf75dd1c26c947b7826d270a16e2658a)



Those trustworthy
avatars...






Those trustworthy
avatars...

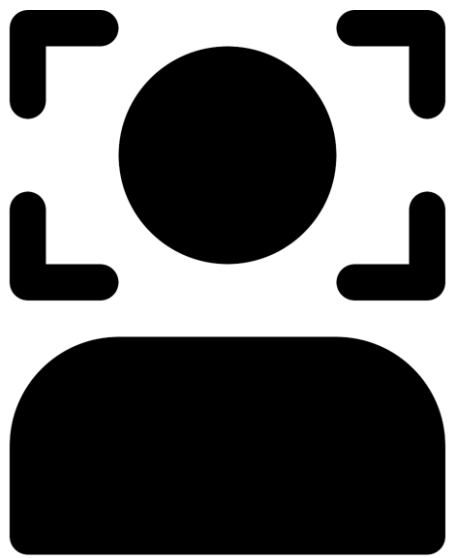


INSTRUCTION: press +/- to adjust feature, toggle feature name to lock the feature



random face

Male	Age	Skin_Tone
- +	- +	- +
Bangs	Hairline	Bald
- +	- +	- +
Big_Nose	Pointy_Nose	Makeup
- +	- +	- +
Smiling	Mouth_Open	Wavy_Hair
- +	- +	- +
Beard	Goatee	Sideburns
- +	- +	- +
Blond_Hair	Black_Hair	Gray_Hair
- +	- +	- +
Eyeglasses	Earrings	Necktie
- +	- +	- +



Automated face mapping...



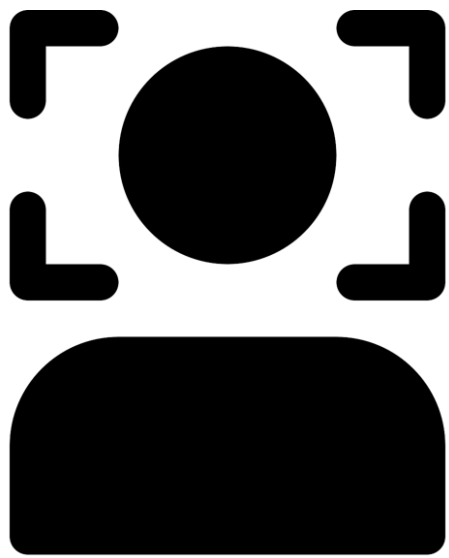
Social Mapper has a variety of uses in the security industry, for example the automated gathering of large amounts of social media profiles for use on targeted phishing campaigns. Facial recognition aids this process by removing false positives in the search results, so that reviewing this data is quicker for a human operator.

Social Mapper supports the following social media platforms:

- LinkedIn
- Facebook
- Twitter
- GooglePlus
- Instagram
- VKontakte
- Weibo
- Douban

Social Mapper takes a variety of input types such as:

- An organisations name, searching via LinkedIn
- A folder full of named images
- A CSV file with names and url's to images online



Once you are
known...



- Create fake social media profiles to 'friend' the targets and send them links or malware. Recent statistics show social media users are more than twice as likely to click on links and open documents compared to those delivered via email.
- Trick users into disclosing their emails and phone numbers with vouchers and offers to make the pivot into phishing, vishing or smishing.
- Create custom phishing campaigns for each social media site, knowing that the target has an account. Make these more realistic by including their profile picture in the email. Capture the passwords for password reuse.
- View target photos looking for employee access card badges and familiarise yourself with building interiors.



www.ericvander.com

AI for humans



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 visual is a photograph of a diverse group of children and a man with glasses looking at a laptop. Below the photo, the heading "AI for good" is centered. A paragraph of text follows, describing the impact of AI in education, healthcare, and environmental advocacy. At the bottom, a horizontal menu lists four categories: "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...



- 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...



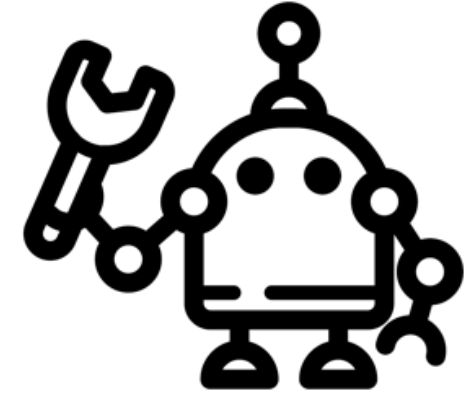
- Automation
- Error prevention
- Data reduction / Muffling the noise
- Prediction based on historical data
- Ploughing through massive amounts of data
- **Creating more human interfaces**

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

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.

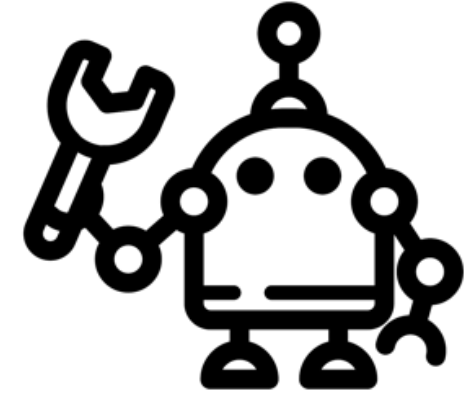
Humans



- 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



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

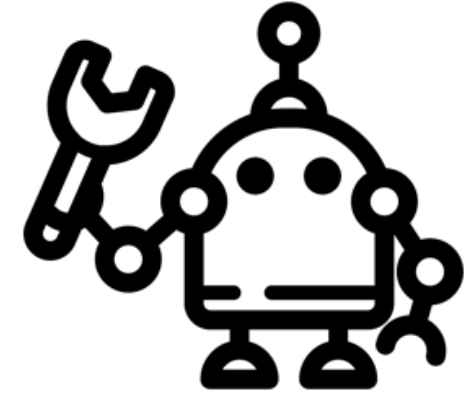


- Messy and prone to mistakes
- Forget things and filter them by their biases
- Bored when doing repetitive tasks
- When bored create more errors
- Non-optimised communication, lots of nuances and misunderstanding



Data
Insights
Patterns

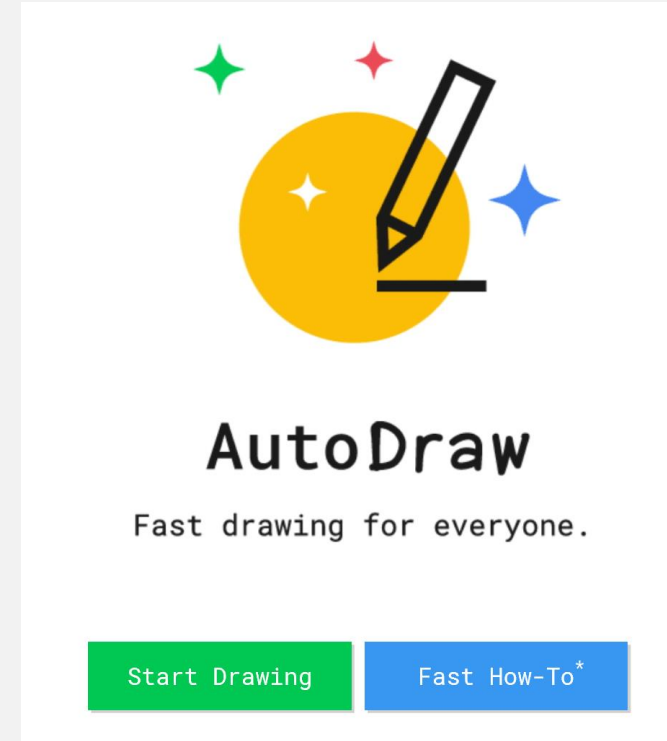
Bots and computers...

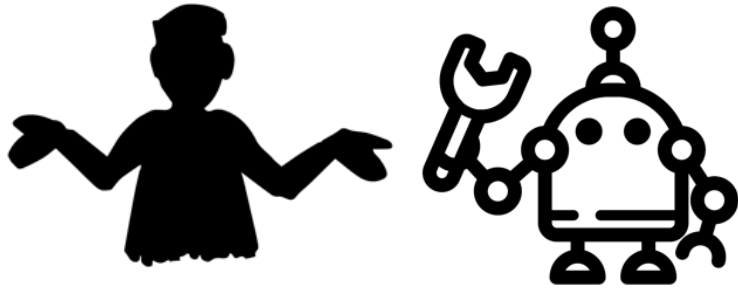


- Make no mistakes, other than physical fatigue
- Never forget, don't judge
- Great at tedious, boring tasks
- Repeat things with minor changes on iterations till a result is met
- Highly optimised, non-nuanced communication.

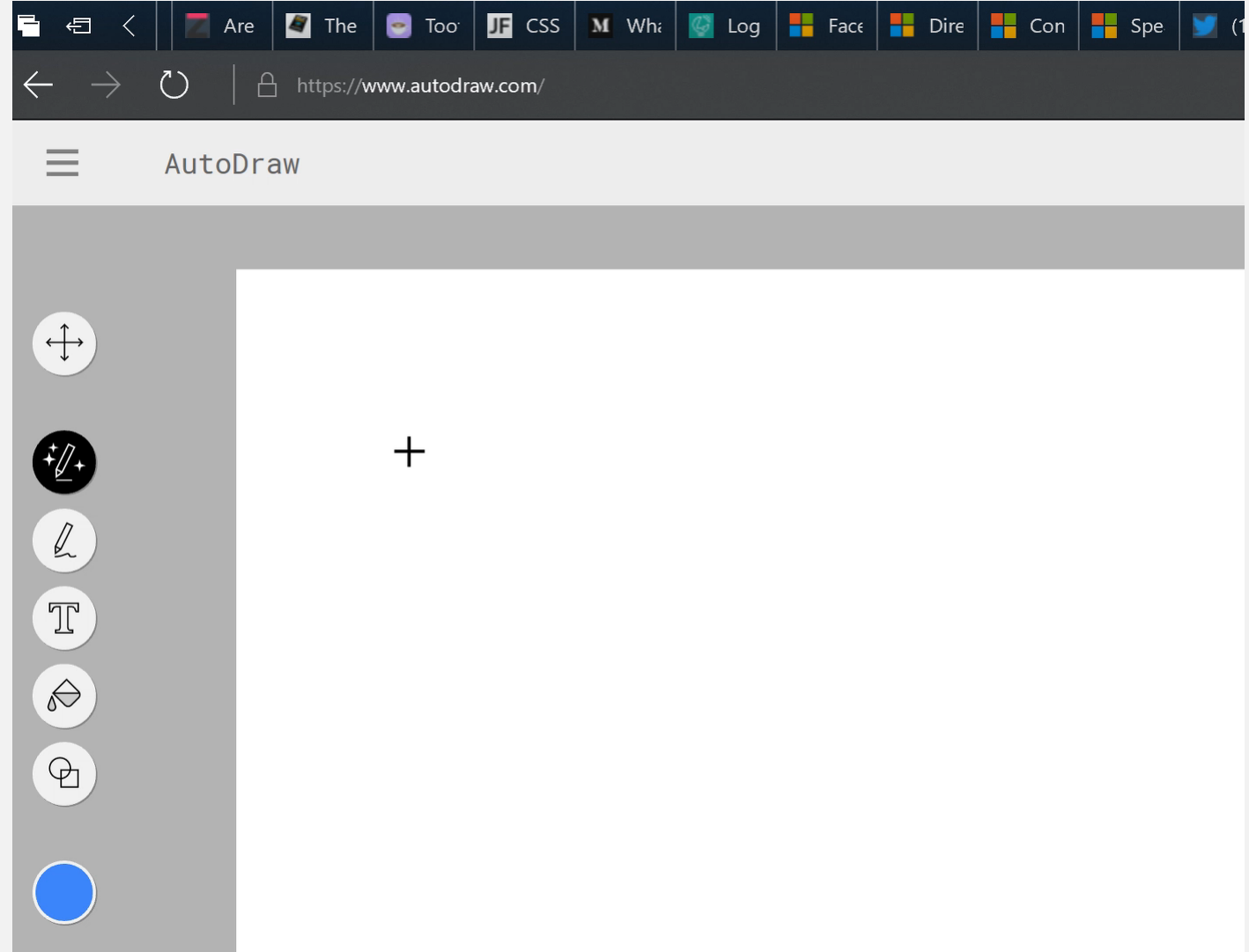


Humans
and
Bots/Computers





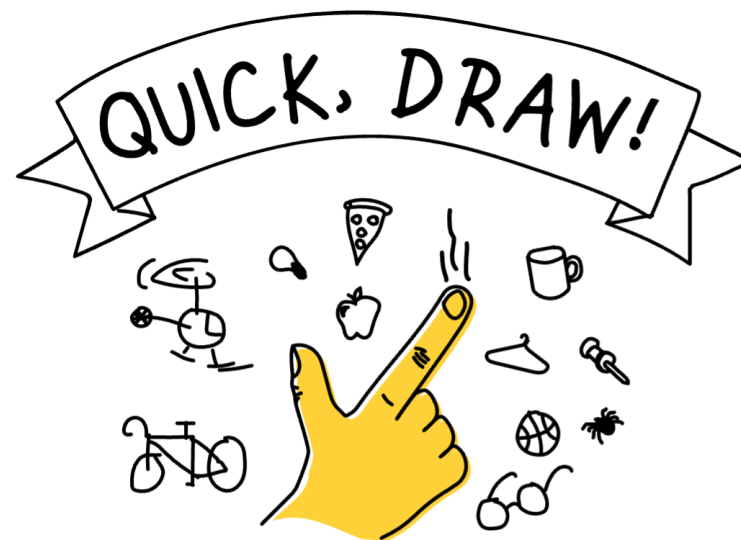
Humans
and
Bots/Computers



autodraw.com



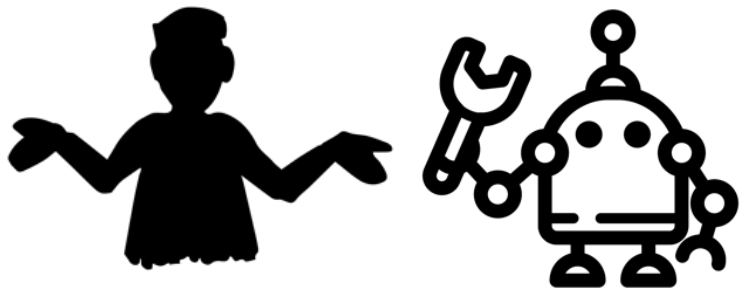
Humans and Bots/Computers



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



Security Check
(Required)

☐ I'm not a robot

reCAPTCHA
Privacy - Terms

Select all images with rivers.

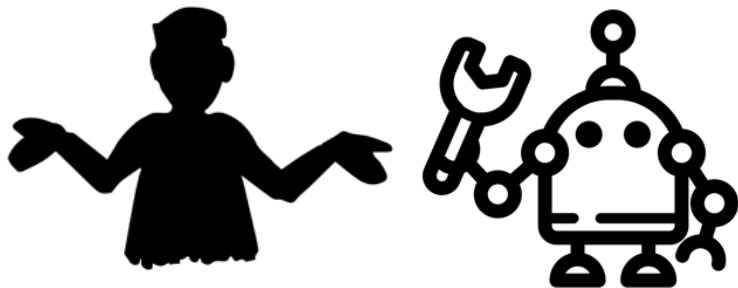
Report a problem

Verify




“Learning” from
lots of images





Humans and Bots/Computers






NEWS CENTER

ACCELERATED COMPUTINGARTIFICIAL INTELLIGENCEAUTONOMOUS VEHICLESDESIGN & VISUALIZATIONGAME

Comments 13.3K



New AI Imaging Technique Reconstructs Photos with Realistic Results

April 22, 2018


Researchers from NVIDIA, led by Guilin Liu, [introduced](#) a state-of-the-art [deep learning](#) method that can edit images or reconstruct a corrupted image, one that has holes or is missing pixels.

The method can also be used to edit images by removing content and filling in the resulting holes.

The method, which performs a process called “image inpainting”, could be implemented in photo editing software to remove unwanted content, while filling it with a realistic computer-generated alternative.

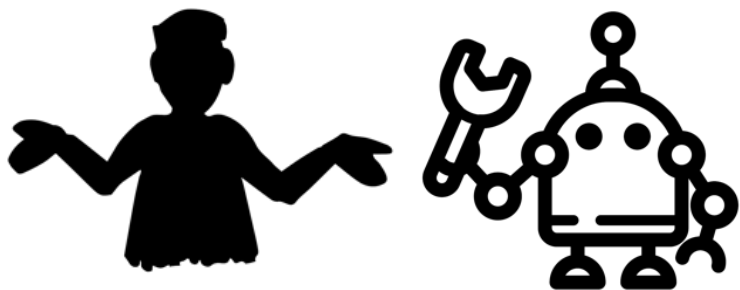
“Our model can robustly handle holes of any shape, size location, or distance from the image borders. Previous deep learning approaches have focused on rectangular regions located around the center of the image, and often rely on expensive post-processing,” the NVIDIA researchers stated in their [research paper](#). “Further, our model gracefully handles holes of increasing size.”

To prepare to train their neural network, the team first generated 55,116 masks of random streaks and holes of arbitrary shapes and sizes for training. They also generated nearly 25,000 for testing. These were further categorized into six categories based on sizes relative to the input image, in order to improve reconstruction accuracy.



An example of the masks generated for training.

aka.ms/nvidia-fix-image



Humans and Bots/Computers



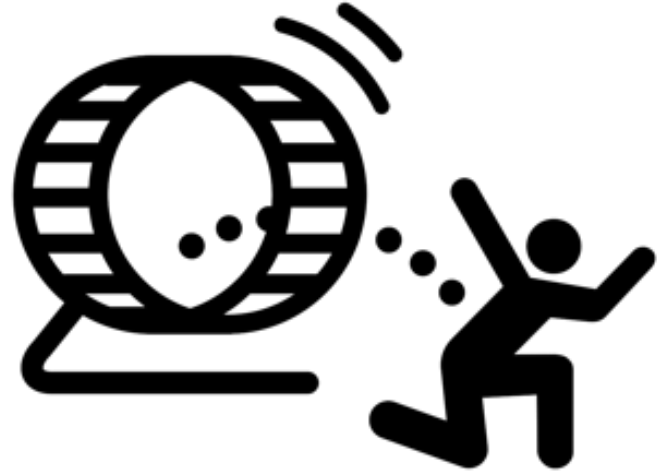
aka.ms/nvidia-fix-image



Humans
and
Bots/Computers



aka.ms/nvidia-fix-image



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**

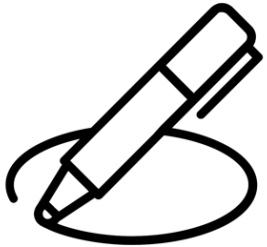
Google: cloud.google.com/products/machine-learning

Amazon: aws.amazon.com/machine-learning

Microsoft: azure.microsoft.com/en-us/services/cognitive-services



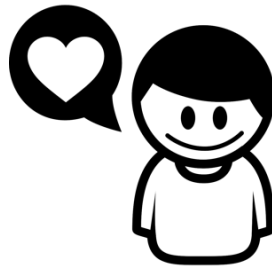
Our toolkit for more human interfaces



Natural
language
processing



Computer
Vision



Sentiment
analysis



Speech
conversion
and analysis



Moderation



Language and Writing

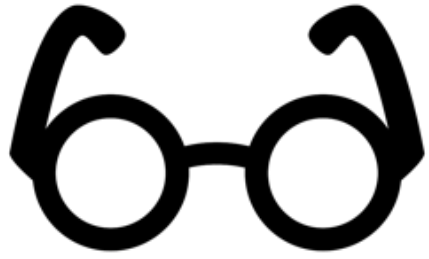


- Probably the oldest task on the web was **translation**
- This moved deeper into **Natural Language Processing** and **Language Detection**
- Using these, we can allow for **human commands** and finding out tasks by analyzing texts.

"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



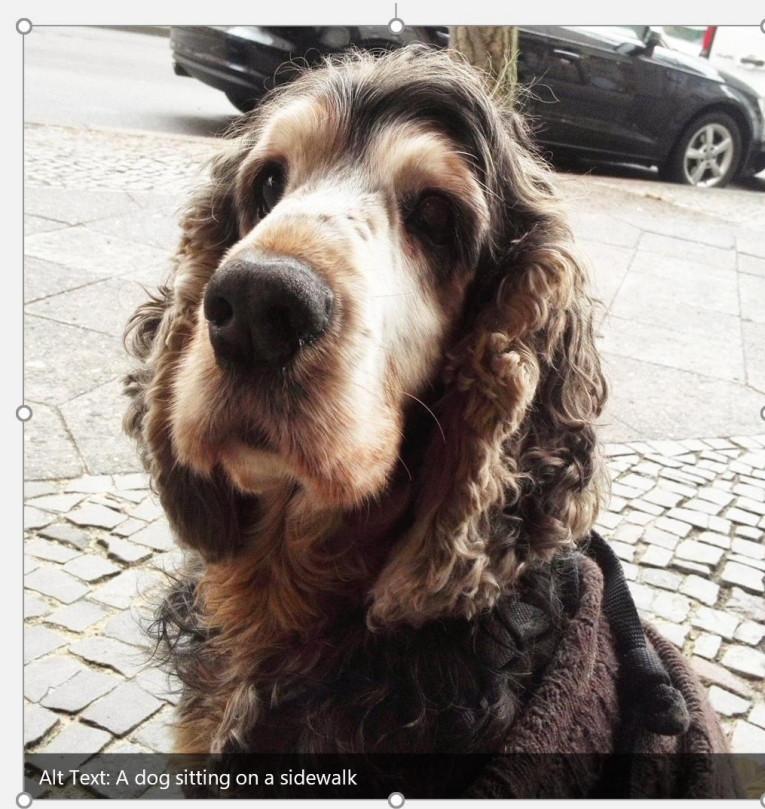
Vision and image analysis...



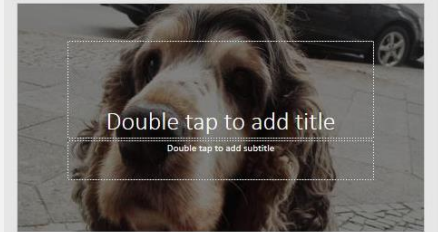
Alt Text: A dog sitting on a sidewalk



Vision and image analysis...



Design Ideas





Vision and image
analysis...



dallas
@mixedhunty

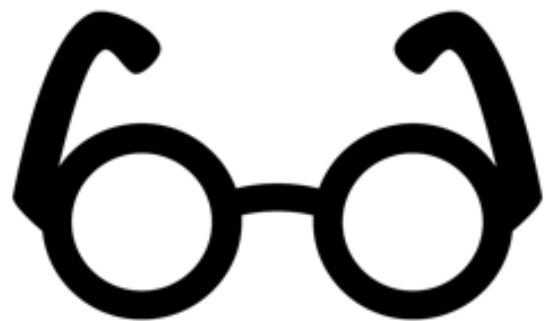
Follow



mom said you had to let me use the xbox



10:03 PM - 1 Apr 2018



Vision and image analysis...



dallas
@mixedhunty

Follow

mom said you had to let me use the xbox



Burke Holland ✓ @burkeholland · Apr 4

Replying to @mixedhunty @codepo8

#vision_api



1



#vision_api @vision_api · Apr 4

Confidence: 86.74 %

Vision API: Ed Sheeran standing in a room

Full API Result: vision-api.azurewebsites.net/api/tweetmedia...

#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

Submit

Browse





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.

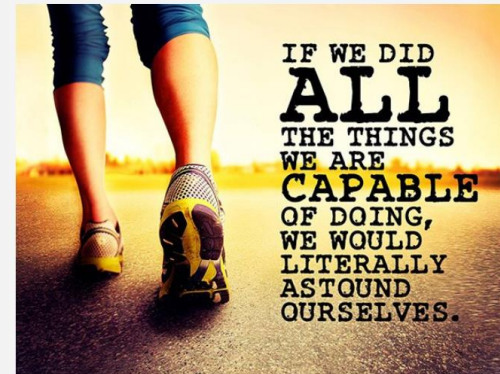


Image URL

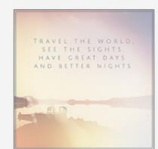
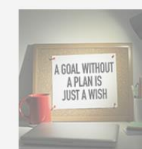
Submit

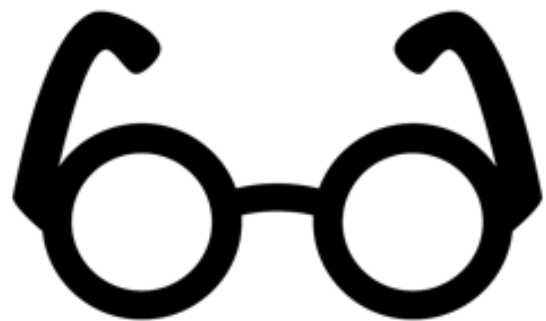
Browse

Preview

JSON

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





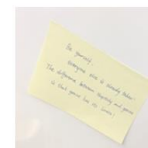
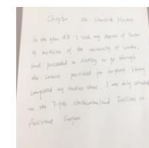
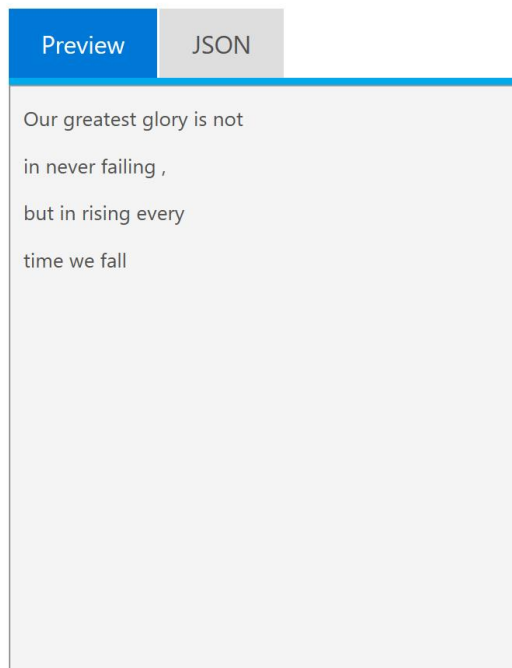
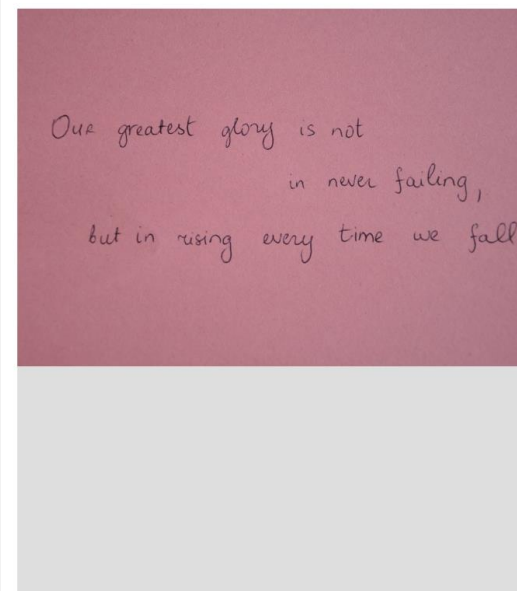
Vision and image analysis...



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.





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.



Image URL



```
all"
tions": [
  "text": "a group of people in front of Colosseum",
  "confidence": 0.84631330287730377

  "id": "c7112a85-6b04-44fd-96d6-057e70fb8763",
  "data": {
    "width": 600,
    "height": 399,
    "format": "Jpeg"
  },
  "tags": [],
  "properties": {
    "inantColorForeground": "Grey",
    "inantColorBackground": "White",
    "inantColors": [
      "Grey",
      "White"
    ]
  }
}
```



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.



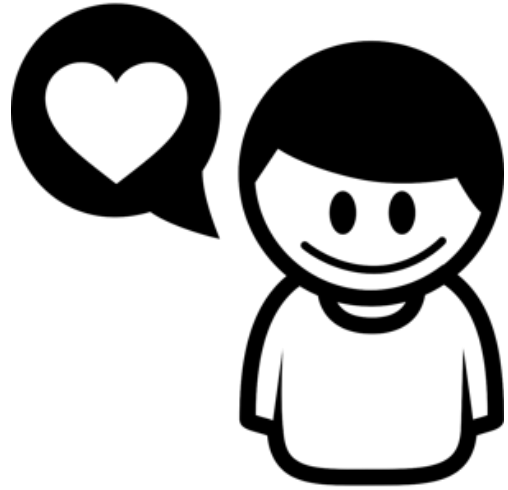
```
all"
tions": [
  "text": "a group of people in front of Colosseum",
  "confidence": 0.84631330287730377

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

': [],
': {
  nantColorForeground": "Grey",
  nantColorBackground": "White",
  nantColors": [
    "Grey",
    "White"
  ]
}
```

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**



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.



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

<input type="checkbox"/> Utterance	Labeled intent ?
<input type="checkbox"/> turn Power the lights in the Location	Control Lights -1 ...
<input type="checkbox"/> turn Power the Location lights	Control Lights -1 ...
<input type="checkbox"/> turn Power the Location lights	Control Lights -1 ...
<input type="checkbox"/> Location lights Power	Control Lights -1 ...

luis.ai
aka.ms/luis-api



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



Conversation as an interface



The Rise Of Intelligent Conversational UI

[UI](#) 54 # [Visual Design](#) 59 # [Interfaces](#) 32 # [User Interaction](#) 49



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



TextJSON

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

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



Want to go deep?



Demystifying Artificial Intelligence: Understanding Machine Learning | ORIGINAL

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

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

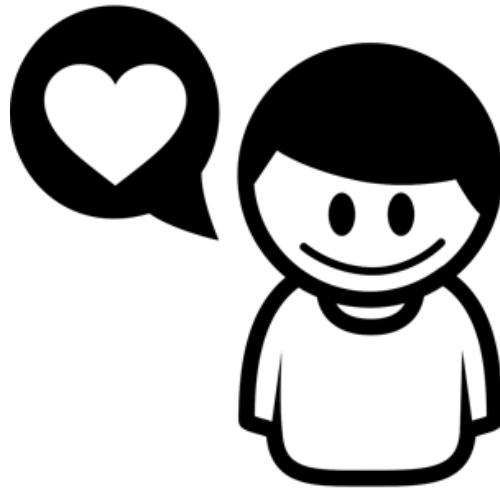
▶ 1. Introduction	1:44
🔒 2. What is Machine Learning	5:25
🔒 3. How We Teach Machines	5:48
🔒 4. Machine Learning to Help Humans	5:28
🔒 5. Tools for Machine Learning	3:44
🔒 6. Visual Uses	7:54
🔒 7. Speaking Human	6:07
🔒 8. Audio & Video	6:32

97 students are watching this class

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

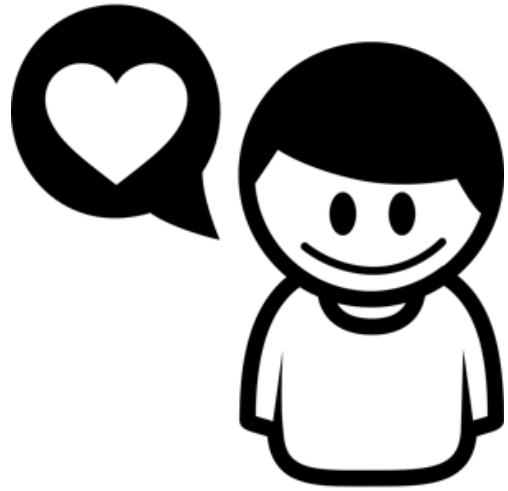
skl.sh/christian

Free with trial sign-up



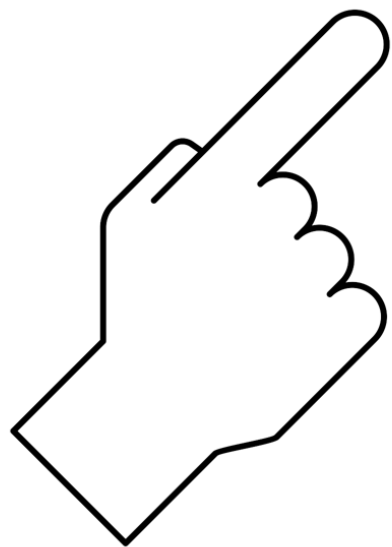
Don't forget to have fun!





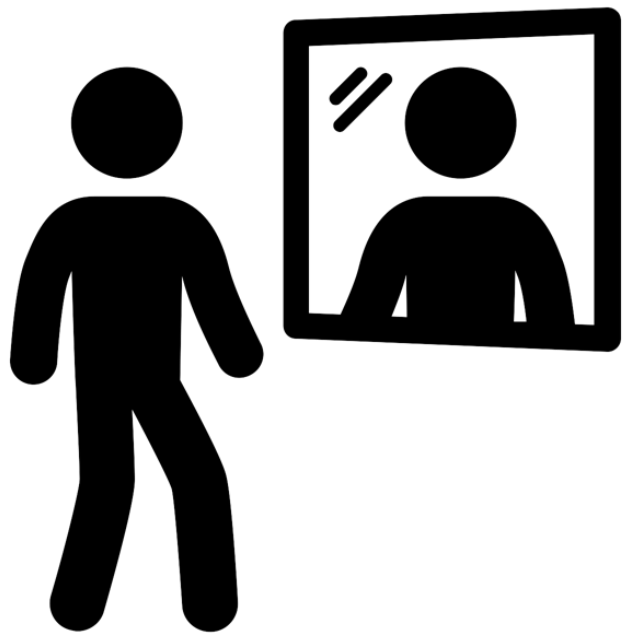
Suz Hinton





Categorising images by gesture

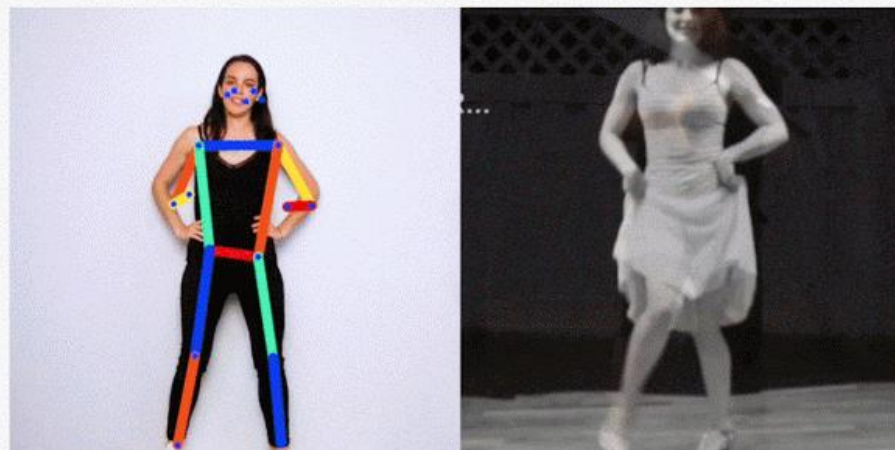




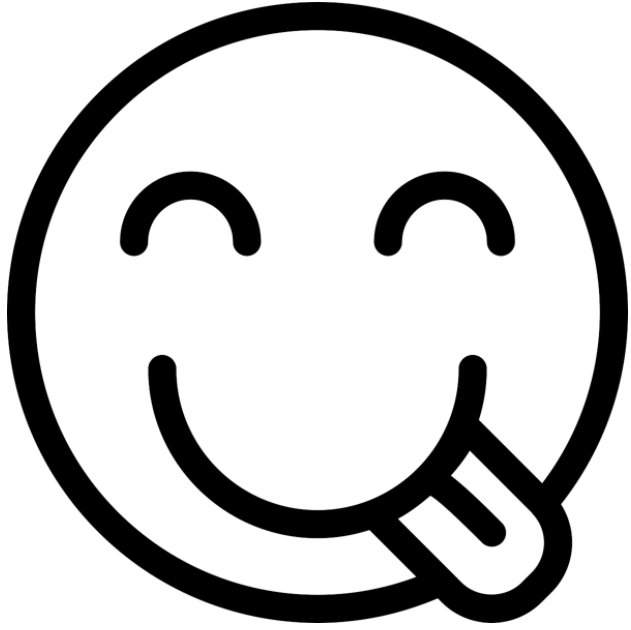
Find your moves



MOVE MIRROR



Make a GIF

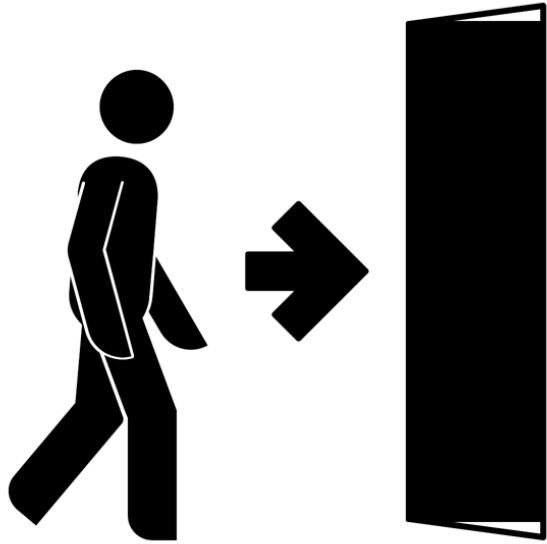


Stay silly...



Cassie Evans

<https://codepen.io/cassie-codes/pen/jKaVqo/>



Help the human

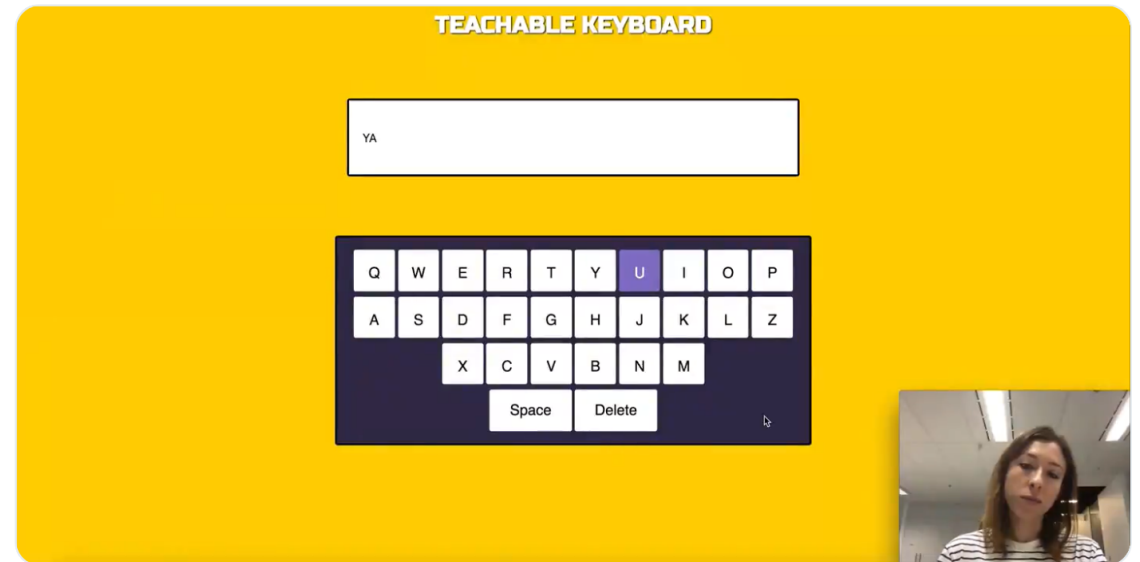


Charlie Gerard 
@devdevcharlie

Following



I used Tensorflow.js to build a quick prototype of head-controlled keyboard to allow people to communicate with simple head movements. Might not be much but it's amazing that you can do stuff like that in JS 🥰 [charliegerard.github.io/teachable-keyb...](https://charliegerard.github.io/teachable-keyboard/) [#tensorflowjs](#) [#javascript](#) [#ML](#) [#accessibility](#)



TEACHABLE KEYBOARD

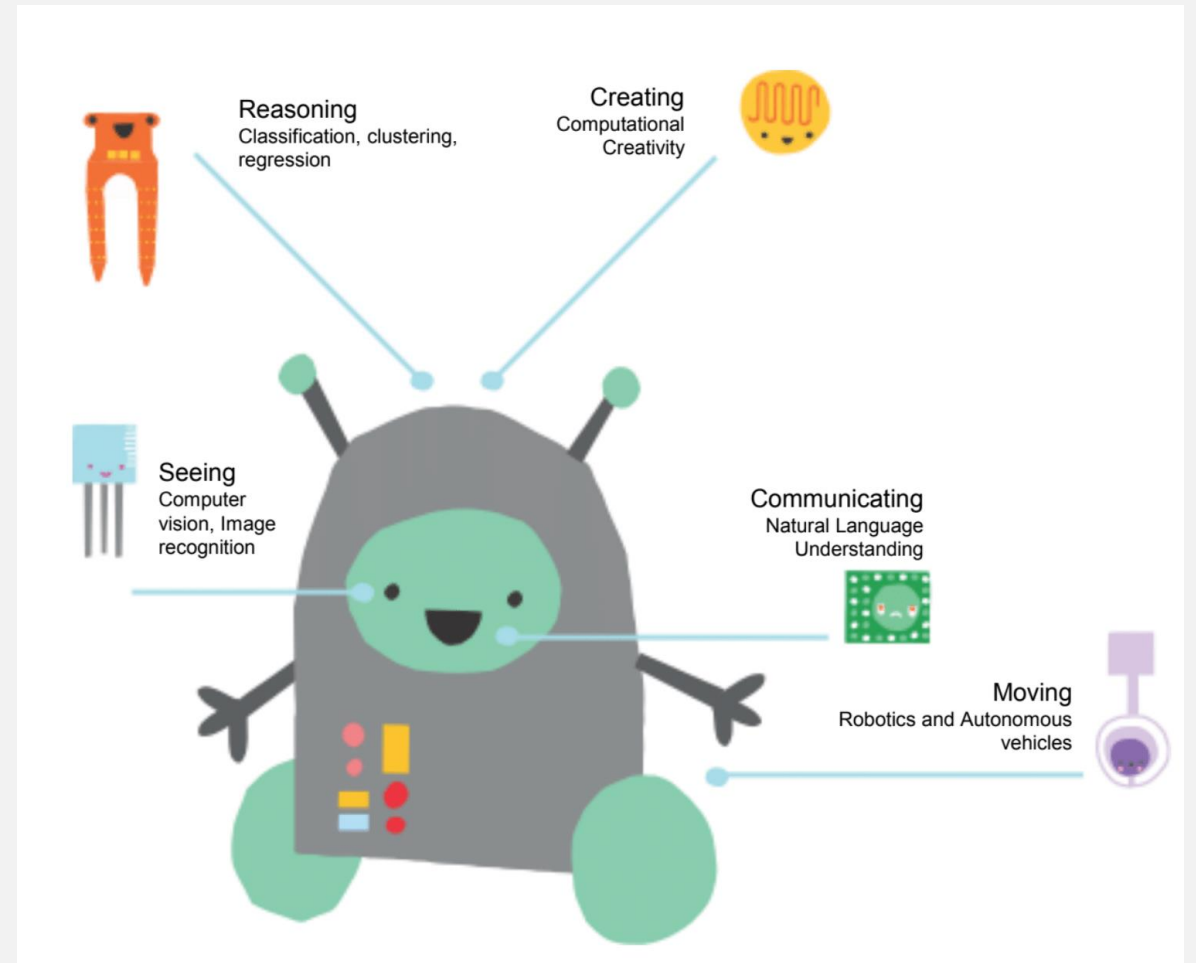
Write using motion control

Start Training





Preparing the next generation



Thanks!

Chris Heilmann

Christianheilmann.com

Developer-evangelism.com

@codepo8

