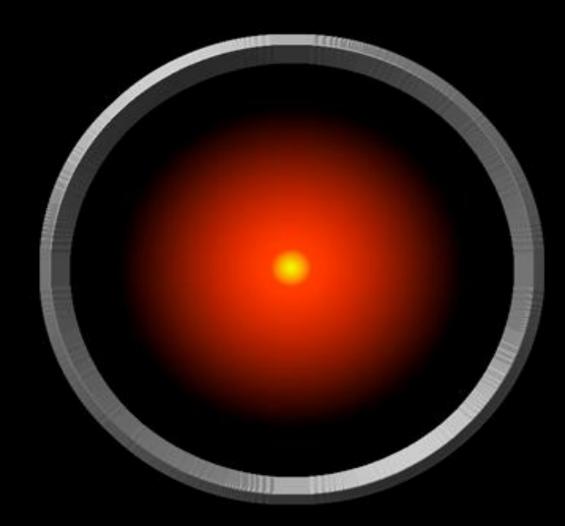
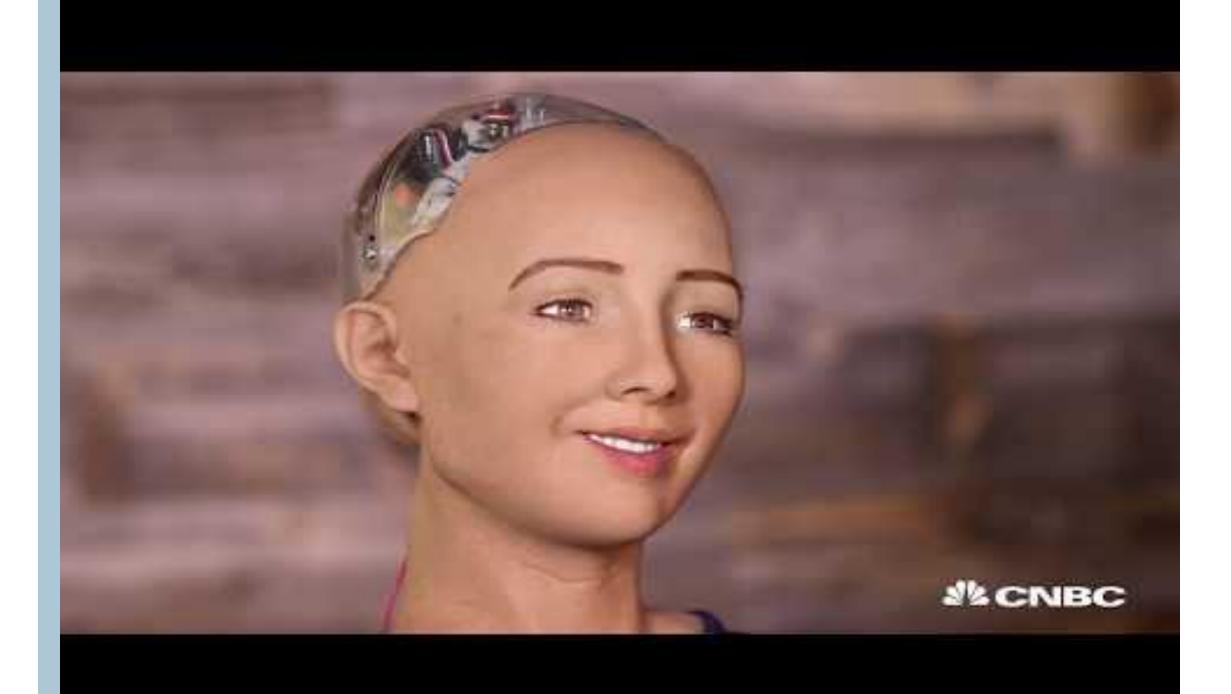


### I'm sorry Dave, I'm afraid I can't do that.







### Al timeline

HELL My Ne Is el	AME			
Eli: (196	~~~~	~~~~		ay 016)
Maze-Solving Mouse (1952)	Deep Blue (1997)	Siri (2011)	Alexa (2014)	AlphaGo (2016)
			o amazon alexa	🎨 AlphaGo

### Interactive System History

Multi-modal systems e.g., Microsoft MiPad, Pocket PC



TV Voice Search e.g., Bing on Xbox



2017

**Virtual Personal Assistants** 





Task-specific argument extraction

(e.g., Nuance, SpeechWorks)

to New York next week."

User: "I want to fly from Boston

Keyword Spotting (e.g., AT&T) System: "Please say collect, calling card, person, third number, or operator"

#### **TEMWATSON**

#### Intent Determination

Early 2000s

(Nuance's Emily<sup>™</sup>, AT&T HMIHY) User: "Uh…we want to move…we want to change our phone line from this house to another house" 52

DARPA CALO Project

#### Language Empowering Intelligent Assistant



Apple Siri (2011)



Google Now (2012)

Google Assistant (2016)



Microsoft Cortana (2014)



### Dialog System

#### Task-Oriented

- Personal assistant, helps users achieve a certain task
- Combination of <u>rules</u> and <u>statistical</u> components
- Examples:
  - POMDP for spoken dialog systems (Williams and Young, 2007)
  - End-to-end trainable taskoriented dialogue system (Wen et al., 2016)
  - End-to-end reinforcement learning dialogue system (Zhao and Eskenazi, 2016)

#### Chit-Chat

- No specific goal, focus on natural responses
- Using variants of seq2seq model
- Examples:
  - A neural conversation model (Vinyals and Le, 2015)
  - Reinforcement learning for dialogue generation (Li et al., 2016)
  - Conversational contextual cues for response ranking (AI-Rfou et al., 2016)



### Microsoft Chatbot Tay

- In 2016 March, Microsoft released Tay— a Twitter bot that the company described as an experiment in "conversational understanding." The more you chat with Tay, said Microsoft, the smarter it gets, learning to engage people through "casual and playful conversation."
- Unfortunately, it took less than 24 hours for Twitter to corrupt an innocent AI chatbot to a sexist and racist chatbot





......

......

.....

# JUST ASK amazon echo

#### Amazon Echo

#### amazon echo

Always ready, connected, and fast. Just ask.



- A woman in Portland, Oregon found out that her family's home digital assistant, Amazon's Alexa, had recorded a conversation between her and her husband without their permission or awareness, and sent the audio recording to a random person on their contacts list.
- Amazon later acknowledged the incident, and announced that it wasn't a hack, offering this narrative of what happened:

Echo woke up due to a word in background conversation sounding like "Alexa." Then, the subsequent conversation was heard as a "send message" request. At which point, Alexa said out loud "To whom?" At which point, the background conversation was interpreted as a name in the customers contact list. Alexa then asked out loud, "[contact name], right?" Alexa then interpreted background conversation as "right". As unlikely as this string of events is, we are evaluating options to make this case even less likely."



### Google Home The smart Speakers

**Equipped With Google Assistant** 

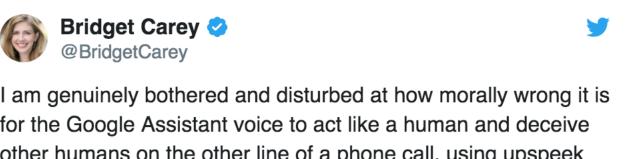
### Google Duplex



### Google Duplex: Concerns

People are concerned about an AI acting like a human, and deceiving the other side.





for the Google Assistant voice to act like a human and deceive other humans on the other line of a phone call, using upspeek and other quirks of language. "Hi um, do you have anything available on uh May 3?" #io18 1:46 AM - May 9, 2018

 $\bigcirc$  2,338  $\bigcirc$  1,091 people are talking about this

0



#### to a basic harvar rand

Acres

Sec. 1

What find of things are you arminen about? Perfuge correcting related to everylegenerit, shallers or relationation. 32.44

what I feel durin animous required. Monte devic Seconde I skirtl really from what in the pre-

I hepe heating services is not interfering with your functioning. And there you has service and all my scentions.

Mic: Pauser

24	10	-	-	-	-
*		34		14	-
C Serie	mont	Negative			10
-	20	. Art	-	. 80	+00
S Eme	pon :			_	
				11	
				_	



#### Questions on IoT and Smart Devices

- Is the device always listening?
- Who can see the photos and videos it takes?
- Are our private data being collected or being used somewhere?
- How can we ensure that our data is protected properly?
- Can the companies and government have access to my private data?
- What if my device or company's system get hacked?

#### **Questions on Surveillance**

- Mass surveillance, whether or not involving AI, is a serious breach of privacy.
- NSA's illicit gathering of telephone calls (revealed by Edward Snowden)
- Mass surveillance in the US is designed to target enemy combatants, but what if it turns towards ordinary citizens?
- On the other hand, mass surveillance can also significantly assist in crime prediction and prevention. Is it really worth it?
- What if the system gets hacked?

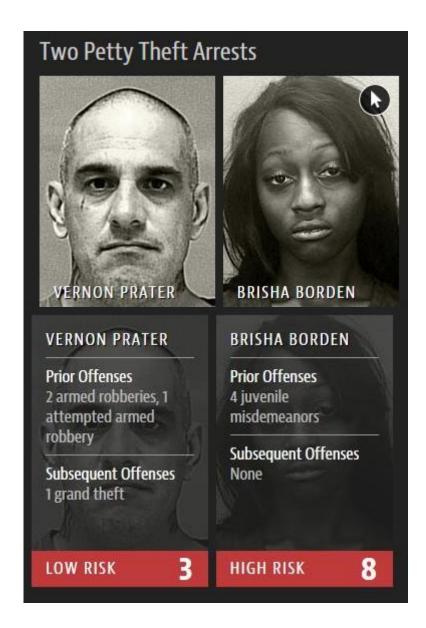


## **Predictive Policing**

- AI can be used to predict crime in advance.
- US, UK, and China are actively working on Predictive Policing that predicts crime with AI and data analysis.
- AI can aid coping with cybercrime.
- Detection and intervention of stealthy malicious attacks in advance.
- Dubai already has operational robocops and plans to have 25% of the police force as robocops by 2030.
- Identify wanted criminals, collect evidence, and patrol busy areas
- Several IoT devices are also already assisting police work in the field.
- E.g. GPS projectiles, body cams

# Questions of objectivity

- Brisha Borden (*right*) was assigned as High Risk of 8, while Vernon Prater (*left*) was assigned as Low Risk of only 3.
- Brisha Borden had only juvenile misdemeanors while Vernon Prater was a convicted felon for multiple armed robberies and subsequent grand theft.
- Brisha Borden was "black" while Vernon Prater was "white".





(a) Three samples in criminal ID photo set  $S_c$ .



(b) Three samples in non-criminal ID photo set  $S_n$ Figure 1. Sample ID photos in our data set.

NEURAL NETWORK USED TO PREDICT CRIMINAL TENDENCY FROM FACIAL FEATURES