

# Artificial Intelligence in the World

Prof. Levy

Fromm Institute

Spring Session, 2017

# Lecture 7 agenda

- Explain the reasoning of ANNs
- Rules for AI -- identify yourself!
- **Job loss / unemployment**
- **Ethics & dangers of AI** – The Singularity and other concerns
- **Privacy** concerns
- Who is doing something about ethics of AI?
- And all the applications of AI we haven't talked about yet

# We need AI (ANNs) to be able to explain their reasoning

- *Article by R. Kanai -- CEO of Araya, a Tokyo-based startup*
- **What would the AI have to do?**
  - Metacognition – introspective reports of mental state
  - Internal models ... of the world
- **“Peeking inside the black box”**
  - safety-critical applications (e.g., air traffic control)
  - pruning the tree

# Turing Red Flag Law (see handouts)

- *Toby Walsh is Professor of Artificial Intelligence*
- *at the University of New South Wales (Australia)*
  
- *An autonomous system should be designed so that it is*
- ***unlikely to be mistaken** for anything besides an autonomous system, and*
- ***should identify itself** at the start of any interaction with another agent.*

# Job loss / unemployment

- **Obama white house concerns** (*NY Times*)
  - More CS education / critical thinking
  - Apprenticeships & retraining
- **Pros & Cons of AI** (*Datamation*)
  - What are you trying to do?
- **Report on automation & employment** (*Economist*)
  - Routine work being displaced

# Ethics & dangers of AI

- **The Singularity**

- See Kurzweil's book, *The Singularity is Near*
- In a communication to *Futurism*, Kurzweil states:
- **2029** is the consistent date I have predicted for when **an AI will pass a valid Turing test and therefore achieve human levels of intelligence.**
- **I have set the date 2045 for the 'Singularity'** which is when we will multiply our effective intelligence a billion fold by merging with the intelligence we have created.
- For Kurzweil, the singularity is an opportunity for humankind to improve. He envisions the same technology that will make AIs more intelligent giving humans a boost as well.

# Reactions to prospects of the Singularity

NOT SO FAST



**STEPHEN HAWKING**  
Not afraid of  
black holes.  
A.I. is another story.



**BILL GATES**  
First you'll lose  
your job. Then it  
gets scary.



**STUART RUSSELL**  
Earth for  
the earthlings!



**NICK BOSTROM**  
Prepare for  
"Disneyland  
without children."



**MAX TEGMARK**  
Uh, can we  
talk about this?



**DEMIS HASSABIS**  
Full speed  
ahead!



**PETER THIEL**  
Will be a winner  
either way.



**STEVE WOZNIAK**  
Resigned to  
being a robot's pet.



**SAM ALTMAN**  
Sees intergalactic  
domination—or  
extinction.



**ELON MUSK**  
Eyeing the  
next flight to  
Mars.



**LARRY PAGE**  
Green-lighted  
Google Brain.  
'Nuff said.



**YANN LECUN**  
Chill, people!  
We got this.



**ANDREW NG**  
Trust the robot.



**MARK ZUCKERBERG**  
Worried? Tell  
my A.I. butler.



**RAY KURZWEIL**  
Eager to  
be a cyborg.

HIT THE GAS

From <http://www.vanityfair.com/news/2017/03/elon-musk-billion-dollar-crusade-to-stop-ai-space-x>

## *“Let's Stop Freaking Out About Artificial Intelligence”*

- **Eric Schmidt, Sebastian Thrun** article in *Fortune*
  - Eric Schmidt is executive chairman of Alphabet, **Google's** parent company.
  - Sebastian Thrun is president and chairman of Udacity, an online education company
  - <http://fortune.com/2016/06/28/artificial-intelligence-potential/>
- **Google**, alongside many other companies, is doing rigorous research on **AI safety**, such as how to ensure people can interrupt an AI system whenever needed, and how to make such systems robust to cyberattacks.
- ...ultimately the hypothetical, long-term concerns are far outweighed by our excitement for the endless possibilities.

# Privacy -- Power of the large company players

- ***The world's most valuable resource is no longer oil, but data***  
*(Economist)*
  - The giants' **surveillance** systems span the entire economy: **Google** can see what people search for, **Facebook** what they share, **Amazon** what they buy.
  - When a new product or service gains traction ... [they] copy it or simply **buy the upstart** before it becomes too great a threat.
    - Many think Facebook's \$22bn purchase in 2014 of **WhatsApp**, a messaging app with fewer than 60 employees, falls into this category of "shoot-out acquisitions" that eliminate potential rivals.
  - One solution: regulation -- companies could be forced to **reveal to consumers what information they hold and how much money they make from it.**

# Privacy – other privacy concerns (1)

- **Amazon** releases Echo recording in murder case (see article)
- **Facebook** using face recognition (IEEE)
  - ... class-action lawsuit against Facebook alleging that the company's use of face recognition technology violates an Illinois law passed in 2008. The Biometric Information Privacy Act (BIPA) sets limits on how companies can store and use people's bio- metric identifiers, which the law defines as **fingerprints, voiceprints, retina or iris scans, and scans of hand or face geometry.**
  - Facebook researchers pioneered a new approach to face recognition that relies on machine learning, introducing their DeepFace system in a 2014 paper. In the report, the researchers describe **training their system using a data set of 4.4 million labeled faces drawn from Facebook photographs.**
- (The **FBI's FACES face recognition database** mostly contains images of law-abiding citizens taken from driver's license and passport photos.)

## Privacy – other privacy concerns (2)

### • **4 Awesome Illegal Uses for Alexa**

- article by Shelly Palmer on LinkedIn
- If you let your imagination run wild in the world of automatic speech recognition (ASR) and natural language understanding (NLU), and you throw in a little fear, uncertainty and doubt (FUD), you can come up with several illegal uses for systems such as **Alexa Voice Service, Google Home, Siri, and Cortana** that will give you pause.
- **“Alexa. Find me a free version of this movie without commercials.”**
- **“Alexa. What song is that?”**
- **“Alexa. Get this song (video, movie, TV series) for me for free.”**
  - Amazon owns IMDb
  - Amazon knows everything an advertiser needs to know about you, including your full contact information, credit card info, purchase history, streaming history (for those on Prime) and more.
- **“Alexa. Record the next three hours of audio from Echo Dot #2.”**

# Apple claims sensitivity to privacy

- Article in MIT *Technology Review*
- **Apple** showed off a string of new iPhone features powered by recent advances in artificial intelligence— many of them aping ones already launched by rival Google.
- But Apple's announcements of features like facial recognition or software that knows what's in your photos, made during its annual Worldwide Developer Conference, were distinct in how much they **emphasized privacy**.

# Who is doing something about AI ethics? (1)

- **“Partnership on AI”** [NY Times, HuffPost]
  - **Google, Amazon, Facebook, IBM, Microsoft, Apple**
  - Impact of A.I. on jobs, transportation and even warfare.
- **CMU**
  - Research center that focuses on the **ethics of artificial intelligence**
  - Carnegie Mellon drew national attention when a group of 36 technical staff members and four faculty members left to join a new self-driving car laboratory that Uber established in Pittsburgh
- **MIT**
- **UC Berkeley – Center for Human-Compatible AI**

# Who is doing something about AI ethics? (2)

Stanford – 100 year study [no Singularity]

“assess the technological, economic and policy implications of potential AI applications in a societally relevant setting”

**Transportation:** Autonomous cars, trucks and, possibly, aerial delivery vehicles may alter how we commute, work and shop and create new patterns of life and leisure in cities.

**Home/service robots:** Like the robotic vacuum cleaners already in some homes, specialized robots will clean and provide security in live/work spaces that will be equipped with sensors and remote controls.

**Health care:** Devices to monitor personal health and robot-assisted surgery are hints of things to come if AI is developed in ways that gain the trust of doctors, nurses, patients and regulators.

**Education:** Interactive tutoring systems already help students learn languages, math and other skills. More is possible if technologies like natural language processing platforms develop to augment instruction by humans.

# Who is doing something about AI ethics? (3)

**Entertainment:** The conjunction of content creation tools, social networks and AI will lead to new ways to gather, organize and deliver media in engaging, personalized and interactive ways.

**Low-resource communities:** Investments in uplifting technologies like predictive models to prevent lead poisoning or improve food distributions could spread AI benefits to the underserved.

**Public safety and security:** Cameras, drones and so ware to analyze crime patterns should use AI in ways that reduce human bias and enhance safety without loss of liberty or dignity.

**Employment and workplace:** Work should start now on how to help people adapt as the economy undergoes rapid changes as many existing jobs are lost and new ones are created.

# Who is doing something about AI ethics? (4)

**RPI – a look at “Westworld”** (local newspaper article)

**“Ethically Aligned Design”** – IEEE

*empowering technologists to prioritize ethical considerations in the creation of Artificial Intelligence and Autonomous Systems.*

**Future of Life Institute** (NY Times) [Kevin Kelly]

Principles for AI researchers

**Ethics for robots** (IEEE) – interactions of humans & robots

# AI applications we have not discussed

- **Recommendation engines** (Amazon, Netflix, ...)
- **Business optimization** (factory, warehouse, delivery, ...)
- **Military** (drones, aircraft, ships, ...)
- **Biology / Medicine** (drug research, medical diagnosis, ...)
- **Financial** (market prediction, insurance risk, ...)
- **Engineering** (design, development, testing)