CST3140: Novel Interaction Technologies

Week 17 Lecture:

Human Interaction with Artificial Intelligence

Plan

- AI is one of the hot topics of CS
- Used in lots of areas for lots of things
  - Many of which we're not really aware of

- How do people interact with AI systems?
- Is AI anything special?
- Does AI pose particular challenges for HCI, Interaction Design, etc
- And what can we do about these challenges?

What is AI?

- any human-like intelligence exhibited by a computer or other machine
- … the ability of a computer to mimic capabilities of the human mind—learning from examples and experience, recognizing objects, understanding language, making decisions, solving problems
- … to perform functions a human might perform, such as greeting a hotel guest or driving a car.

"AI is whatever hasn't been done yet"

https://www.ibm.com/cloud/learn/what-is-artificial-intelligence

”AI will help humanity reach its full potential”

- Mark Zuckerberg –

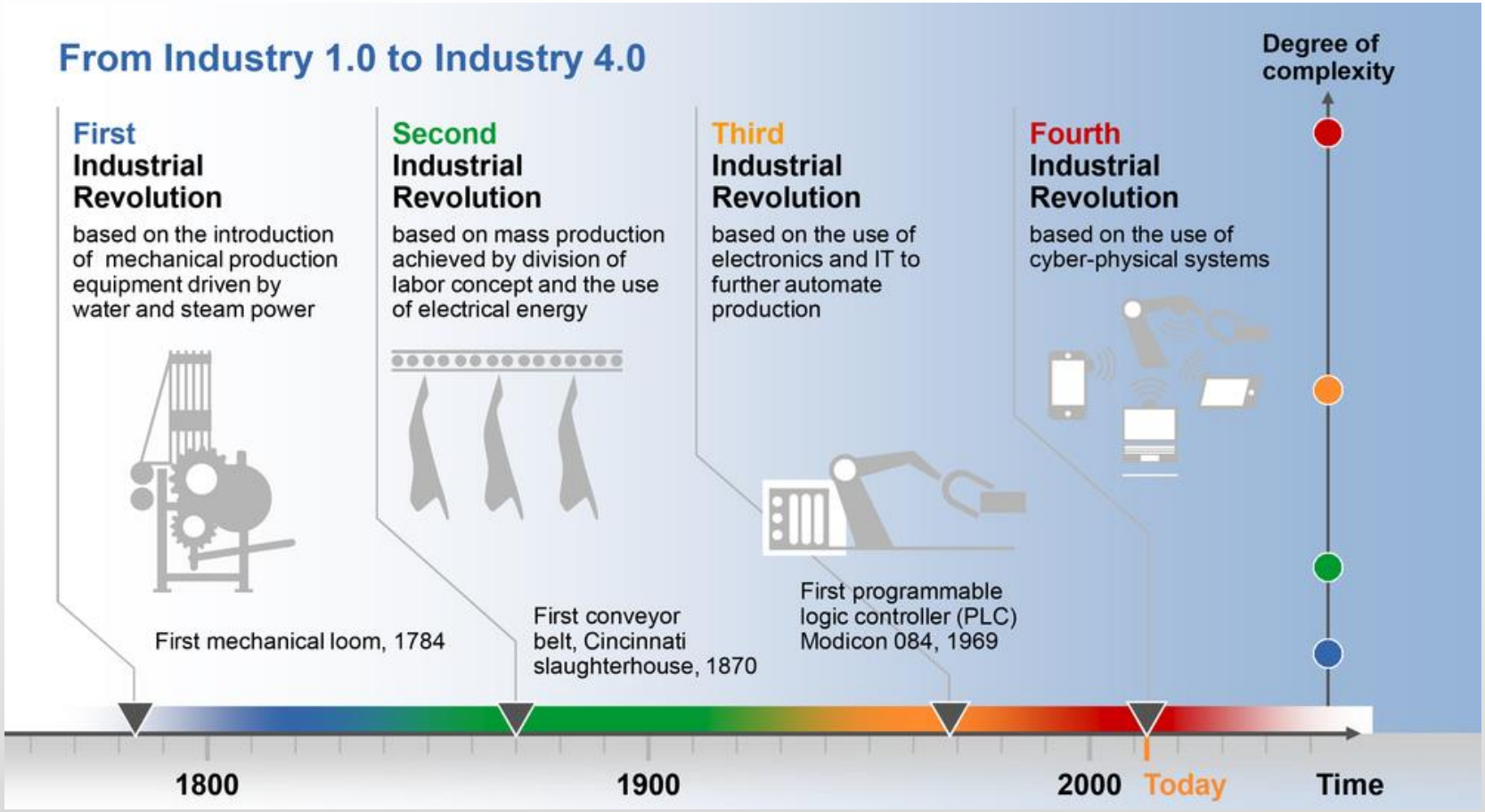
“AI is the biggest existential threat to humanity”

- Elon Musk –

“The development of full AI could spell the end of the human race.”

- Stephen Hawking -

History


From Industry 1.0 to Industry 4.0
Degree of complexity
First Industrial Revolution
based on the introduction of mechanical production equipment driven by water and steam power
Second Industrial Revolution
based on mass production achieved by division of labor concept and the use of electrical energy
Third Industrial Revolution
based on the use of electronics and IT to further automate production
Fourth Industrial Revolution
based on the use of cyber-physical systems
First mechanical loom, 1784
First conveyor belt, Cincinnati slaughterhouse, 1870
First programmable logic controller (PLC) Modicon 084, 1969
1800
1900
2000
Today
Time

Characteristics of AI

Human intelligence involves….

- Reasoning
- Planning
- Knowledge
- Communication – e.g. through NL
- Perception – gathering information about the environment
- Consciousness

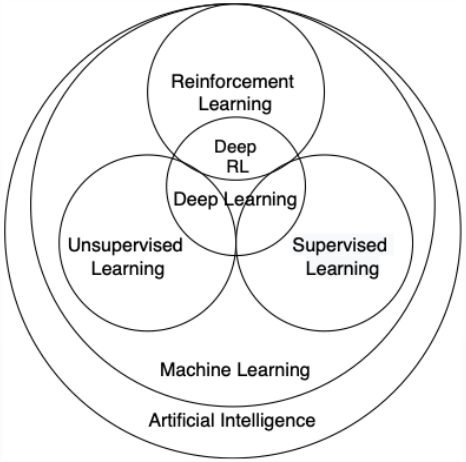
'*Strong AI* aims to create intelligent machines that are indistinguishable from the human mind'

But most AI systems aim, not at 'general intelligence' but at solving specific problems

The 'Turing Test'
https://www.youtube.com/watch?v=3wLqsRLvV-c

https://www.dexlabanalytics.com/blog/5-major-problems-in-ai-artificial-intelligence

What is AI?



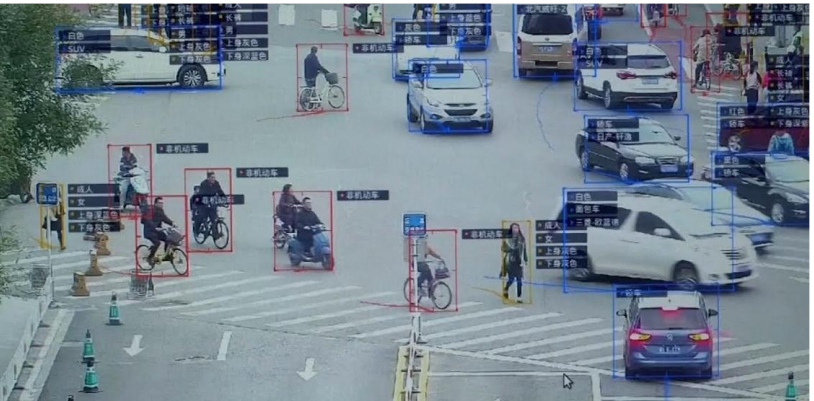
- ‘Strong AI’ vs. ‘Weak AI’?

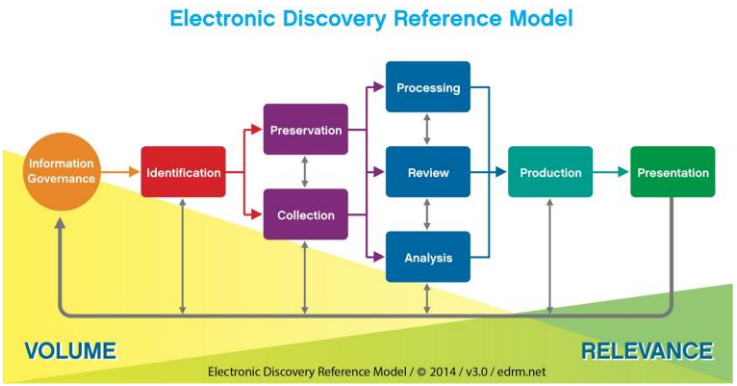
- Machine learning – for the purposes of today, equivalent to AI
- Classification, prediction, pattern detection, etc
- Automation and autonomy

https://www.youtube.com/watch?v=2ePf9rue1Ao

Applications

- Many application areas - involving human & computer working together





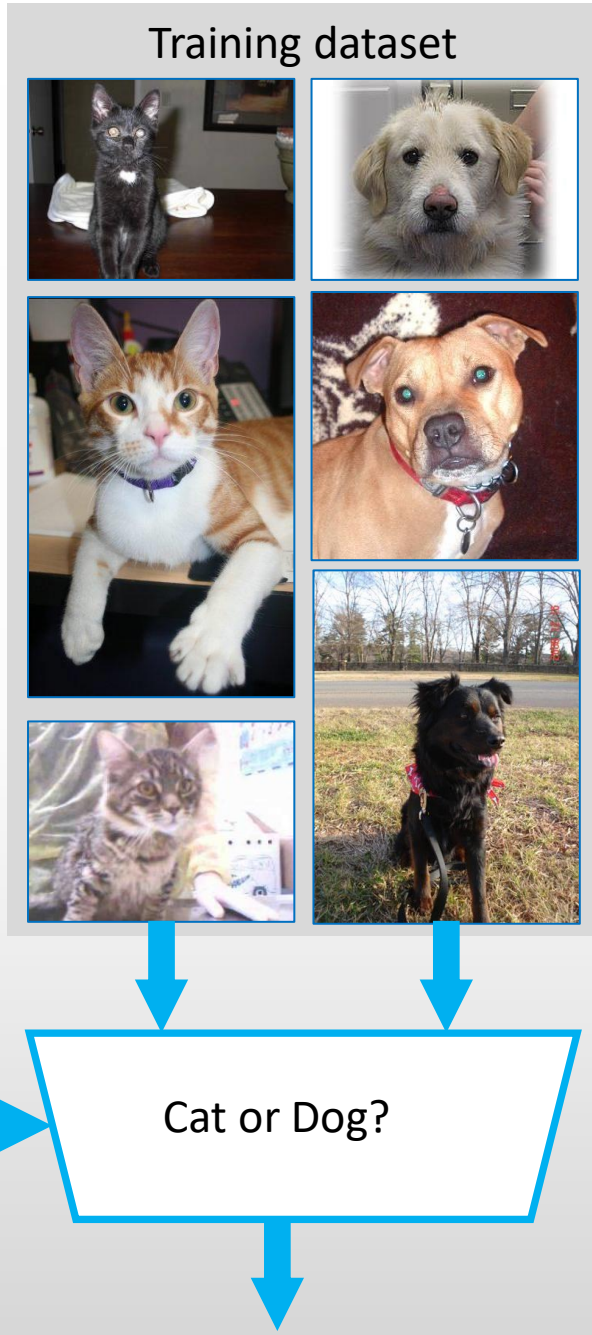
Machine learning

Typical ML problem: classification

Given data items, should they be classified as type A or type B?

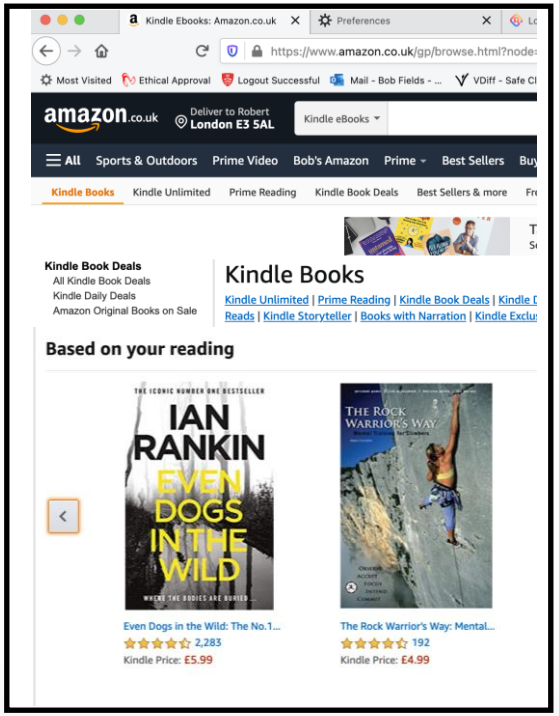
ML system ‘trained’ with classified training data from which it ‘learns’ a model that can decide A or B

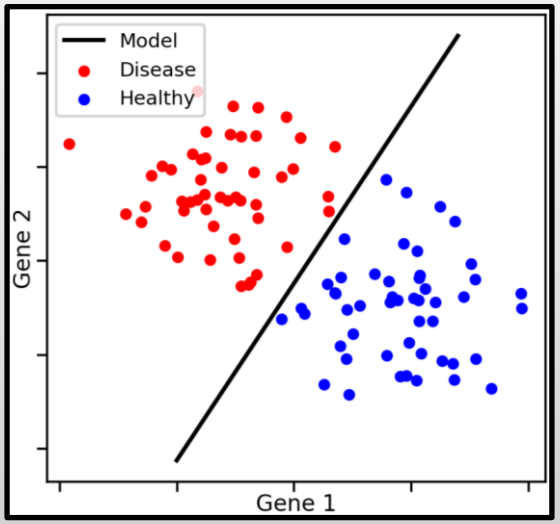
Once trained, the ML system will be able to classify new, unseen data



Uses?

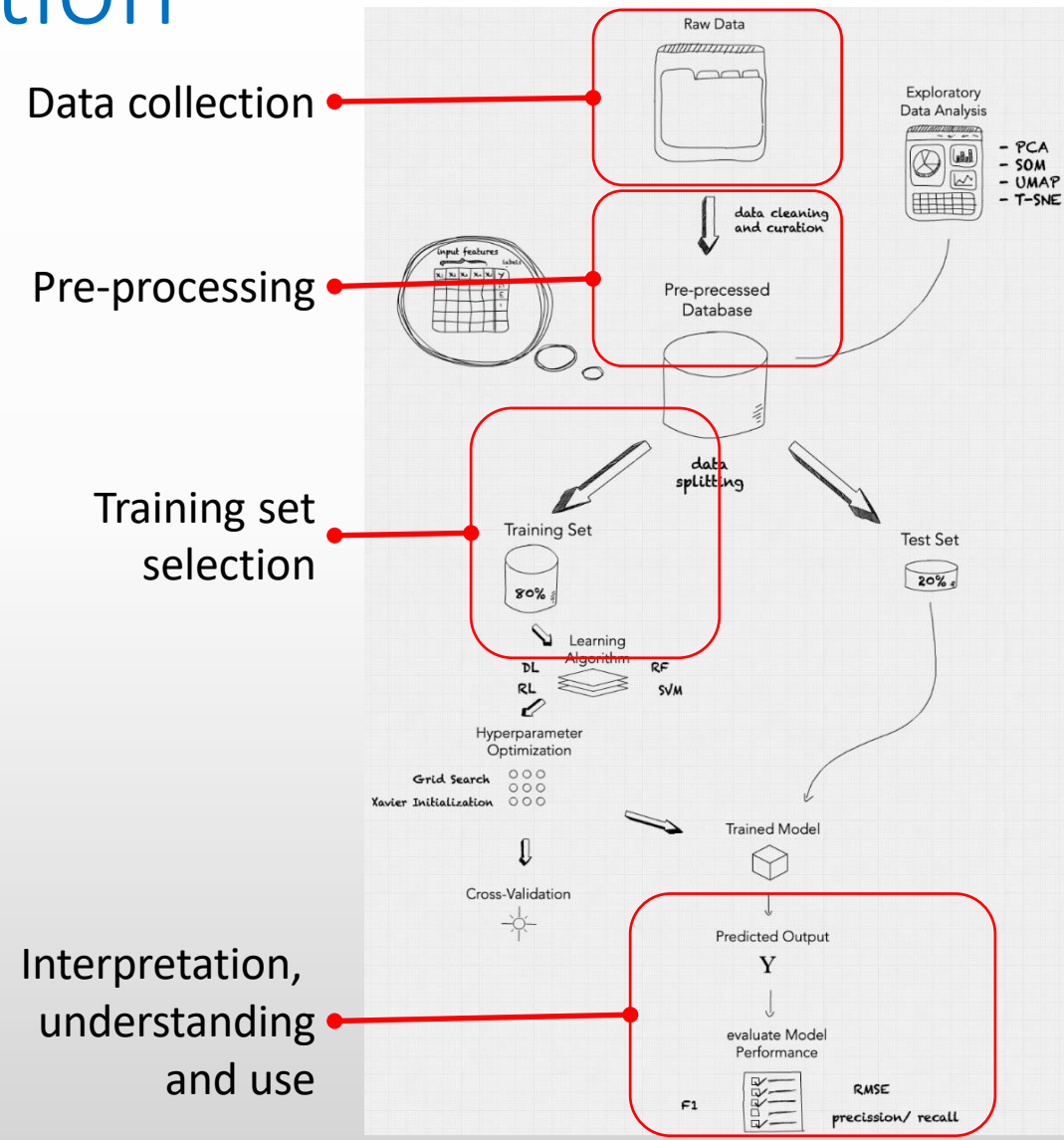
- Recommender systems
- Spam filtering
- Data analysis
- Image recognition
- E-Discovery
- Medical image analysis and diagnosis
- …
- Where's the human interaction?
  - Choosing or creating training data
  - Consuming and understanding the results





Human interaction

- Human interaction throughout the AI / Machine learning ‘process’



https://towardsdatascience.com/tagged/supervised-learning

Human interaction

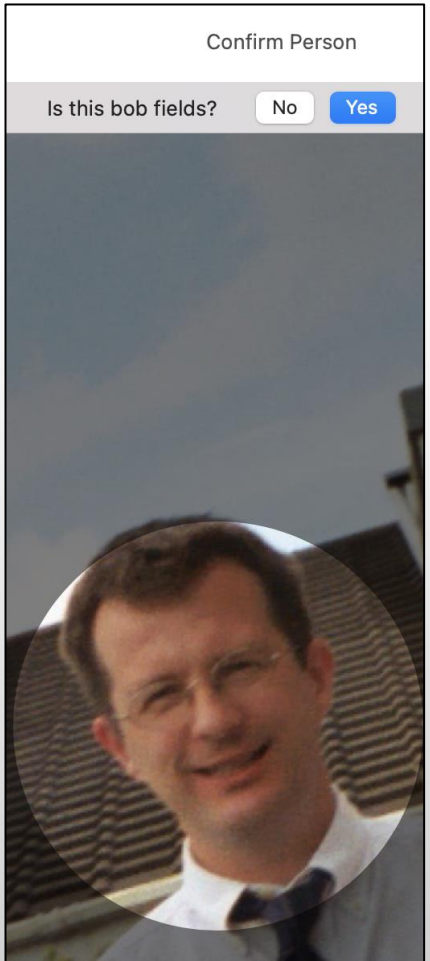
- Humans choose training data sets? How?
- Will the user understand enough about the models and learning algorithms to select data? Will they want to?

Active learning

- Training is an interactive process

E.g. Apple Photos

- Learns to identify people in pictures
- User can identify unknown faces
- Machine prompts user for confirmation or classification
- Chooses ‘most informative’ items for user to classify
- Performance improves over time



Selecting training data

A cautionary tale:

AI system to distinguish friendly & enemy tanks

100% accuracy!

But: friendly photos taken on sunny days

AI can classify the weather – not the tanks!

Many other examples of bias – e.g. face recognition for surveillance

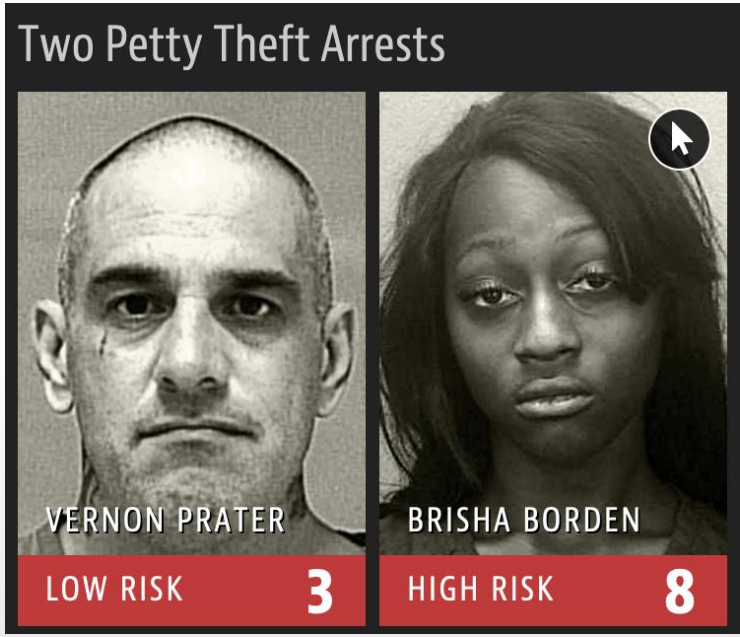
- Training data selection is critical!

https://www.youtube.com/watch?v=ukzFI9rgwfU

See https://www.gwern.net/Tanks

What's the problem with AI?

- Output of an AI / Machine learning system:
  - categorisation or prediction
- E.g. used to assess risk, make decisions
- Biased, potentially wrong
- Solution
  - Make sure the system gets it right?
  - Or something more subtle: explanation



https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

Explainable AI

“Defendants rarely have an opportunity to challenge their assessments. The results are usually shared with the defendant’s attorney, but the calculations that transformed the underlying data into a score are rarely revealed”

Key research area: AI systems should be able to explain their conclusions – just as a human could

https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

Consuming the results of AI

- Users receive recommendations from a ML system
- Decisions
  - Does this make sense? What has the machine done?
  - What does it mean?
  - What should I do about it?
- Or the user interface is adapted based on ML
  - E.g. by learning common tasks for this user, inferring user goal, etc

Problems with AI

- Adaptation
- Understandability & predictability
- Explain-ability
- Trust-ability
- Ethics?

Remember Week 5 …

Usability Principles – guidance for usable design

How do these apply to AI-based systems?

- If a system can adapt or learn from past events
- Will it be predictable?
- Is it consistent?
- Will the user be able to discover how it works?
- Etc.

Visibility
- Let the user see what they can do

Feedback
- Tell the user what's going on

Consistency
- Similar actions & displays should do or mean similar things

Non-destructive operations
- So the user can learn by exploring – e.g. undo

Discoverability
- Let the user find out how things work

Norman & Nielsen: Gestural Interfaces: A Step Backward In Usability. Interaction

The Ironies of Automation

Automation is intended to help, but sometime makes things harder!

Automation does things, automatically while humans monitor …

(This is generally hard to do for any length of time)

… and take over if anomalies arise

also hard to do – system complex; user may not know system state & automation intention

‘Skills fade’ when tasks are automated

Bainbridge, L. (1983). "Ironies of automation". Automatica. ***19*** *(6): 775–779*

The Ironies of Automation

Automation, and its problems, discussed for many years – not only in relation to AI

- Complexity
- Lack of transparency – difficulty to understand
- Changes human task - not always for the better!

https://www.youtube.com/watch?v=VgQwHDFohTo

The 'ironies of automation' refers to a set of unintended consequences as a result of automation, that could detrimentally affect human performance on critical tasks. Automation might increase human performance issues, rather than eliminate them

Bainbridge, L. (1983). "Ironies of automation". Automatica. ***19*** *(6): 775–779*

Re-shaping HCI?

- Does AI change understandings of what's going on?
- Interaction with computers: largely stimulus-response: do something, get an answer
- What's the computer doing when not interacting?
  - Autonomously doing things in the background
- From Human-Computer Interaction … to *Human-Computer Integration*
- Design for integration of human & computer activities, working in partnership, 'human-machine teaming'

Farooq and Grudin "Human-Computer Integration" ACM Interactions. Nov 2016.

Summary

- Increasingly, interactive systems involve AI/ ML elements
  - Many application areas – self-driving vehicles, fitness trackers, data analysis, …
- Raises important issues – human choice, autonomy, adaptation, …
- Creates potential usability problems: predictability, consistency, visibility, etc
- May lead to rethinking the relationship between humans and machines
  - Integration as well as interaction