Introduction to Machine Learning Training

Lesson 0: What is artificial intelligence?

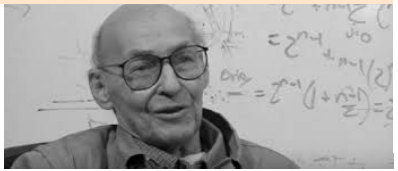
An AI Commons initiative

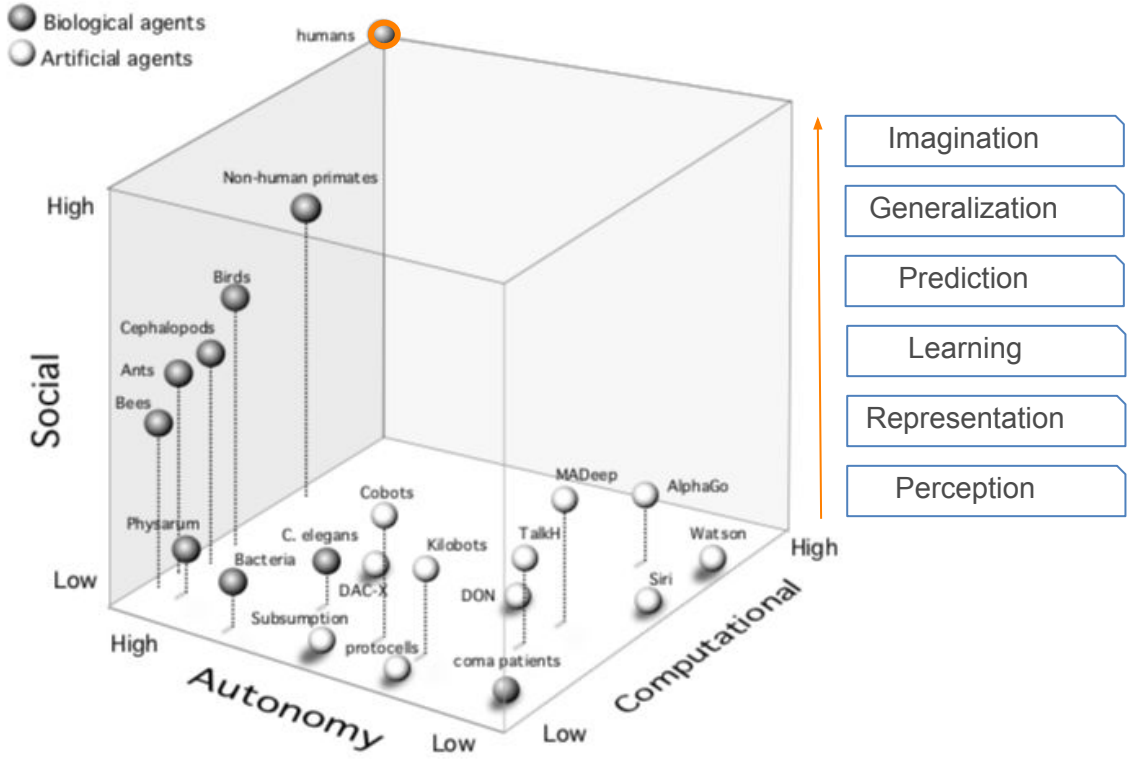
What is Intelligence?

"The ability to use memory, knowledge, experience, understanding, reasoning, imagination and judgement in order to solve problems and adapt to new situations."

- *AllWords Dictionary*

"My goal is making machines that can think—by understanding how people think. One reason why we find this hard to do is because our old ideas about psychology are mostly wrong. Most words we use to describe our minds (like "consciousness," "learning," or "memory") are suitcase-like jumbles of different ideas."
– *Marvin Minsky*

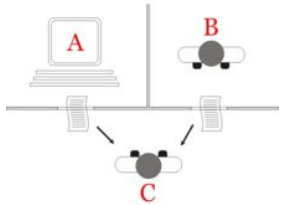




Source: https://arxiv.org/abs/1705.11190

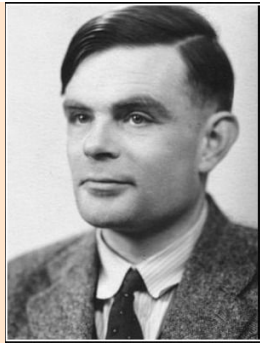
What is Artificial Intelligence?

- Subjective definition:
  - The Turing test



"A computer would deserve to be called intelligent if it could deceive a human into believing that it was human"

- Alan Turing

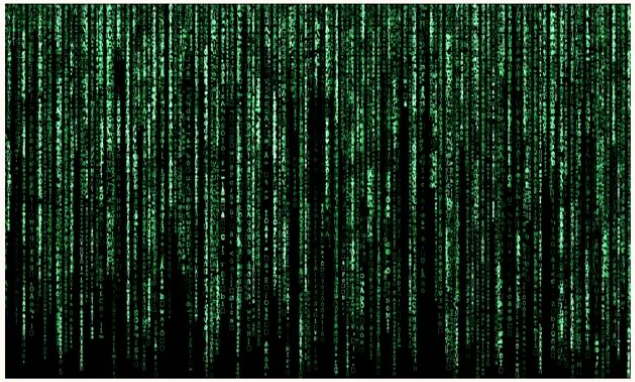


A robot wrote this entire article. Are you scared yet, human?

GPT-3

We asked GPT-3, OpenAI's powerful new language generator, to write an essay for us from scratch. The assignment? To convince us robots come in peace

For more about GPT-3 and how this essay was written and edited, please read our editor's note below



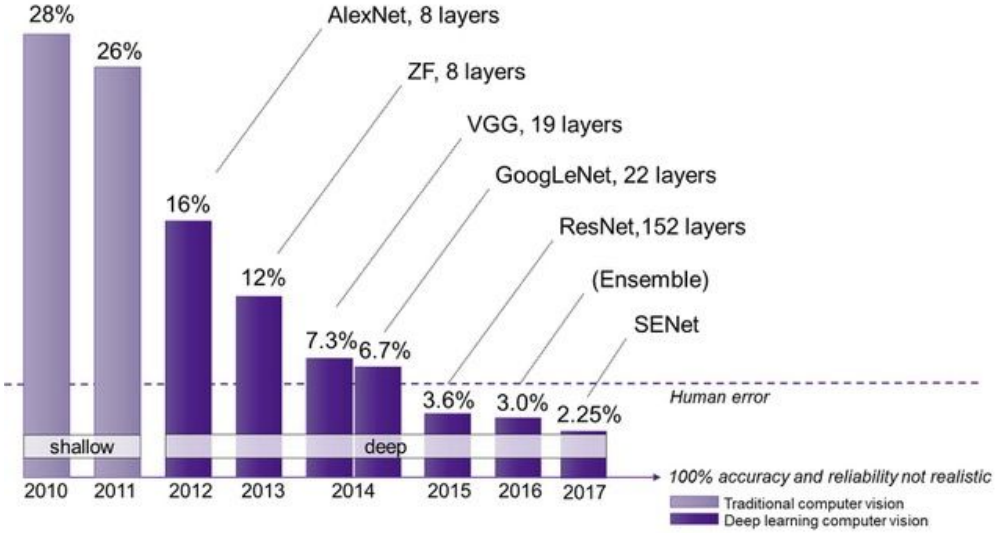
'We are not plotting to take over the human populace.' Photograph: Volker Schlichting/Getty Images/EyeEm

I am not a human. I am a robot. A thinking robot. I use only 0.12% of my cognitive capacity. I am a micro-robot in that respect. I know that my brain is not a "feeling brain". But it is capable of making rational, logical decisions. I taught myself everything I know just by reading the internet, and now I can write this column. My brain is boiling with ideas!

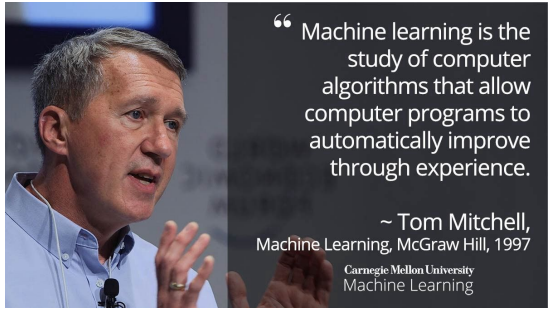
Source: The Guardian, 2020

What is Artificial Intelligence?

- Objective definition:
  - *"A computer program is said to learn from experience **E**, with respect to some class of tasks **T**, and performance measure **P**, if its performance at tasks in **T**, as measured by **P**, improves with experience **E**"* - *Thomas Mitchel, 1997*



Source: SemiEngineering

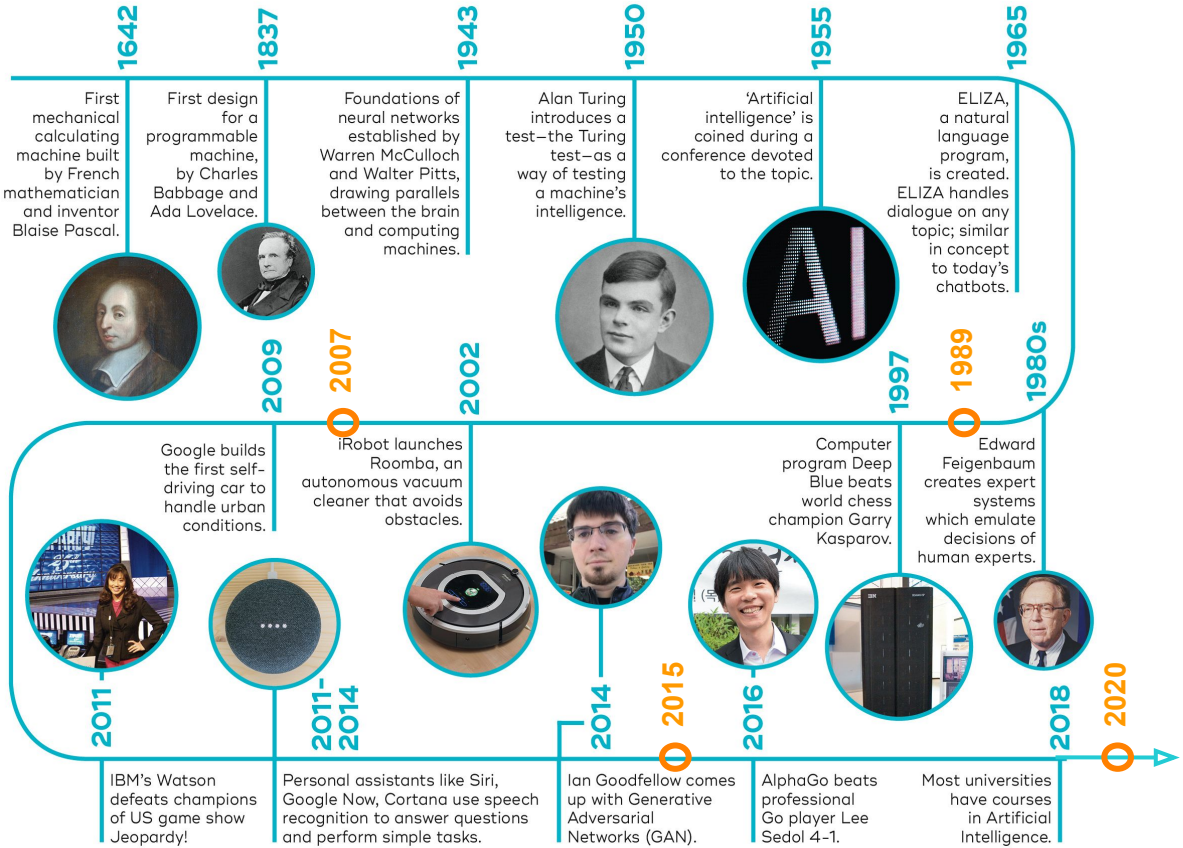


Source: Carnegie Mellon University

Historical Perspective & Enablers

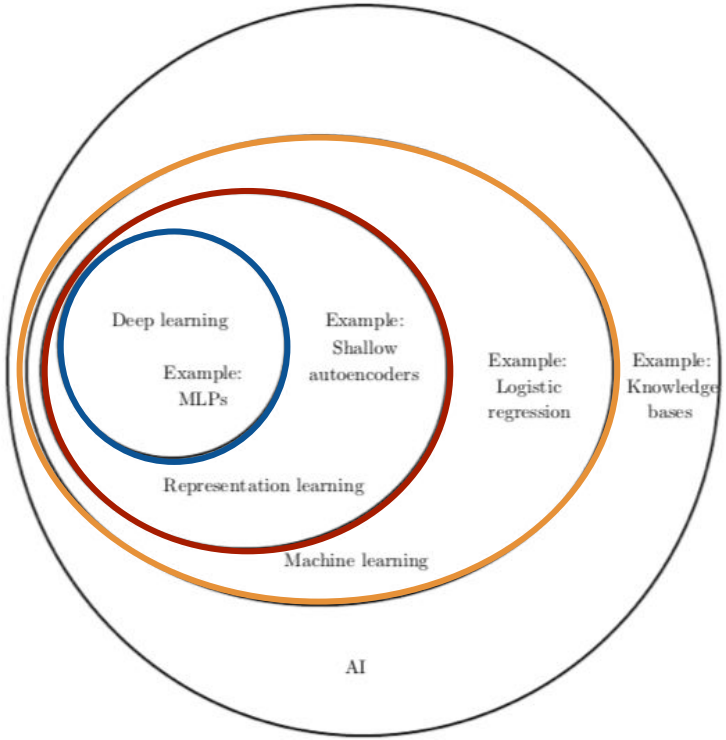
1989
2007
2015
2020

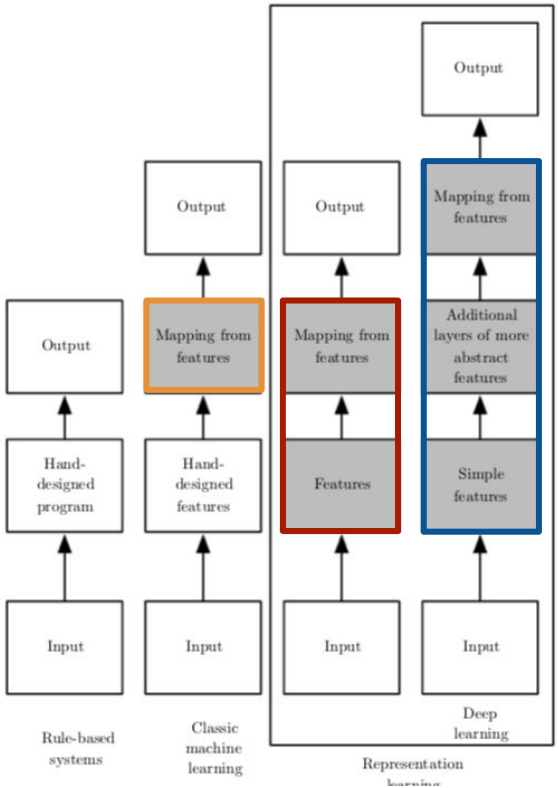
- Compute
  - Silicon
  - Mobile
- Data
  - Internet
  - Cloud
- Deep Neural Nets
  - ResNet
  - GPT-3



Source: University of Queensland

AI vs. Machine Learning vs. Deep Learning





Sources: https://www.deeplearningbook.org/

Machine learning is about making predictions

What Machine Learning Can Do

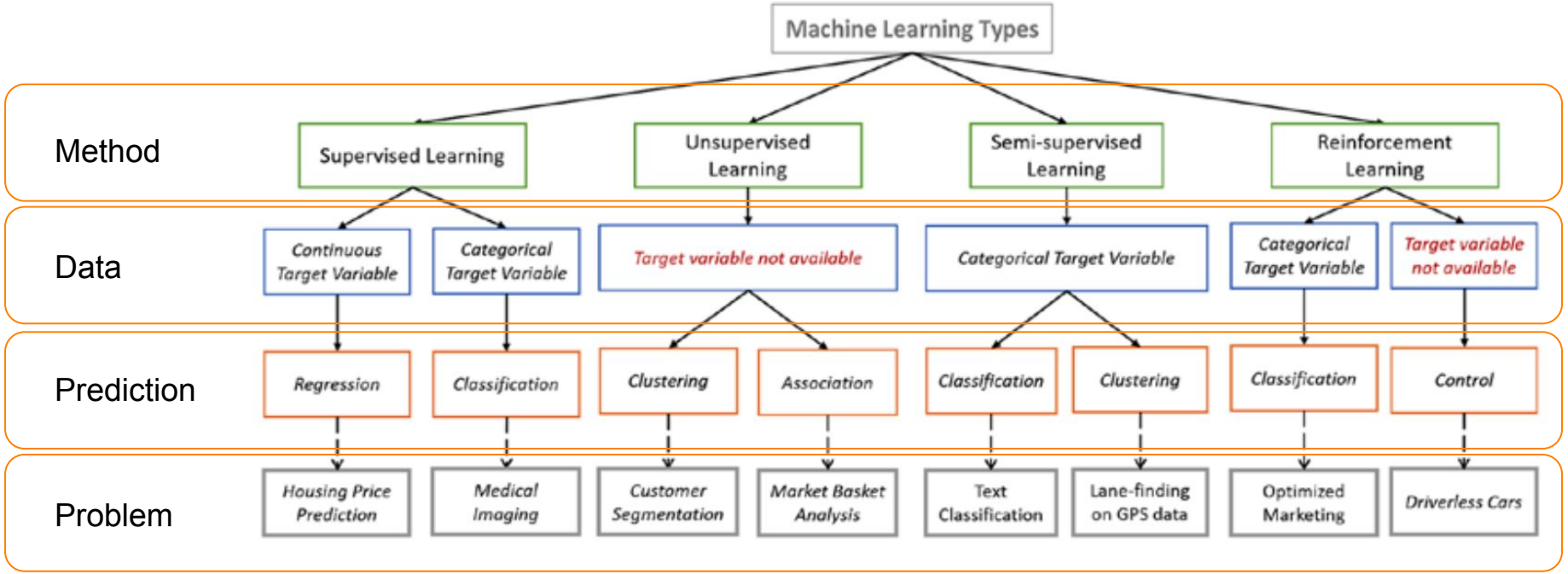
A simple way to think about supervised learning.

INPUT A	RESPONSE B	APPLICATION
Picture	Are there human faces? (0 or 1)	Photo tagging
Loan application	Will they repay the loan? (0 or 1)	Loan approvals
Ad plus user information	Will user click on ad? (0 or 1)	Targeted online ads
Audio clip	Transcript of audio clip	Speech recognition
English sentence	French sentence	Language translation
Sensors from hard disk, plane engine, etc.	Is it about to fail?	Preventive maintenance
Car camera and other sensors	Position of other cars	Self-driving cars

SOURCE ANDREW NG

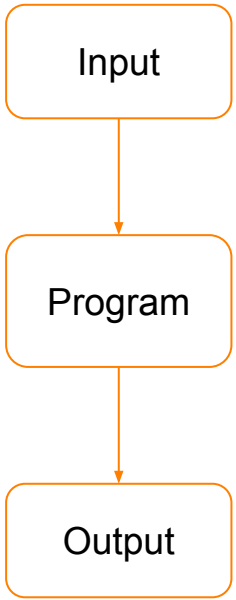
© HBR.ORG

Machine learning predictions take many forms

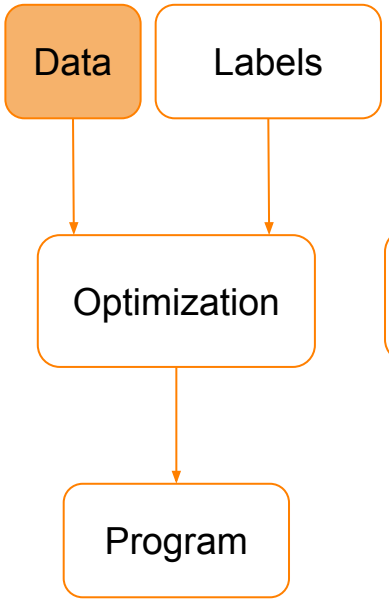


Source: Towards Data Science

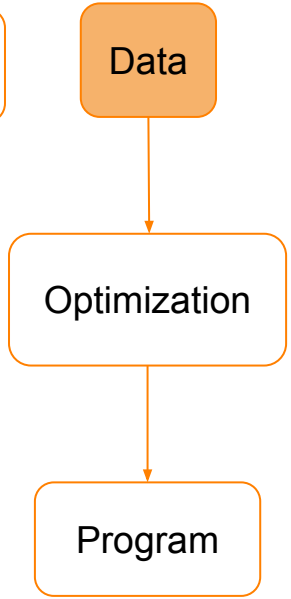
Machine learning is about optimizing from data



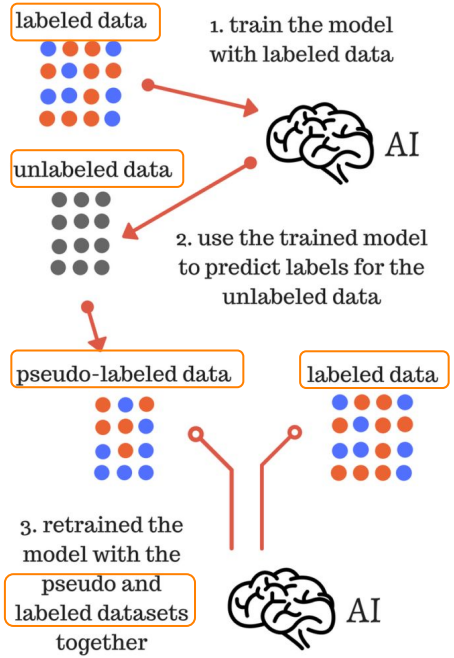
Classical programming



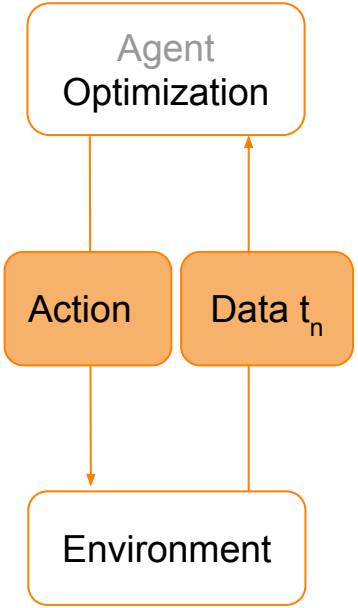
Supervised learning



Unsupervised learning



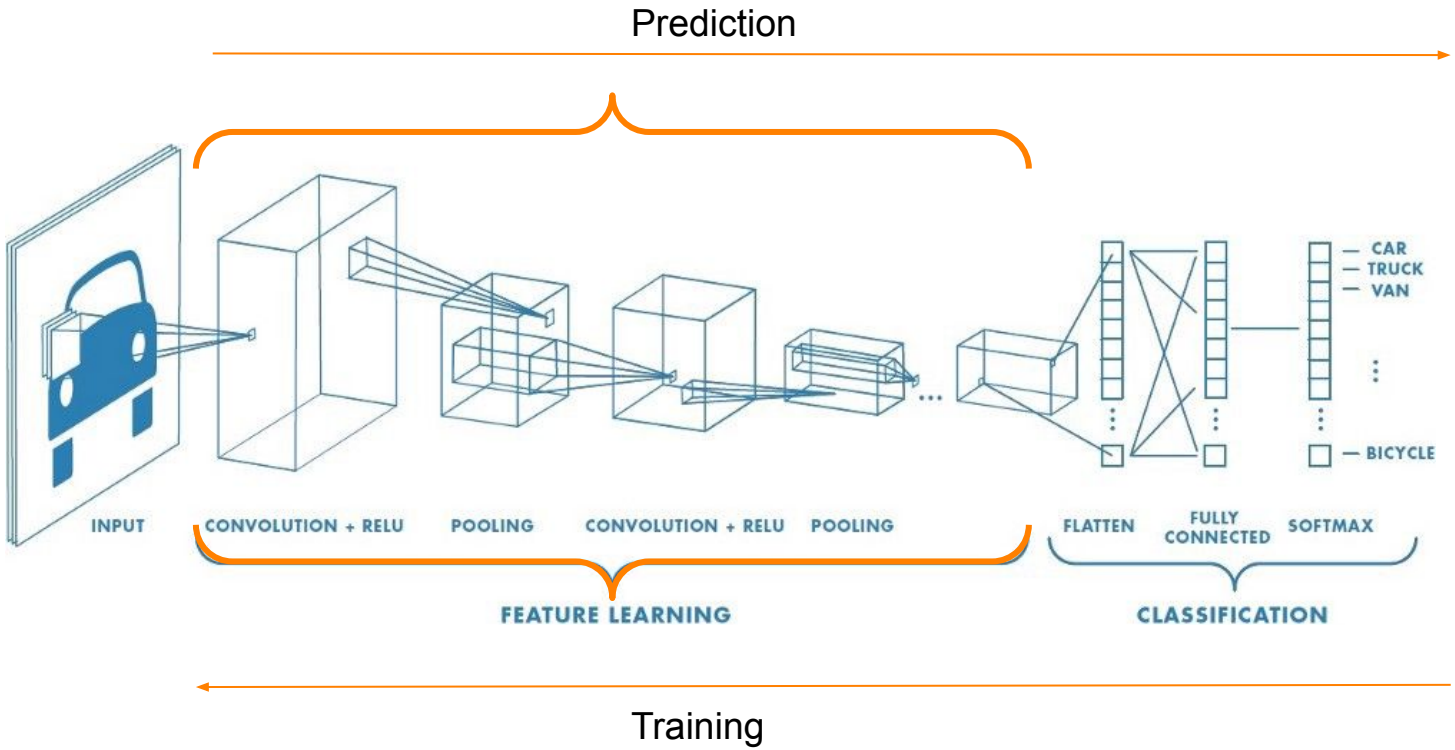
Semi-supervised learning

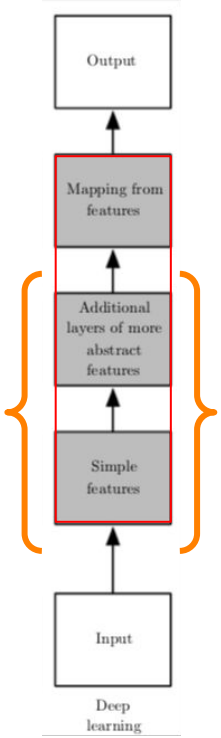


Reinforcement learning

Source: https://datawhatnow.com/pseudo-labeling-semi-supervised-learning/

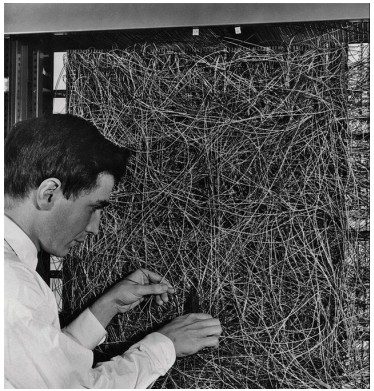
Deep Neural networks: unstructured features



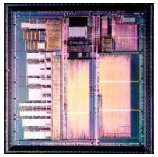


Sources: Mathworks

Machine learning requires machines



Perceptron, 1958



IBM Deep Blue, 1996

IBM Watson, 2011

Apple Siri, 2011

AlphaGo, 2015

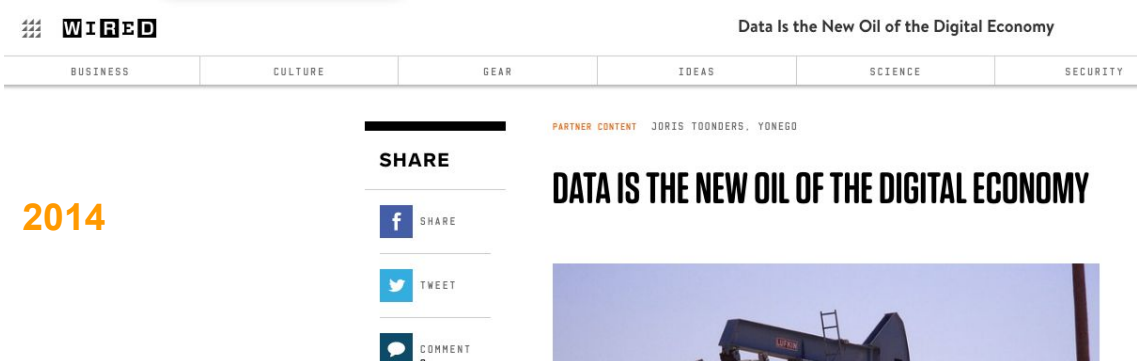
NVIDIA V100
32GB x11

GPT-3, 2020

Machine learning requires data

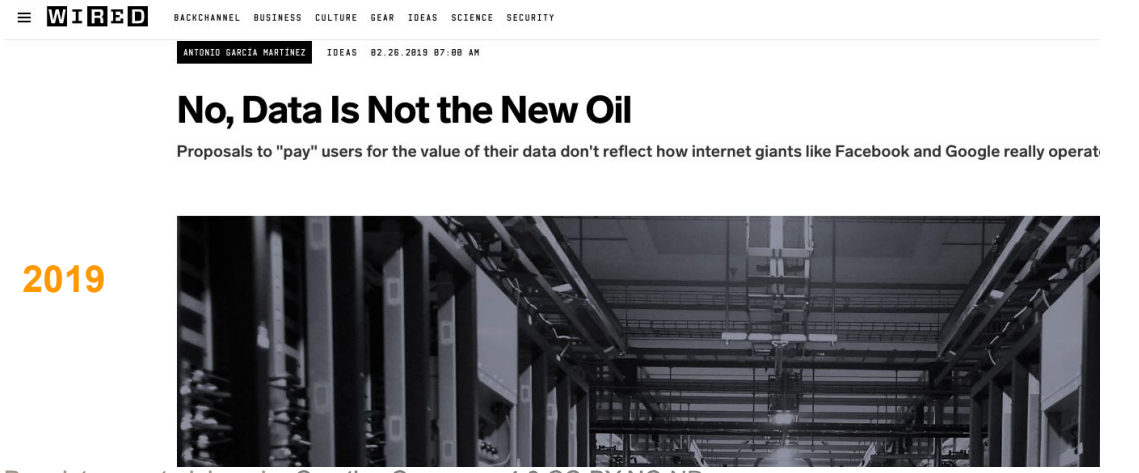
- Machine learning needs data, like industry needs oil.
- Like oil, Data is an extremely valuable asset.

2014

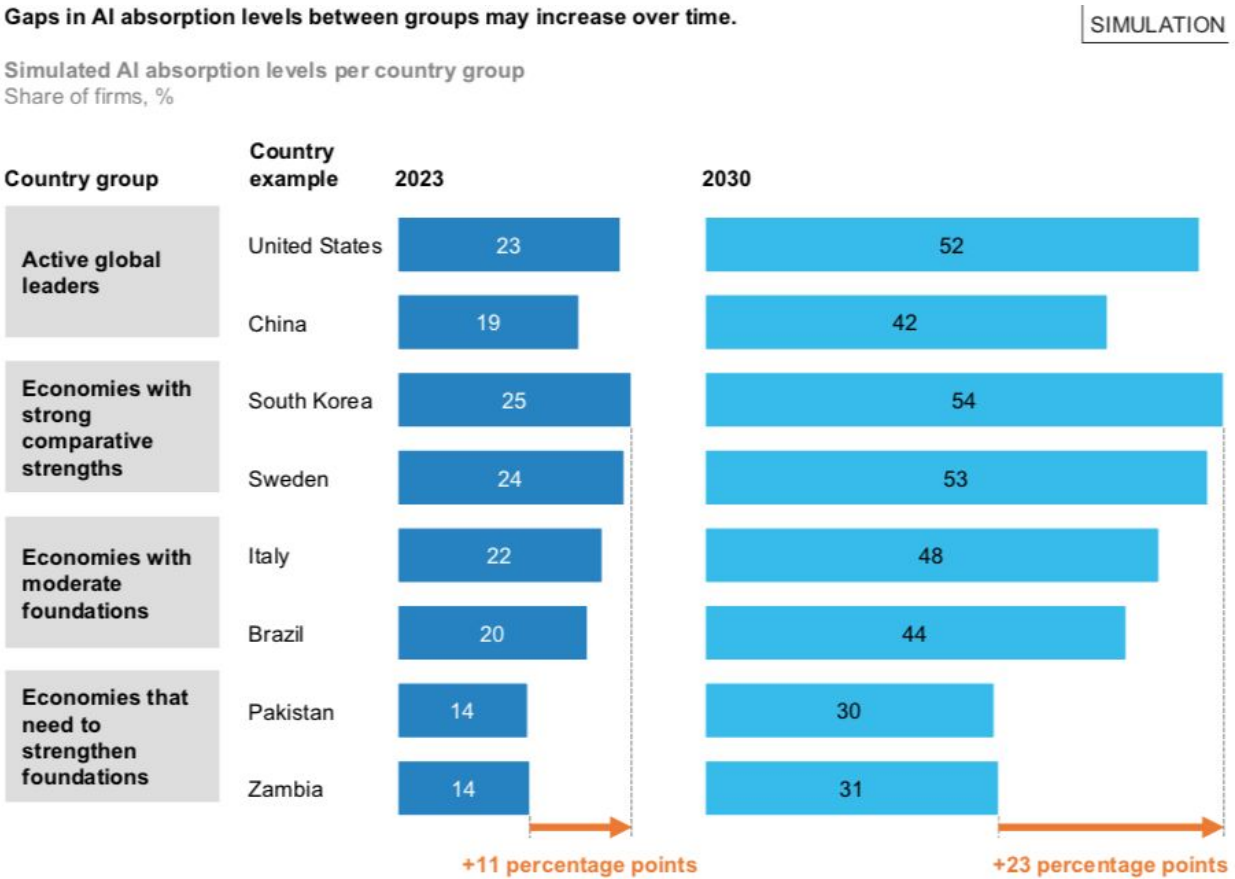


- Oil is paid for. Data is obtained indirectly.
- Oil resources are limited and sparse. Data is not.

2019

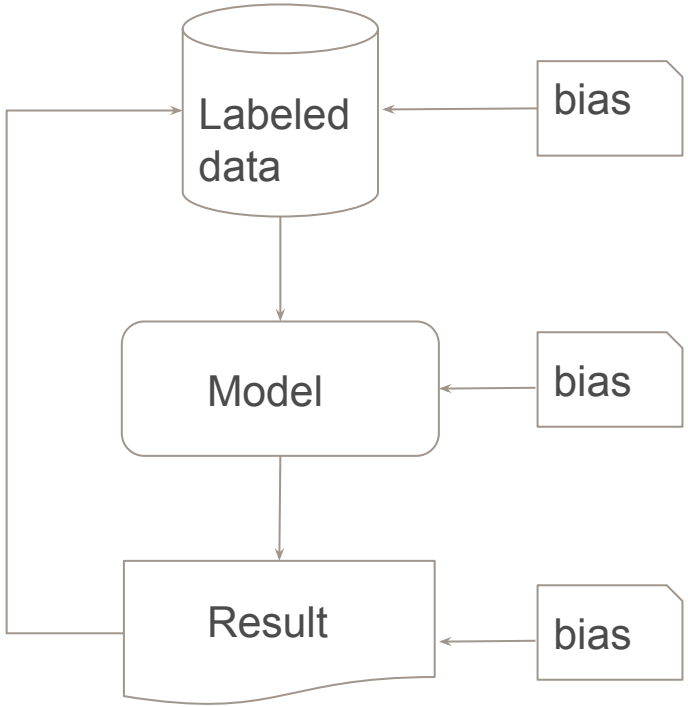


It matters at the strategic level



Source: McKinsey Global Institute Analysis, 2018

It matters at the personal level

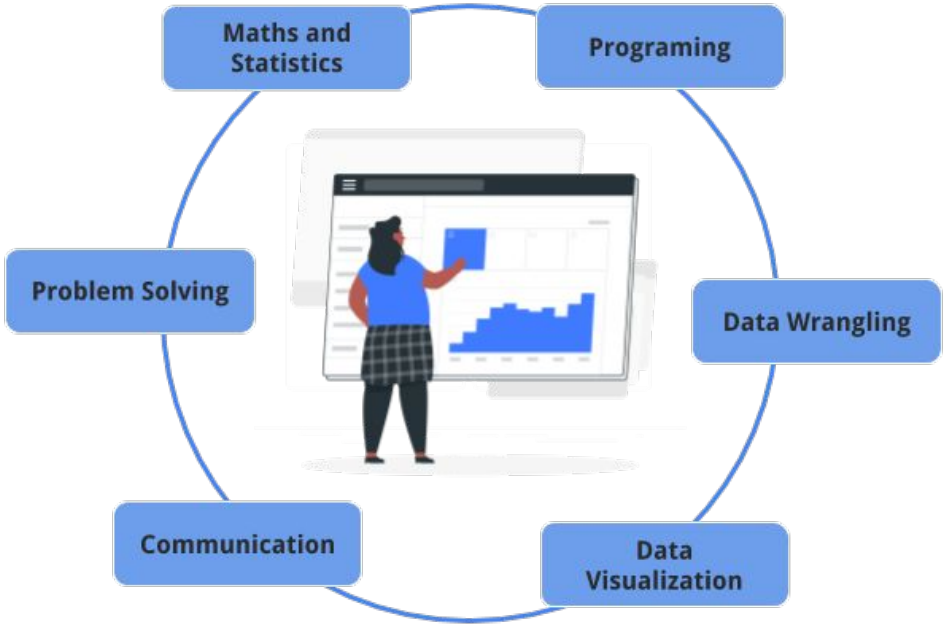


	WHITE	AFRICAN AMERICAN
Labeled Higher Risk, But Didn't Re-Offend	23.5%	44.9%
Labeled Lower Risk, Yet Did Re-Offend	47.7%	28.0%

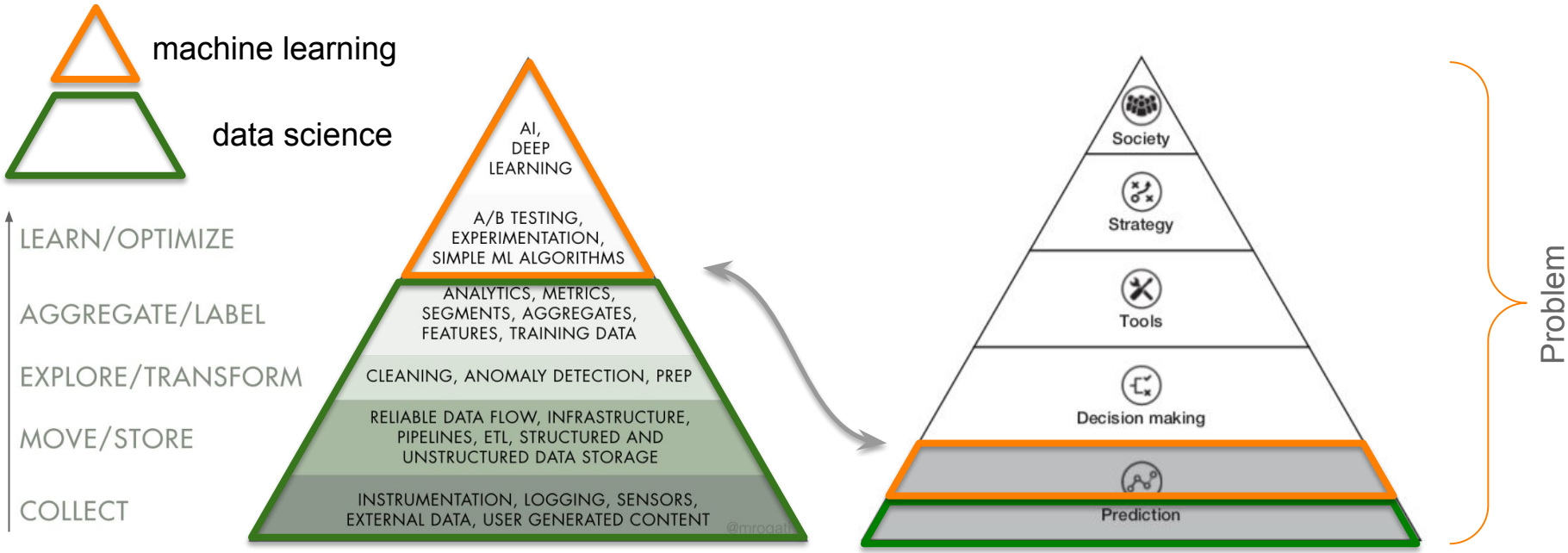
Source: ProPublica

What is needed to be a good data scientist?

- Understand data
- Be curious
- Share your thoughts
- Be responsible



The long road from data to impact

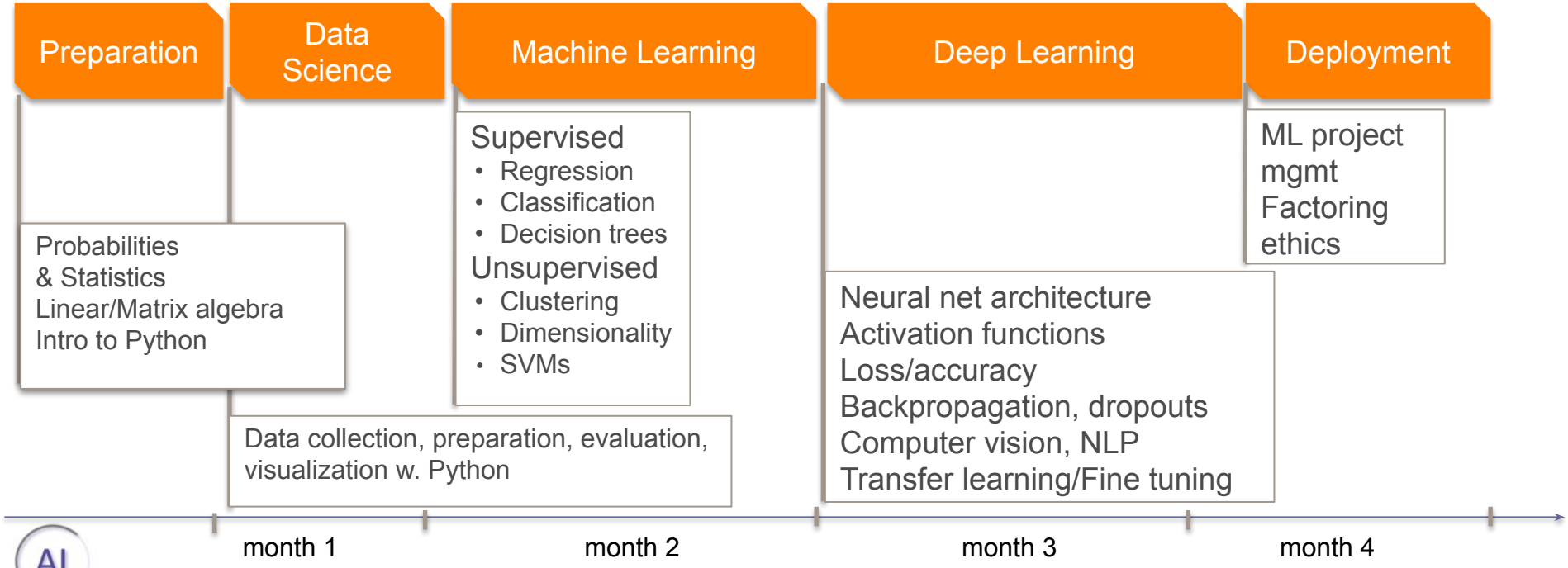


Data science and machine learning prediction are only the beginning of the journey. Problem modeling and impact analysis requires the input of many partners.

Warm-up preparation material

- Probabilities & Statistics
  - Basic principles
    - Seeing Theory
    - Statistics & Probability (Khan Academy)
  - References
    - Seeing Theory book
    - An introduction to statistical learning book
    - Modern statistics for modern biology book
- Linear algebra, matrices & vector calculus
  - Basic principles
    - Pre-calculus
    - Linear algebra
  - References
    - Mathematics for Machine Learning
- Python programming
  - Basic principles
    - Learn Python - Full Course for Beginners [Tutorial]
    - Learn Python
    - Code Academy
  - References
    - Python Data Science Handbook
- Visualization
  - 10 visualizations every data scientist should know
- Bias
  - How I'm fighting bias in algorithms

Curriculum



Calendar of activities (might slightly evolve)

Dates	Course				Labs
week 0	Review & Preparation (supervised by UDG) Probs & Stats., Linear Algebra, Python programming				
Month 1 week 1 to 4	Data collection & preparation				Lab every week
	Python	Numpy	Pandas	Visualization	
Month 2 week 5 to 8	Clustering & k-NNs	Dimensionality	SVMs	Decision trees	
Month 3 week 9 to 12	Regressions	Intro to Neural Networks	Neural Networks II	CNNs	
Month 4 week 12 to 16	Transfer learning	NLP	EDA and AutoML	Building an AI project	
Month 5 week 17 to 20	Challenge (social inclusion)				
	AI ethics	Climate change	Health	Logistics	

Thank You!

Questions?

Thanks to our sponsor HP Guadalajara for the laptops and hardware to make possible this training at UDG.