



UNIVERSIDADE
NOVA
DE LISBOA

Artificial Intelligence and Ethical Issues

UNICA Rectors Seminar, Zagreb, March 24, 2019

José Júlio Alferes

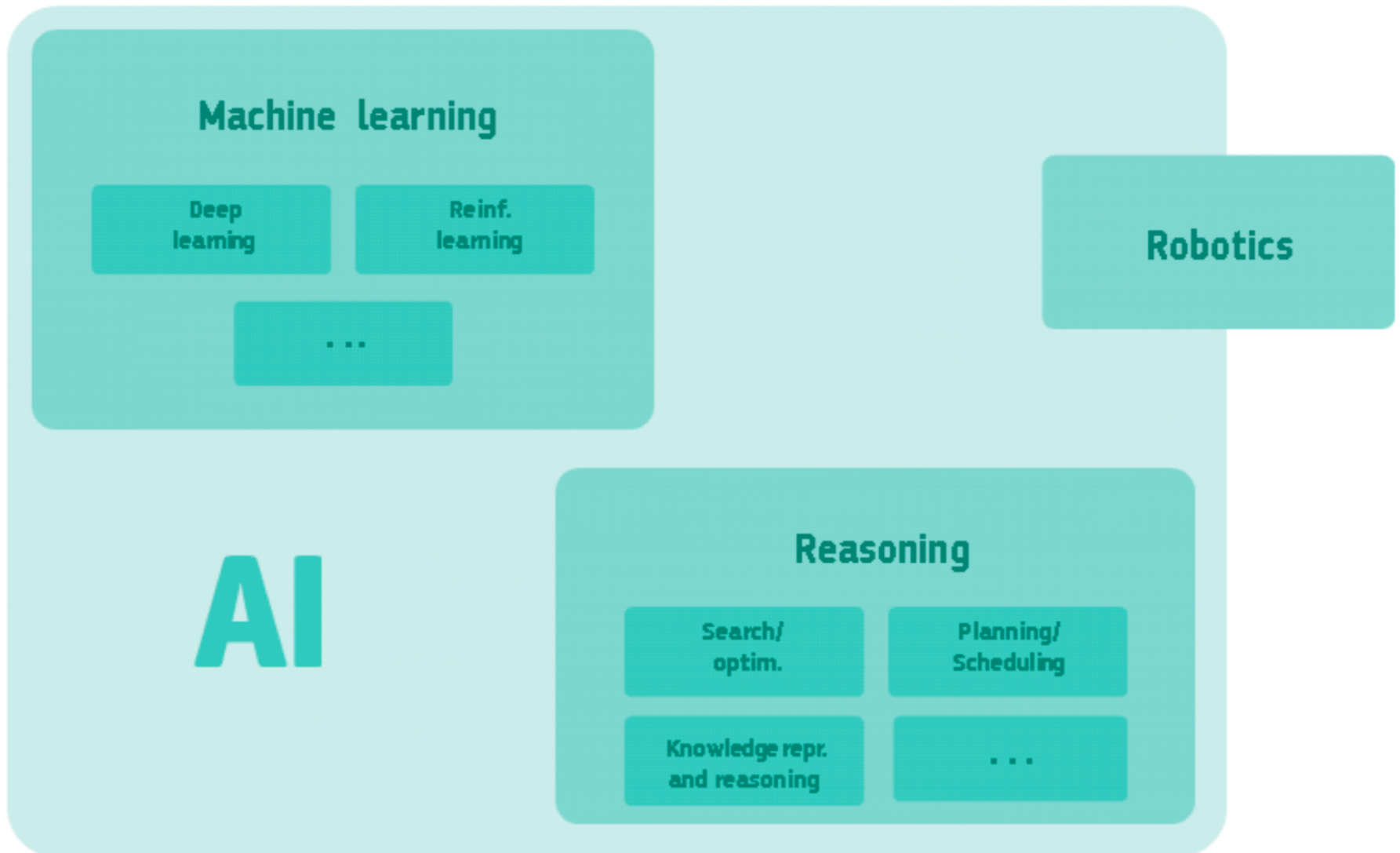
Artificial Intelligence

Aims at building systems that display **intelligent behaviour** and are capable of **learning**.

An AI system is given complex goals and **decides and acts**, possibly **autonomously**, in the physical and/or digital world by perceiving its environment, by learning about it and by reasoning over the learnt knowledge

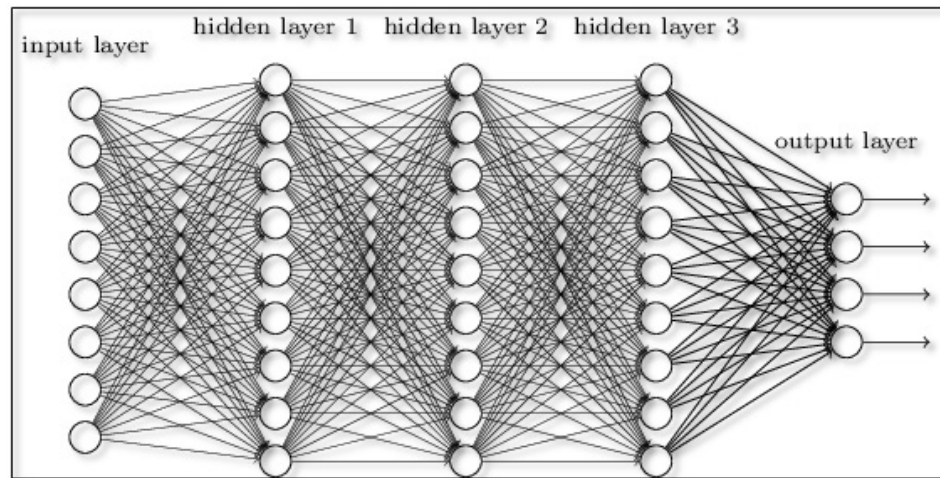
- ✓ It started as a scientific area more than half a century ago, but is gaining momentum just recently, and is having greater and greater impact in the society.

Artificial Intelligence



Technological advances

AI systems learn based on data



- ✓ For that, we need a lot of data
- ✓ **We now have huge amounts of data!**
 - ✓ Web
 - ✓ IoT
 - ✓ Huge databases

Technological advances

AI systems decide and act on complex environments based on reasoning

- ✓ Reasoning requires a lot of processing power
- ✓ **We now have a lot of processing power!**
- ✓ So, the stage is set for huge technological advances in AI systems, with real impact in the society
- ✓ **BUT:**
 - ✓ We must be careful with this impact. Ethical issues ensue
 - ✓ We need a more transparent and trustworthy AI

The data we have

- ✓ 3 white teenagers
- ✓ 3 black teenagers
- ✓ There is a **lot of bias** in existing data!



- ✓ We should learn **how the world ought to be** rather than **how the world is**
 - ✓ But how can we do that? Is there enough good data?



Magnifying Bias

Once an AI starts acting based on learnt knowledge it encodes and magnifies existing bias

- ✓ Correctional Offender Management Profiling for Alternative Sanctions - COMPAS
 - ✓ Advise judges on sentencing and probation decisions
 - ✓ Black people twice as likely to be incorrectly labelled as being at high risk of committing violent repeat offences than white people
- ✓ Predictive policing
 - ✓ If police keeps going more to places that were spotted based on bias, more arrests will be made in those area

On the news (22.05.19)

The screenshot shows the BBC News website interface. At the top, there's a navigation bar with the BBC logo, a 'Sign in' button, and menu items for News, Sport, Weather, Shop, Reel, Travel, and More. A search bar is on the right. Below this is a red 'NEWS' banner with sub-menu items: Home, Video, World, UK, Business, Tech, Science, Stories, Entertainment & Arts, Health, World News TV, and More. The main content area features a 'Technology' sub-section with the headline 'Amazon set for facial recognition revolt' by Leo Kelion, dated 22 May 2019. The article includes social media share icons and a large image showing a woman's face being matched to a database of many other faces. A caption below the image reads: 'Rekognition can match photos to databases holding millions of people's faces'. To the right, there are 'Top Stories' and 'Features' sections. The 'Top Stories' section lists: 'Calls growing for UK PM May to resign' (3 minutes ago), 'Trump hits back at Democratic claims of 'cover-up'' (21 minutes ago), and 'Huawei faces break with chip giant ARM' (7 hours ago). The 'Features' section has a video thumbnail for 'India's six-week election in two minutes' with the text 'Have you Voted?'.

BBC Sign in News Sport Weather Shop Reel Travel More Search

NEWS

Home Video World UK Business Tech Science Stories Entertainment & Arts Health World News TV More

Technology

Amazon set for facial recognition revolt

By Leo Kelion
Technology desk editor

22 May 2019

f Messenger Twitter Email Share

Rekognition can match photos to databases holding millions of people's faces

Top Stories

Calls growing for UK PM May to resign
3 minutes ago

Trump hits back at Democratic claims of 'cover-up'
21 minutes ago

Huawei faces break with chip giant ARM
7 hours ago

Features

India's six-week election in two minutes

Have you Voted?

Transparent AI

Trust is essential, and can only be gained by transparency!

- ✓ AI systems should provide explanations for decisions
 - ✓ Why have you decided that I'm guilty?
 - ✓ Why should I take this drug/treatment?
 - ✓ Why has the car decided to turn (in case of accident)?
- ✓ State of the art AI is far from this
 - ✓ But "Explainable AI" is a hot area of research
- ✓ But what if it decreases accuracy?
 - ✓ Would you prefer an "AI doctor" 99% accurate explaining nothing, or one that is 70% accurate and explains?

What about autonomy?

Should we always keep humans in the loop?

- ✓ Sometimes, most likely yes
 - ✓ Doctors, Judges, ...
 - ✓ But what if it decreases accuracy?
- ✓ Sometimes it is highly debatable (Warfare)
- ✓ Sometimes, definitely no!
 - ✓ Driverless cars
- ✓ If the AI decides by itself, it must function **responsibly, ethically, and securely**

An ethics for machines

Autonomous machines co-exist ethically among themselves, and with humans. They should be provided with an ethics of its own

- ✓ Do we know how to program ethics and moral rules? Do we know how to express them?
- ✓ Upon a given situation
 - ✓ Car A chooses to destroy itself (and kill the driver), to save a group of 10
 - ✓ Car B chooses to save itself (and the driver), possibly killing a group of 10
- ✓ Which one would you support?
- ✓ Which one would you buy for yourself?



Autonomous weapons

Keep humans in the loop?



is at following rules of war
it

- ✓ Arguments against killer robots
 - ✓ It will spark an arms race
 - ✓ There is no way to prevent terrorists
 - ✓ It was possible to deter in chemical and bio weapons
 - ✓ Better safe than sorrow!

Socio-economic impact

AI is changing the economy and the society!

- ✓ It raises intricate questions about unemployment and work (and wealth) distribution
- ✓ Leaving it just to the market:
 - ✓ May accumulate a lot of wealth for big AI developers
 - ✓ Increase inequality (a society of castes)
- ✓ Danger of
 - ✓ Devaluating human skills
 - ✓ Devaluating human individuality
 - ✓ Removing human responsibility
 - ✓ Reducing human control and privacy

A Humane trustworthy AI

A Humane, trustworthy, transparent AI, interacting with humans augments our capacity.

It has the power to change the world for good!

- ✓ Such an AI is much more difficult to develop! We need:
 - ✓ More **research in AI**: explainability, verification, security...
 - ✓ Ethical guidelines and **machine ethics**
 - ✓ **Multidisciplinary work**: philosophy, law, psychology, anthropology, evolutionary biology, economy, AI, Computer Science, ...
 - ✓ Legislation
 - ✓ Policy
 - ✓ **Education**



Thank You

The machine is only a tool after all, which can help humanity progress faster by taking some of the burdens of calculations and interpretations off its back. The task of the human brain remains what it has always been; that of discovering new data to be analyzed, and of devising new concepts to be tested.

I, Robot, Isaac Azimov