This E-Book has been developed to support your learning process during the Training of TOs. It provides more in-depth information on the content of each Module. It also includes some useful resources for TOs on each Topic.

Module 1: What is AI?

The first module introduces to the basic concepts of Artificial Intelligence, starting from defining Artificial Intelligence and discussing the different domains related to AI (computer science, statistics, philosophy, etc.). The module explores the basic concepts related to artificial intelligence such as intelligence and imageries of AIs, addressing false myths and unrealistic expectations around AI technology through an honest and critical discussion, spanning different disciplines. Through this training module, teachers will acquire competencies to guide class debates and activities focused on expectations and beliefs about AI, and to start a critical reflection on challenges and opportunities of the new technologies

Topic 1 .1 - Defining Artificial Intelligence

According to the high-level expert group on Artificial Intelligence[1] set up by the European Commission "*Artificial intelligence (AI) refers to systems that display intelligent behaviour by analysing their environment and taking actions – with some degree of autonomy – to achieve specific goals*". It is important to note that AI-based systems can be purely software-based, acting in the virtual world (e.g. voice assistants, image analysis software, search engines, speech and face recognition systems) or AI can be embedded in hardware devices (e.g. advanced robots, autonomous cars, drones or Internet of Things applications). Artificial intelligence (AI) systems are software or hardware systems designed by humans that, given a complex goal, act in the physical or digital dimension by perceiving their environment through data acquisition, interpreting the collected structured or unstructured data, reasoning on the knowledge, or processing the information, derived from this data and deciding the best action(s) to take to achieve the given goal (for a more detailed explanation see the EU definition).

The term AI contains a direct reference to the notion of intelligence that is quite hard to define: thus it is important to explore this concept in class and create a common ground for defining what makes a machine "intelligent".

---

1 https://digital-strategy.ec.europa.eu/en/library/definition-artificial-intelligence-main-capabilities-and-scientific-disciplines

Generally speaking, Artificial Intelligence is the name given to any computer system taught to mimic intelligent human behaviours. However, the concept of intelligence is difficult to define and it is hard to differentiate between "types" of intelligences, such as human, animal and machine intelligence. There are indeed certain tasks that are very easy to do for humans (such as learn to walk or to speak) and that are very difficult to replicate in machines. On the other side, there are tasks that are very easy for machines and difficult for humans, such as activities that are repetitive or require a lot of calculation and data elaboration, such as playing chess or translating a text.

Given the renewed general interest in the field of AI, there is a lot of excitement but also a lot of unnecessary hype about AI. One of the reasons for this is because the concept of AI is actually related to two separate ideas: Artificial Narrow Intelligence and Artificial General Intelligence. Almost all the progress we are seeing in the field of AI today are related to Artificial Narrow Intelligence. These are applications that perform one single task such as an algorithm that can recognise a human face in a picture, a smart speaker that can recognise the human voice or a self-driving car that can drive in the streets. Unfortunately, AI also refers to a second concept of Artificial General Intelligence, that is the goal to build AI that can do anything a human can do or maybe even be super-intelligence and do even more things than any human can.

It is important to consider that we are witnessing great progress in artificial narrow intelligence and almost no progress in artificial general intelligence. Unfortunately the rapid progress in artificial narrow intelligence, even though it is incredibly valuable, has caused people to conclude that there is a lot of progress in AI, which is true. But that has caused people to falsely think that there might be a lot of progress in artificial general intelligence as well, which is leading to some irrational fears and expectations around AI.

Useful resources

European Commission High-level expert group on artificial intelligence. A definition of Artificial Intelligence: main capabilities and scientific disciplines. https://digital-strategy.ec.europa.eu/en/library/definition-artificial-intelligence-main-capabilities-and-scientific-disciplines (English only)

Element of AI. A series of free online courses created by Reaktor and the University of Helsinki and it is available in multiple European languages (including Bulgarian/Greek/Italian/Romanian). The site helps teach what AI is, what AI can and cannot do, and how to start creating AI methods. The courses combine theory with practical exercises and can be completed at one's own pace. https://www.elementsofai.com/ (Multiple languages)

A-Z guide (by Oxford Internet Institute and Google) offers a series of simple, bite-sized explainers to help anyone understand what AI is, how it works and how it's changing the world around us. https://atozofai.withgoogle.com/intl/en-GB/ (English only)

An article discussing the **different types of "intelligence" between humans and machines**. https://time.com/4960778/computers-smarter-than-humans/ (English only)

The article presents **60-years developmental history of AI** and analyzes the external environment promoting the formation of AI 2.0
Pan, Y. (2016). Heading toward artificial intelligence 2.0. Engineering, 2(4), 409-413. https://reader.elsevier.com/reader/sd/pii/S2095809917300772?token=86492C08D8D3F92DE82182916ABD8E2D63DDB1659D7C0BE2D22EE666229171BB52FC442CE57403DCB4DA557B19D876B0&originRegion=eu-west-1&originCreation=20210928124530

Topic 1.2 - Interdisciplinarity of AI

The use of artificial intelligence (AI) in a variety of research fields is drastically accelerating multiple digital revolutions; indeed many paradigms are shifting in healthcare, precision medicine and wearable sensing, but also in public services and education offered to the masses around the world, to future cities made optimally efficient by autonomous driving and self-sustainability. These systemic revolutions are bound to deeply change the future configuration of our societies and the consequences of the foreseeable systemic changes are not obvious straight away. To date, there is neither consensus nor uniformly adapted framework to guide AI research to ensure a sustainable societal transition. A good example of these debates within AI studies, is represented by the ethical implication of AI powered robotics. For instance, the general area of research of robotics includes research and development in mechanical engineering to build the hardware components of the robots, computer science and cognitive science to create the AI that can control the different parts of the robot, applying research to problems involving computer vision (how the robot can perceive the environment), language processing (how the robot can talk and understand words) and reasoning (how the robot can plan actions and reacts to the environment). However, it is also important to understand how the robots can positively interact with other humans (Ried 2019), so it is important to consider research on psychology and design, and evaluating the economic and social consequences of using robots for automating certain tasks and how this might impact certain jobs. For instance, as AI makes prediction abundant and inexpensive, it is important to explore how AI can help us solve problems that were not previously prediction oriented, or how the value of some human skills will rise while others fall (Agrawal et al. 2017). Moreover, recent debates have highlighted the significant environmental

impact of AI (Griffin 2020), i.e. the carbon footprint resulting from the use of AI technologies. Another pressing and interdisciplinary aspect of AI related to development is that of the exploitation of workers in the Global South, including refugees, who perform “clickwork” as part of machine learning processes that serve big tech companies such as Google, Amazon, Facebook and Tesla (Jones 2021). Indeed, as we will further explore in Module 3, the implications of the use of AI technologies on different areas of human and, more generally, on sustainable development are complex and require collaboration among researchers, developers, businesses, civil society actors and governmental agencies.

AI applications do not fluctuate in a vacuum but lie at the intersection of great societal problems. How to strike a balance between technological enhancement and ethical concerns? How can AI assist with interdisciplinary research for the public good? Interdisciplinary research in the field of AI poses large cognitive challenges for human researchers (Baum 2020) : the challenges include but are not limited to epistemic divides between disciplines.

Useful resources
Will robots take my job? This is a website that lets you find out whether a job is in danger of extinction due to machines and digital transformation.
https://willrobotstakemyjob.com/ (English only)

Some scientific papers on Interdisciplinarity of AI:

Baum, S. D. (2020). **Artificial Interdisciplinarity: Artificial Intelligence for Research on Complex Societal Problems**. Philosophy & Technology, 1-19.
http://gcrinstitute.org/papers/052_ai-interdisciplinarity.pdf

Kusters, Remy, Dusan Misevic, Hugues Berry, Antoine Cully, Yann Le Cunff, Loic Dandoy, Natalia Díaz-Rodríguez et al. "Interdisciplinary Research in Artificial Intelligence: Challenges and Opportunities." *Frontiers in Big Data* 3 (2020): 45.

Definition of the stages in the design process
https://www.interaction-design.org/literature/article/5-stages-in-the-design-thinking-process

Humans are increasingly coming into contact with AI and ML systems. Given the potential for intelligent systems to impact people's lives, it is important to design intelligent systems with this in mind. The paper lays out arguments that AI and ML systems that are human-centered (a) have an understanding of human sociocultural norms as part of a theory of mind about people, and (b) are capable of producing explanations that non expert end-users can understand.It reaches the conclusion that the pursuit of human-centered AI presents a research agenda that will improve our scientific understanding of fundamen-tal AI and ML while simultaneously supporting the deployment of intelligent products and services that will interact with people in everyday

contexts.
Riedl, M. O. (2019). Human-centered artificial intelligence and machine learning. Human Behavior and Emerging Technologies, 1(1), 33-36.
https://onlinelibrary.wiley.com/doi/pdfdirect/10.1002/hbe2.117

Learning through human feedback: DeepMind co-founded initiatives like the Partnership on AI to Benefit People and Society and why we have a team dedicated to technical AI Safety. Research in this field needs to be open and collaborative to ensure that best practices are adopted as widely as possible, which is why we are also collaborating with OpenAI on research in technical AI Safety. https://deepmind.com/blog/article/learning-through-human-feedback

The article maintains that to understand how advances in artificial intelligence are likely to change the workplace — and the work of managers — in coming years, you need to know where AI delivers the most value.
Agrawal, A., Gans, J., & Goldfarb, A. (2017). What to expect from artificial intelligence.
https://static1.squarespace.com/static/578cf5ace58c62ac649ec9ce/t/589a5c99440243b575aaedaa/1486511270947/What+to+Expect+From+Artificial+Intelligence.pdf

Baum, S.D., "Artificial Interdisciplinarity: Artificial Intelligence for Research on Complex Societal Problems", Philosophy and Technology, 2020

Sahota, Neil. "Beyond AI: The Need For Solutionists In Creating A Sustainable World". Forbes. 24 April 2019.
https://www.forbes.com/sites/cognitiveworld/2019/04/24/beyond-a-i-the-need-for-solutionists-in-creating-a-sustainable-world/

BBC The Documentary Podcast (2019). "Will AI kill development? Will robotisation prevent poorer countries taking the traditional route to prosperity?"
https://www.bbc.co.uk/programmes/p075nqyy

Jones, Phil. "Refugees help power machine learning advances at Microsoft, Facebook, and Amazon". Restofworld. 22 September 2021.
https://restofworld.org/2021/refugees-machine-learning-big-tech/

Griffin, Thomas. "Why We Should Care About The Environmental Impact Of AI". Forbes. 17 August 2020.
https://www.forbes.com/sites/forbestechcouncil/2020/08/17/why-we-should-care-about-the-environmental-impact-of-ai/

Topic 1.3 - Recognizing AI

This topic focuses on the recognition of AI technologies in daily life, e.g. when shopping online or browsing through social media. Being aware of AI around them, students will be able to visualize the more technical explanations of the second module in these identified instances of AI in their lives.

From Alexa and Sort to Google Search and Facebook's Newsfeed, we interact with versions of "artificial intelligence" constantly, but few of us really recognise them as AI. We tend indeed to think that AI is something that requires a bodily representation, such as robots or a virtual avatar, but, in reality, AI is widely deployed in many different tools that we use everyday. For example, Siri, Alexa and Google Now are voice-activated assistants that use AI to recognise the human voice, elaborate the meaning of the information they collect and provide an answer to that. Amazon, eBay and many other commerce services use AI to personalise the recommendations provided to their customers. Google Image and many photo applications use AI to extract meaning from images, recognising for example if you took a picture of a dog or a cat, or if you just took a selfie of yourself.

Useful resources

Humans vs. Machines: Seven Myths (and Realities)
https://glocalthinking.com/en/humans-vs-machines-seven-myths-and-realities/

What is AI? An executive guide to artificial intelligence, from machine learning and general AI to neural networks.
https://www.zdnet.com/article/what-is-ai-heres-everything-you-need-to-know-about-artificial-intelligence/

An article helping to investigate **what is AI using a flowchart**
https://www.technologyreview.com/2018/11/10/139137/is-this-ai-we-drew-you-a-flowchart-to-work-it-out/