Concepts, definition, and history of AI and interaction with AI

Module 1

The History of AI

Governments funded the building of expensive computers at a few universities after seeing their potential due to its role in code breaking during World War II. In 1949, Alan Turing, a leading code breaker created a sensation when he indicated how these computers may be used. He wrote in the London Times that he thought that the computer could eventually simulate human intellect and be able to compete on equal terms. (Grudin, 2009).

The term "artificial intelligence" was used for the first time in 1956 by mathematician and logician John McCarthy when he called for participation to his workshop (Grudin, 2009). The workshop was called the Dartmouth Summer Research Project on Artificial Intelligence and it was held to discuss what would ultimately become the field of AI and to clarify and develop the concepts around "thinking machines" (Marr, 2018).

Defining Artificial Intelligence

John McCarthy, who first coined the term Artificial Intelligence shared his definition of AI on a "layman's level" in an article posted at Stanford in 2004. This was his answer to the question: "What is artificial intelligence?":

> *"It is the science and engineering of making intelligent machines, especially intelligent computer programs. It is related to the similar task of using computers to understand human intelligence, but AI does not have to confine itself to methods that are biologically observable."* (McCarthy, 2004)

The Encyclopedia Britannica defines artificial intelligence as such:

> *"Artificial intelligence (AI) is the ability of a computer or a robot controlled by a computer to do tasks that are usually done by humans because they require human intelligence and discernment. Although there are no AIs that can perform the wide variety of tasks an ordinary human can do, some AIs can match humans in specific tasks"* (Britannica, n.d)

The global technology company IBM defines artificial intelligence as such:

> *"Artificial intelligence leverages computers and machines to mimic the problem-solving and decision-making capabilities of the human mind."* (IBM, 2020)

IBM's and Britannica's definitions connects artificial intelligence to human intelligence – specifically how AI mimics tasks that are originally performed by humans. McCarthy's definition describes AI as making intelligent machines and computer programs using science and engineering, and he later defines intelligence as the computational part of the ability to achieve goals in the world. If I were to make my own definition based on these three definitions, I would define it like this:

> *"Artificial intelligence is the use of science and engineering to create systems that imitate human intelligence."*

A brief review: Does AI make PD obsolete? Exploring challenges from Artificial Intelligence to Participatory Design.

An article written by Bratteteig and Verne (2018) explores how Artificial Intelligence potentially challenges Participatory Design. PD aims to understand technology and to open for future users to have a say in choices concerning the technology during its design and use. It can be challenging to employ PD techniques when developing AI – for example, it is difficult to create concrete AI prototypes because they are time consuming to build and rely on an adequate data set (preferably accumulated over a longer period) for it to be accurate. An algorithm will also change over time as is adapts to new data input which makes it unpredictable to design for.

The correct answer needs to be provided for the computer algorithms to improve its recognition, and the algorithm could be developed in a bad direction if someone intentionally feeds it wrong, sensitive, or possibly offensive information. The designer has a certain responsibility when designing for AI systems, for example to provide users with a choice when interacting with autonomous AI to avoid being them overruled by a system. A critique of autonomous AI is that it potentially poses a threat to the development of the human experience, sensation, and decision-making abilities because we need to make less decisions

for ourselves – the systems do it for us. Using PD techniques could help to avoid that this would happen.

Apple's Siri

Apple has an integrated voice-controlled virtual assistant called Siri in their products which gives their users the opportunity to utilize their devices' features using only their voice. On their website, they describe Siri as a service who can help you with just about anything. "Siri does more than ever. Even before you ask." They explain how Siri uses machine learning to continuously learn and become able to personalize your experience – but they also reassure you that your personal data is kept private and off limits to advertisers or other organizations who might benefit from it. Apple is transparent with the users on how their personal data is used to develop their service – and they even encourage them to interact with Siri in new ways to make her smarter. (Apple, n.d).

Do you trust this computer? – Human interaction with AI

"Do you trust this computer" is a documentary from 2018 which in short talks about the pros and cons with the use of artificial intelligence – or as they articulate it in the film, the "miracles and horrors" of it. Artificial intelligence is not good or evil, but it can be utilized for both depending on who is writing the code. AI can be utilized in healthcare to make surgery easier and help save more lives, it can be used in robots that can help clean your house or guide a blind man trough the street – and it helps millions of people every day to locate information through the search-engine Google.

As of now though, it is a multimillion-dollar industry with almost no regulations, which opens opportunities for it to be used for less noble actions. In the film, machines are described as natural psychopaths. They have no emotions – and if they are given a task that could take human lives they will complete it with no remorse. This is especially concerning regarding the development of autonomous weapons, which is part of the reason why people want a to see a global ban of using this technology in warfare. They mention concerns around the possibility of artificial intelligence gaining more control than us humans – because the technology is accelerating quicker than us, and nobody fully understands the advanced algorithms used in many AI systems. (Paine, 2018) After watching the documentary I feel that we could exploit

AI and ML-techniques to advance humanity and achieve great things – but at the same time I am left with the feeling that it could affect us horribly if the technology is not regulated properly.

Robots and AI systems

The origin of the word Robot.

The word robot was first introduced by Czech writer and journalist Karel Čapek in his play "R.U.R" (Short for Rossum's Universal Robots) in 1921. It stems from the word *Robota* in Czech and can be translated to work of duty. The play illustrates robots as scary beings – as they eventually run amok and exterminate mankind. (Store Norske Leksikon, n.d)

Defining "Robot"

The Encyclopedia Britannica defines a robot as such:

> *"Robot, any automatically operated machine that replaces human effort, though it may not resemble human beings in appearance or perform functions in a humanlike manner."* (Britannica, n.d)

The dictionary of Cambridge defines a robot as such:

> *"[a robot] is a machine controlled by a computer that is used to perform jobs automatically."* (Cambridge Dictionary, n.d)

If I were to make my own definition based on these two, I would define it like this:

> *"A robot is an automatically operated machine that is used to perform labor, and it may or may not have features that resemble human behavior and/or appearance."*

AI vs. Robots

Artificial intelligence and robots are closely related technologies. I think that they are similar in the way that both can and often aim to imitate human behavior to perform a task. They might differ because I perceive a robot as more of a physical object that does the job, rather

than the possible non-physical artificial intelligence that makes it behave in the way that it does. I perceive AI as a feature that a robot can have, but not the other way around.

The robot lawn mower

Having a robot lawn mower has become an increasingly popular item in the ordinary household. Professor Guri Verne (2020) has written an article on how she adapted to hers to make it work in the manner it was supposed to. The robot is supposed to replace the human physical labor of mowing the lawn – but Verne explains how she was surprised over how much labor she still had to do for the robot to cut the grass in her garden. For example, the garden had to always be clean when the robot started, which meant there could not be a hose or a shovel laying around. The article entails that there is minimal human interaction with the lawn mower robot itself, but the garden must be prepped for it to work correctly. It starts automatically, then proceeds to do its job by cruising around the garden and cutting the grass, while potentially destroying things or getting stuck along the way depending on how well the garden is prepped.

Universal design and AI systems

Defining Universal Design

The Centre for Excellence in Universal Design posted a definition of Universal design that originates from The Disability Act 2005. They define Universal design as such:

1. *The design and composition of an environment so that it may be accessed, understood and used*
    - i. *To the greatest possible extent*
    - ii. *In the most independent and natural manner possible*
    - iii. *In the widest possible range of situations*
    - iv. *Without the need for adaptation, modification, assistive devices or specialised solutions, by any persons of any age or size or having any particular physical, sensory, mental health or intellectual ability or disability, and*
2. *Means, in relation to electronic systems, any electronics-based process of creating products, services or systems so that they may be used by any person.*

(The Centre of Excellence in Universal Design, n.d)

Universal design in electronic systems aims to give all members of society an equal opportunity to utilize these systems – regardless of whether they have the same physical or mental abilities.

The potential of AI with respect to human perception, human movement, and human cognition/emotions.

AI has great potential in extending and accelerating human capabilities. It is possible for amputees to get a robotic prosthesis to replace their missing arm or leg – which then makes them more capable of participating in society like "normal" because they can now walk up the stairs or wave to people on the street. Pet-robots have also been created to accompany elderly who suffer from dementia. They are furry robots who look like real animals, and they provide emotional support.

The potential of AI for both including and excluding people

AI can only learn from data is receives and interacts with – which makes it prone to appear biased. For example, if a filter on Snapchat is only tested with white people, the AI is trained to better recognize their features which can exclude black people from being able to interact with the filter. AI can also be very inclusive, as there are programs that makes blind people able to interact with a smartphone.

Do machines understand?

I think "to understand" entails that you have learned something, and that you are able to interpret that knowledge and use it in another context. I do not think that a machine can understand, because I think it only acts according to the direct data it is given and has learned from. If the machine was put in a new environment with no new data and was expected to figure out how to succeed, I assume it would fail.

Guideline for Human-AI interaction

Microsoft Guideline 12

Microsoft's 12th guideline for Human-AI Interaction Design reads "Remember recent interactions. Maintain short-term memory and allow the user to make efficient references to that memory". An example where this rule is practiced is in any web-browser. The browser remembers your search history and will recommend a page you have previously interacted with when you start typing in the search bar.

Similarities and differences: HCI design guidelines and the Human-AI interaction guidelines.

Nielsen and Molich's 10 user interface design guidelines and Microsoft's guidelines for Human-AI interaction design share several similarities and some differences. For example, both highlight the fact that you should be able to understand the systems' intended use immediately, and that the system shows you task-relevant information – often based on your previous choices. The user interface design guidelines differ because they also mention the preferred aesthetic of the system – which is to keep clutter at a minimum.

References

Apple. (n.d) *Siri.* Available at: https://www.apple.com/siri/ (Accessed: 7.9.21)

Bratteteig, Tone. Verne, Guri. (2018). *Does AI make PD obsolete? Exploring challenges from Artificial Intelligence to Participatory Design*.

Britannica. (n.d) *Artificial Intelligence.* Available at: https://www.britannica.com/technology/artificial-intelligence (Accessed: 7.9.21)

Britannica. (n.d) *Robot.* Available at: https://www.britannica.com/technology/robot-technology (Accessed: 7.9.21)

Cambridge Dictionary. (n.d) *Robot.* Available at: https://dictionary.cambridge.org/dictionary/english/robot (Accessed: 7.9.21)

The Centre for Excellence in Universal Design. (n.d) *Definition and overview.* Available at: http://universaldesign.ie/What-is-Universal-Design/Definition-and-Overview/Definition-and-overview.html (Accessed: 7.9.21)

Grudin, Jonathan. (2009). *AI and HCI: Two Fields Divided by a Common Focus* Association for the Advancement of Artificial Intelligence.

IBM. (2020) *What is Artificial Intelligence?* Available at: https://www.ibm.com/cloud/learn/what-is-artificial-intelligence (Accessed: 7.9.21)

Marr, Bernard. (2018) *The Key Definitions Of Artificial Intelligence (AI) That Explain Its Importance.* Available at: https://www.forbes.com/sites/bernardmarr/2018/02/14/the-key-definitions-of-artificial-intelligence-ai-that-explain-its-importance/?sh=5c1fb9124f5d (Accessed: 7.9.21)

McCarthy, John. (2004) *What is Artificial Intelligence?* Available at: https://homes.di.unimi.it/borghese/Teaching/AdvancedIntelligentSystems/Old/IntelligentSystems_2008_2009/Old/IntelligentSystems_2005_2006/Documents/Symbolic/04_McCarthy_whatisai.pdf (Accessed: 7.9.21)

Paine, Chris. (2018) *Do you trust this computer?* [Documentary]

Store Norske Leksikon. (n.d) *Karel Čapek.* Available at: https://snl.no/Karel_%C4%8Capek (Accessed: 7.9.21)

Verne, Guri. (2020) *Adapting to a Robot: Adapting Gardening and the Garden to fit a Robot Lawn Mower*.

Module 2

Characteristics of AI infused systems

Amershi et al. (2019) defines AI-infused systems as ' systems that have features harnessing AI capabilities that are directly exposed to the end user'. This definition on it own can be quite vague, and it would be wise to look at the key characteristics of an AI-infused system to better understand what it entails.

The lecture in this class presents four characteristics of an AI-infused system; learning, improving, black-box and being fuelled by large data sets. An AI-infused system is dynamic and will change over time based on what it is taught by the people who are interacting with it. This also means that there is great potential for the system to improve and become better – that is if the system is designed to handle possible errors and unwanted inputs.

The users might perceive an AI-infused system as a Black-box, which entails that the functionality of the system (collecting, processing and generating data) is hidden from plain-sight and it can be difficult to understand how it works. Kocielnik et al. (2019) explains how the users often expect that the system is more knowledgeable than it is, and how it is important for designers to clearly state to them what the system can do to ensure a better experience with the product.

An AI-infused system is dependent on a large data set to be able to learn and improve – but it can also learn “too much” to the point where the generated output no longer matches the purpose of the system. Yang et al. (2020) explains how designers struggle to set appropriate expectations to the users when it may deliver such unpredictable outputs.

Spotify: An AI-infused system

Spotify is an AI-infused system which I am very familiar with. I listen to music everyday on this app, and everyday Spotify learns more about what music I might want suggested to me and which songs I might want to put together in a playlist. Spotify constantly learns about its users and improves, but I would not consider Spotify to be a black-box. The main data input that we provide as users is what music we choose to interact with, and the main generated output is what music Spotify thinks that we would like.

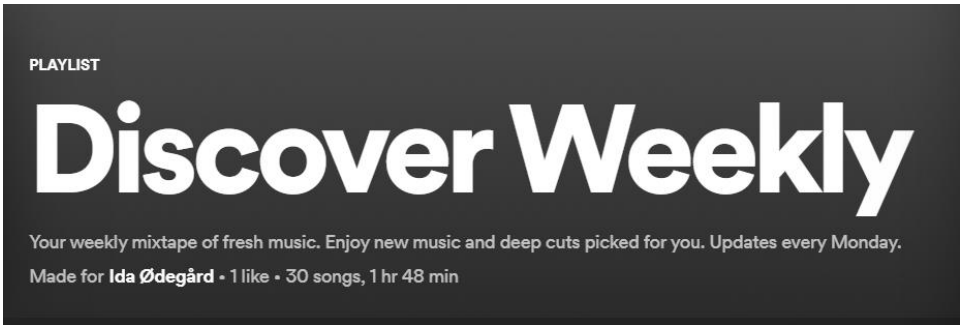
I can see where Spotify might be be affected negatively by gathering too much data. For example; if a user is unaware that Spotify is playing in the background, or if a group of friends is listening to music and the owner of the account doesn't normally vibe with that type of music – Spotify might learn in the wrong direction and suggest irrelevant music to this user.

Human-AI Interaction Design

Amershi et al. (2019) proposes 18 design guidelines for human-AI interaction. The reasoning for establishing these guidelines is that fast advances in artificial intelligence frames new opportunities and challenges for user interface design. The guidelines are thoroughly tested and validated by 49 design practitioners and aims to support those who want to create a great user-experience with their AI-infused systems. They are categorized by when they likely are to be applied during interaction with users; "Initially", "During Interaction", "When wrong" and "Over Time".

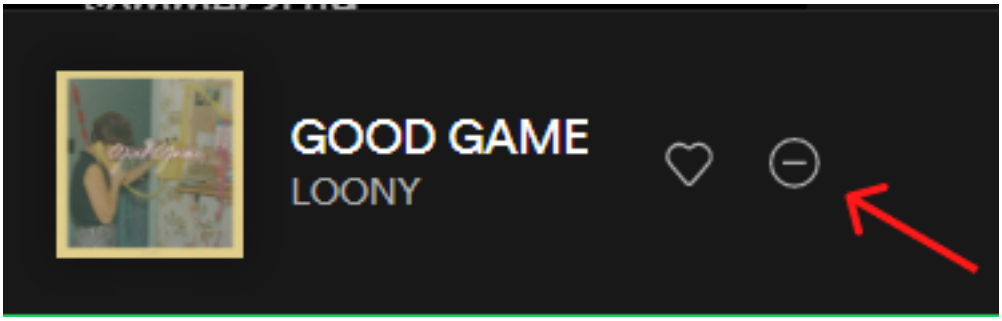
Kocielnik et al. (2019) explains how expectations impact how accepting end-users are of the technologies they use, and that they wish to examine how the impact can differ by using several methods of expectation settings in an AI-infused scheduling assistant. They suggest adjustment techniques that will increase acceptance in such technology based on their results; Including an accuracy indicator, giving an explanation based on examples and providing the user with the ability to adjust the performance of the system. The design expectation adjustment techniques will prepare the users for AI imperfections and hopefully result in an increase in acceptance.

Human-AI Interaction Design Guidelines: Spotify



G13: Learn from user behaviour.

Spotify states that I will get new music every Monday which is picked out specifically for me. The system bases all their suggestions on music that I have previously listened to, which shows that they learn from my behavior.



G15: Encourage granular feedback

In the playlists that are suggested to me, I get the opportunity to either “favorite” or “remove” the songs. The “remove” function only appears in this situation, which provides me with the ability to give feedback to the system on whether they are getting my preferences right or not.

AI-infused systems based on large language models

Bender et al. (2021) raises a critical discussion on AI-infused systems based on large language models. They aim to create awareness on how enlarging language models could both threaten the environment, and that they will require huge financial resources.

"Training a single BERT base model (without hyperparameter tuning) on GPUs was estimated to require as much energy as a trans-American flight." (Bender et al., 2021, s. 619).

Another issue is that the data that the models are trained with is full of bias that exists in our world. This is problematic because the system can easily be trained to think that sexism or racism is okay, and it will not be applicable in different countries with different cultures and norms.

Chatbots/Conversational User Interfaces

According to Følstad and Brandtzæg (2017), we are headed in a direction where we are not only interacting with machines by swiping, scrolling and clicking – but also through strings of text that represent our language. There are several challenges that can occur when designing chatbots. When interacting with chatbots, we input either text or speech and receive an output of generated data based on what the chatbot has learned from previous input. If the system is not prepared to handle unwanted input, it might learn from "bad" input and turn into a chatbot that is offensive or answers entirely different matters than its purpose. It can also be hard to establish the possibilities and limitations of a chatbot to the user, which can negatively affect their experience because they initially expect more from it.

G1: Make clear what the system can do

G2: Make clear how well the system can do what it can do

By adhering to these design guidelines when developing chatbots, one could avoid giving the users false expectations and instead deliver a service that they immediately understand and know how to interact with. A solution to making it clear what a simple chatbot could do could be to limit functionality by making the user have to interact with statements on buttons, instead of typing themselves.

Appendix

In this assignment I made sure to make my references clearer, and I have sharpened some sentences to make it easier to read.

References

Amershi, S., Weld, D., Vorvoreanu, M., Fourney, A., Nushi, B., Collisson, P., ... & Teevan, J. (2019). Guidelines for human-AI interaction. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (paper no. 3). ACM. (https://www.microsoft.com/enus/research/uploads/prod/2019/01/Guidelines-for-Human-AI-Interaction-camera-ready.pdf)

Bender, E. M., Gebru, T., McMillan-Major, A., & Mitchell, M. (2021). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?. In Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (pp. 610-623). ACM. (https://dl.acm.org/doi/pdf/10.1145/3442188.3445922)

Følstad, A. Brandtzæg, P. B. (2017). *Chatbots and the new world of HCI. interactions*, 24(4), 38-42. Lest: 19.10.2021

Kocielnik, R., Amershi, S., & Bennett, P. N. (2019). Will You Accept an Imperfect AI?: Exploring Designs for Adjusting End-user Expectations of AI Systems. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (paper no. 411). ACM. (https://www.microsoft.com/enus/research/uploads/prod/2019/01/chi19_kocielnik_et_al.pdf)

Yang, Q., Steinfeld, A., Rosé, C., & Zimmerman, J. (2020). Re-examining Whether, Why, and How Human-AI Interaction Is Uniquely Difficult to Design. In Proceedings of the 2020 CHI conference on human factors in computing systems (Paper no. 164). (https://dl.acm.org/doi/abs/10.1145/3313831.3376301)