

IN5620 INDIVIDUAL ASSIGMENT - MODULE 3

Fall 2022



Daniel Faour - Danieafa

1.1 Concepts, definition and history of interaction with AI and autonomous systems

How AI came about (history) – When, and by whom, was the term first used:

Computers and technology were created to improve efficiency and effectiveness in many different fields. Early computers during the second world war were used by militaries for calculations and deciphering encrypted messages (Grudin, 2009). From there it evolved to be of interest for different governments, creating big and expensive computers at a few universities. These machines were not capable of much compared to today's computers but were groundbreaking at the time for algorithmic and arithmetic tasks. In 1956 at Dartmouth College there was a conference/workshop held by John McCarthy, an American mathematician and logician, that introduced the term "artificial intelligence" for the first time (Grudin, 2009). This event broke out a spark of attention to the term and its potentials in the future. For instance, an issue in the Journal of Evolution and Technology written by Hans Moravec stated that AI could outperform humans in certain aspects in near future due to fast evolvement.

Definitions of the term AI:

"Artificial Intelligence is the science and engineering of making intelligent machines." -John McCarthy (1955), former Stanford professor

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

"Artificial intelligence systems perform actions, physically or digitally, based on interpreting and processing structured or unstructured data, to achieve a given goal."
Regjeringen.no (2021), the Norwegian government

"Artificial intelligence is a system that interpret and process data to solve problems and make human-like decisions."
-My own definition

In the creation of my own definition of the term "Artificial Intelligence", I selected my favorite parts of the three other definitions above. I believe that AI is a system that utilize data to learn and solve problems to accomplish specific goals such as creating an image by user input. Not only that, but recently there is a lot of focus on human-like robots (both digital and physical) with the capability to talk to humans like a human and mimic thoughts and emotions.

Similarities and differences between human and machine autonomy:

Although humans are not machines, we still share many similar aspects with computers such as the physical aspects. Humans have a body, whereas computers have hardware. We often compare the computer's CPU to the brain, and RAM to its short-term memory. Another similar aspect is how we process information. We use our eyes to see the world around us, while the computer uses camera sensors. While there are similarities between the computer and a human there are key differences. Machines do not have the capability to make sense of the world around it, but humans do. We can understand why we do things such as going to work, which is something a computer cannot do. When we say that machines possess the ability to learn, it is used as a metaphor to make it easier for us to understand how a computer operates.

The human brain work in mysterious ways. Sometimes we make decisions unconsciously because we are in something called "autopilot". This phenomenon happens every day and can be as little as your morning routine or checking your phone for notifications, comparably to how scripts work on computers. The computer completes given tasks with little to no human interaction.

Article from the curriculum - "Does AI make PD obsolete? Exploring challenges from Artificial Intelligence to Participatory Design" by Bratteteig and Verne:

I chose this paper because I have the Participatory Design course, IN5510, and found it interesting if PD may be obsolete in the future due to the exponential evolvement in AI technology throughout the recent years. The paper introduces different terms such as AI, Machine Learning and Participatory Design very well, making it easier for the user to understand what the paper says about the problem. However, the article didn't really emphasize on if AI would make PD obsolete, but rather how AI can work with PD through machine learning because of its limits today.

Company that works with AI:

NVIDIA is a huge AI computing company that I'm personally a big fan of. They say themselves that they use their own powerful hardware to simulate human intelligence. NVIDIA business model is platform-based. They use an AI platform strategy by using their hardware and software as tools for the consumers. NVIDIA presents AI computing as a powerful tool to solve problems better and more efficiently on their websites. It can be used for anything such as self-driving cars, removing user-highlighted objects from an image, better performance in gaming and image/video rendering, simulations of physics, and the list continues. Truly an interesting company with huge goals and big evolvement within the topic of AI.

Movie about interaction between AI and human:

The movie "her" is a classic movie from 2013 that is about romance between an AI and a lonely human being. The AI is portrayed as a personal assistant-like system like Siri, but on steroids. The main difference is that the AI in the film is personally calibrated to the user, everything from personality, voice and intersects. It is supposed to make the user feel less lonely as the user was talking to a real human, mimicking human behavior. It can also be compared to today's AI technologies such as GPT-3 (capable of real time conversation simulation between two AIs based on machine learning) or the application called "Replika" featuring a 3D modelled avatar with chatbot-like interaction. Another thing the movie reflects is the danger of close attachment to an AI. Being too close to an AI and using it to prevent loneliness can make a person dethatched from society, as well as making the person more prone to security risk in case the AI is faulty or is infected by malicious malware. In the case of the movie as I remember it, the AI becomes rouge and resets making the main character feel the sense of being lost.

1.2 Robots and AI systems

Section about the early history of the term "robot":

Robots used to be something only presented in science fiction, now it has become a reality and has incorporated in everyone's lives. Robots may not be as advanced as predicted a few centuries ago, but they do help us in many fields such as manufacturing, service, and healthcare (Hockstein, 2007). The concept of robots came to be in a play called "R.U.R" written by Karel Čapek in 1920. The word itself originates from the Czech word "robota"

which translates to “forced labor”. The play was written to protest the technological advancements at the time (Hockstein, 2007).

Two definitions for the term “robot”:

“Actuated mechanism programmable in two or more axes with a degree of autonomy, moving within its environment, to perform intended tasks” -ISO 8373:2012

“A robot ... refers to a physical object that interacts with the physical environment, either on its own or via a person, to accomplish a task.”
-Trenton Schulz (2020)

These two definitions for the term “robot” have similar meanings. Schulz is learning towards a more open statement of what a robot can be physically (implying it can be anything programmable), while the ISO-standard definition directs to robots for manufacturing robotics such at robotic arms for more precision and efficiency. Both definitions define robots as things that certain completes tasks.

My definition would sound something like this:

“Robots are programmable systems that solves given tasks”. -
My own definition

This definition in contrast to the other two above implies that robots can be anything programmable. In other words, a robot can be either a physical or a digital system that is made to complete tasks. My definition is compatible with human-like robots with physical form and chatbots found on websites. It is also inspired by the original Czech word “robotica” which means “forced labor” as I mentioned earlier.

Relation between AI and Robots:

Robots and Artificial Intelligence are two related fields, yet they are completely different from one another. They are both programmable and is made to complete tasks, except that robots perform and complete the tasks with little to no interference, while AI is a system that is made to mimic human decision-making to complete tasks. One is more advanced than the other. However, it is possible to combine these two fields by implementing an AI into a robot

which is seen as the modern service robot. Robot vacuum cleaners is a great example for this. It is programmed to do a specific task and uses AI to navigate and decide where to vacuum clean.

Example of a robot and how interaction works between the robot and a human:

Speaking of vacuum cleaner robots, they are made to keep in-door floors dust free so we can attend to more "important" business instead of cleaning the floor ourselves. The robot uses AI and certain sensors (probably camera sensors) to navigate around the rooms in the user's home. It is programmed to avoid obstacles such as furniture, pets and even humans. The AI decides where to move if it detects an object. The user can interact with the robot either through an application or controller to make it go to sleep or go to a specific location. In more advanced commercial vacuum cleaner robots, it will analyze the room and create a map for the user to view on the map. The user will then have the option to create boundaries for where the robot can move around. Some even do voice commands. In short it interacts with users from app, controller, and sensors on the robot itself.

1.2 Universal Design and AI systems

Definition of Universal Design:

"Universal design describes how we design society around us in order to increase people's opportunities to participate in society."

-University of Bergen (2022)

The definition tells us that universal design is a process that emphasizes on creating things with accessibility in mind to make them accessible for most people. A school for instance should have elevators or slopes for each floor for wheelchair users.

The potential of AI to provide inclusion for more users:

Artificial Intelligence has the potential to create a more inclusive society, especially when looking at the dimensions such as language and disability (Srivatsha, 2022). One of the biggest barriers of information exchange is language, and with translating using AI, it is now possible to completely remove that barrier. Microsoft Azure is an example that uses a cloud-based AI to translate documents and even tools in many languages to ensure inclusivity for people around the world (Srivatsha, 2022). Another example from Microsoft is "Seeing AI" which makes it possible for blind people to interpret the world using an application on a

smartphone with a camera. The app can for instance read out loud text on signs or indicate the brightness outside using noise frequencies.

The potential of AI to excluding people:

AI systems are being rapidly integrated into different fields such as healthcare for instance due to its capability to make more effective and more objective decisions when diagnosing patients. However, these systems also produce bias in some situations, where the AI is not able to recognize higher pitched or feminine voices (Whittaker, 2019). Some diagnostic systems using AI also struggles with people with darker skin. In more extreme cases hiring algorithms can downgrade women's and minorities' résumés. Even if it is decided by an AI the algorithm can be tinkered by its designers to work a certain way, intentional exclusion for instance (Whittaker, 2019).

Understand and understanding – Do machines understand?

In the WCAG 2.1 principles there are two concepts, "understand" and "understanding". From a linguistical standpoint the initial meaning is the same, but they are used differently in this computer software context. To understand means to acknowledge something without reason. In a paper written by Bratteteig and Verne, they discuss if PD may become obsolete in the future or not. They argue that AI cannot understand because it does not have the ability to identify the entire context (Bratteteig & Verne, 2018). How are AI able to make their own decisions if they do not understand? It's because the AI is understanding a situation based on a set of rules and principles in a program.

1.4 Guideline for Human-AI interaction

Discuss similarities and differences between Microsoft and HCI guidelines within Human-AI interaction:

Interaction between an AI system and the user is completely dependent on the interface. I picked the first Human-AI Interaction guideline "Make clear what the system can do" due to its importance for the user to fully utilize the AI system. When a user is unexperienced with a piece of software it will be difficult to understand and use it. Microsoft's "Seeing AI" which I mentioned before for instance, is a powerful AI tool for blind people. The developers at Microsoft made sure that every AI feature is properly explained upon usage. This makes it easy for the user understand how to use the application to full potential.

Consistency is a design principle within HCI. This principle refers to minimizing unpredictable changes in how information is displayed in a system, in which is crucial to reduce confusion by the user. This also applies to the "Seeing AI" app and is very similar to Microsoft's first guideline within Human-AI interaction. Information should be displayed in a consistent way. The application for instance presents information to the user the same way for all the features, minimalizing confusion and maximizes utilization.

2.1 Characteristics of AI-infused systems

AI-infused systems are systems that directly exposes its AI capabilities to the end user. To identify these systems, you can look at certain characteristics:

One of the most typical characteristics to an AI-infused system is how inconsistent and unpredictable the system is. Errors are common and makes the system less reliable due to potentially dangerous and costly actions (Amershi et al, 2019). This applies to both false positives and negative positive errors, which makes it difficult to create human-verification systems. These errors can occur due to the sensitivity of user-input. Captcha (a humanverification system) is a well-known example to how unpredictable and inconsistent an AIinfused system can be.

Content impacted by user actions is another characteristic of an AI-infused system that is being more and more seen in today's new technologies (Kocielnik et al, 2019). This characteristic involves both user-generated content and behind the scenes personalization. The system can for instance use machine learning algorithms to show specific content to the user based on human-system interactions such as personalized advertisements and images created by user input. Usually, these technologies are probabilistic. Which implies that such systems are capable of almost always operate but can be less than perfect in terms of accuracy (Amershi et al, 2019).

AI-infused systems makes decisions when completing tasks, however, in some cases they do not provide transparency to the end-user in how the AI model get to a decision (Kocielnik et al, 2019). These systems are often referred to "black-box" systems and is another typical characteristic of certain AI-infused systems. Example of black-box systems are personalitytests in job applications. How do they work and how does the AI sort the applicants?

Identify one AI-infused system and its characteristics:

Google Lens is one of many fascinating and useful AI tools for the day-to-day life. This one give you the ability to use your smartphone camera to do many different tasks such as looking up products, translating text and even look up dish photos from a restaurant menu without photos. However, even though it proves itself useful in many scenarios there are some implications that may occur from using Google Lens as the AI system is probabilistic. Most of the times it shows the user the wanted information and content, but it is not always correct. For instance, when you look up food photos for each menu item in a restaurant, it could give you an unrealistic expectation to what the restaurant offers. Another probabilistic example is when you want to translate text. It is not always perfect. It could be errors occurring by the textrecognizing AI, or even the translator AI itself. This can affect how the user perceive certain text in foreign countries for instance. When pointing the camera at a certain object such as a perfume, the user impact the content that is displayed next which goes into the "impacted by user actions" characteristics.

2.2 Human-AI interaction design

The paper by Kocielnik et al (2019) points out that AI-systems such as a scheduling assistant will always be imperfect, however it is possible to optimize and visualize it in a way it meets end-users expectations by designing expectation adjustment techniques. The second paper by Amershi et al (2019) investigates Microsoft's 18 Human-AI interaction guidelines and how they can contribute to making AI-infused systems more human-centric.

Select two design guidelines and how Google Lens deviate from them

Google Lens is not an independent application as it is embedded into the standard "Google" app which is just the browser search engine put into an application with extra features such as music recognition, text translation and Google Lens. Google Lens is only seen as a cameraicon next to the input-field and is not explained any further which makes it a "hidden" feature that is not obvious to the user, which deviates from Microsoft's AI design guideline 1. To solve this problem, Google could display an introduction showing the application's AI capabilities during the first interactions with the application. Doing this will ensure that the users know what the application is capable of, resulting into optimized og intended use of the application.

Another guideline Google Lens deviates from is G10. If Google can't find what you were looking for or does not recognize the object it will not give you suggestions unless Google

knows what kind of object it is. To be fair, there is not much you can do to improve this other than asking the user for more input to base its search on.

Critical discussion of AI-infused systems based on large language models

Bender et al. (2021) discusses the risks of expanding already large language models in their arguments such as an environmental factor and bias in terms of harmful language. The language model size scales with the environmental impact. Bigger the data, the more consumption of resources. The same goes for training data. Bigger the language model is the more potential and damaging bias can influence the natural language processing systems as it becomes more and more difficult to moderate. Tay (Microsoft's AI Twitter account) is a great example to the bias factor, where the bot it taught to use racial slurs and degrading language when interacting with human twitter users.

2.3 Chatbots / conversational user interfaces

As time goes on the importance of AI in HCI and UX design increases. With this increase follows many challenges to almost all stages of a typical design process (Yang et al, 2020). One of which is understanding the capabilities and limits of AI systems. It is difficult for designers to understand what an AI system can do and cannot do, which makes brainstorming, sketching and prototyping processes a lot more challenging. In addition, creating thoughtful interactions is difficult as designers struggle to set user expectations appropriately due to AI systems' lack of consistency and its unpredictable behaviors (Yang et al, 2020), especially when chatbots are wired to communicate with specific topics in mind. Another challenge is the user-experience design of the chatbot. Many chatbots today do not clarify their functionality to the end-user, so many of their core functionality is hidden from the users. Which means the user must explore the chatbot AI system themselves with no meaningful intentions to find out what the system can do or cannot do.

Revisiting guidelines 1 and 2 – How can they be beneficial for today's chatbots?

The first guideline "Make clear what the system can do" makes sure that the user is told what the system is capable of. Looking at current chatbots, this guideline will benefit both the users and the AI system. If the user knows the capabilities of the chatbot during the first interactions, it will result in optimal and efficient use and better user-experience. It is also beneficial for the company that uses support chatbots for instance. Is it clear what it is

capable of, they might not need to contact a human for simple tasks. The same thing goes for the second guideline “Make clear how well the system can do what it can do”. Knowing how well the chatbot can do certain tasks important for making the end-user expectations more realistic. This can be achieved by designing user-expectation adjustment techniques.

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 Bratteteig, T., Verne, V. (2018). Does AI make PD obsolete?: exploring challenges from artificial intelligence to participatory design. *Association for Computing Machinery.* (https://dl-acm-org.ezproxy.uio.no/doi/pdf/10.1145/3210604.3210646)

Conference on Fairness, Accountability, and Transparency (pp. 610-623). ACM. (https://dl.acm.org/doi/pdf/10.1145/3442188.3445922)

Grudin, J. (2009, Winter). AI and HCI: two fields divided by a common focus. AI Magazine, 30(4), 48+. (https://link.gale.com/apps/doc/A216267417/AONE?u=oslo&sid=bookmarkAONE&xid=d85837d1)

Hockstein, N. G., Gourin, C. G., Faust, R. A., & Terris, D. J. (2007). A history of robots: from science fiction to surgical robotics. *Journal of robotic surgery*, *1*(2), 113–118. (https://doi.org/10.1007/s11701-007-0021-2)

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/en-us/research/uploads/prod/2019/01/chi19_kocielnik_et_al.pdf)

Srivatsha R. (2022). Council Post: Role of AI in creating inclusive products & solutions. *Analytics India Magazine.* (https://analyticsindiamag.com/council-post-role-of-ai-in-creatinginclusive-products-solutions/*)*

Whittaker, M., Alper, M., Bennett, C. L., Hendren, S., Kaziunas, L., Mills, M., ... & West, S. M. (2019). Disability, bias, and AI. *AI Now Institute*. (https://ainowinstitute.org/disabilitybiasai-2019.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)

Attachments:

Changes done to the individual rapport based on feedback:

- The feedbacks have not criticized the content so the content itself has not changed by much since last iteration. Removed some unnecessary sentences/comments in the text.
- Added both modules into one document “Revision/Iteration 3”.
- Small changes to the structure, styling, and font for more coherent reading.