Individual assignment - Iteration 3

Innholdsfortegnelse

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

1.1.1 History of AI

The term artificial intelligence was first coined by the American John McCarty during a workshop held in 1956. Though, he was not the first to bring the concept of machines' ability to perform task previously only capable by humans into the daylight. World War II brought with it a lot of technological advancements. As a result of this the scientist Alan Turing wrote about the future of computers entering the realm of the human intellect in The Times in 1949 (Grudin, 2009).

1.1.2 Definitions of AI

Definition 1

The Norwegian dictionary SNL defines AI like this:
"Artificial intelligence is information technology (IT) that adjusts its activity and therefore appears intelligent."(SNL, n.d.)

Definition 2

The consultant agency Deloitte describes a broad definition of AI in an article from 2017:

"In general terms, AI refers to a broad field of science encompassing not only computer science but also psychology, philosophy, linguistics and other areas. AI is concerned with getting computers to do tasks that would normally require human intelligence. […]." (Deilotte,. n.d.)

Definition 3

Tone Bratteteig and Guri Verne, both with backgrounds from research and design, defines AI in their article about PD and AI:

«AI is a subfield of computer science aimed at specifying and making computer systems that mimic human intelligence or express rational behavior, in the sense that the task would require intelligence if executed by a human." (Bratteteig & Verne, 2018).

Based on the three definitions above I have concluded with my own definition:
AI is computer technology that seeks to accomplish tasks previously thought to require human cognition to be achievable. This includes tasks that require learning and tasks that mimic human behavior.

My definition focuses on the aspect of the technological progress connected to AI. I think AI differs from "regular computing" not because of its intelligence per se, but rather I think it is

a word we use to describe a more modern kind of “computing”, which can expand our use of computers into new areas of human activity.

1.1.3 Article review

I have focused on the article by D. Norman about the problem of automation (Norman, 1990). This article presents three case studies from the aviation domain on how automation has caused more risk related to unexpected situations. The article argues that to provide automation that is flexible and minimizes risk it has to be designed with feedback and interaction built in. Also, Norman suggests that the issue is not related to over-automations (which is a popular belief), rather it is a issue of under-automation.

The article discusses that in order to provide continual feedback from automated systems, without being a nuisance for the workers, the system needs to mimic human behavior. The article does not bring up the possibility of using AI technology to accomplish this. I think this may be because the article was written in 1990, and the technology has evolved a lot since then. It is intriguing to me that this is something that might be possible today. Equipped with the insights from this article, it could be possible to, for example, design a better autopilot for airplanes that provides tailored information to the crew.

The article presents a paradox that automation is often introduced to reduce workload and interactions, but that this design approach does not take into consideration the importance of feedback and interaction necessary to have a safe and efficient system. Automated systems have to be designed with errors and unexpected behavior in mind.

1.1.4 Tesla and AI

Tesla is a car company that produce electric vehicles, but is considered by many as more of a software company. This is due to the fact that they focus on development of autonomous cars through the use of AI technology.

“We develop and deploy autonomy at scale in vehicles, robots and more. We believe that an approach based on advanced AI for vision and planning, supported by efficient use of inference hardware, is the only way to achieve a general solution for full self-driving and beyond.”(Tesla, n.d.)

In the quote from Tesla they refer to AI as an approach. I think this very accurately describes Tesla as an innovative company. They do not just use the word AI as a “buzz-word”, and I think the company sees AI as a springboard for other innovations and opportunities to expand their reach as a software company.

1.1.5 AI in Ex Machina

In the film Ex Machina by Alex Garland the main theme explores what "artificial intelligence" is. The AI is embodied as a human looking robot called Eve. The plot of the film revolves around testing this robot and its interaction with a human participating in a project. The film defines AI as a completely conscious machine that is indistinguishable from humans.

1.2 Robots and AI systems

1.2.1 Origin of robots

The word "robot" comes from the Slavic word "robota" which can be translated into "forced labor". The word was first used in 1921 in the play Rossum's Universal Robots by Karel Capek. (Wired, n.d.)

1.2.2 Definitions of robots

Definition 1

Thrun (2004) mentions a definition from the Robot Institute of America from 1979, where a robot is defined as:

"A reprogrammable, multifunctional manipulator designed to move materials, parts, tools, or specialized devices through various programmed motions for the performance of a variety of tasks."

Definition 2

A more recent definition formulated by Matt Simons, a science journalist at Wired, defines a robot as such:

"A robot is an intelligent, physically embodied machine. A robot can perform tasks autonomously to some degree. And a robot can sense and manipulate its environment." (Wired, n.d).

These two definitions have many similarities, but also some important differences. The first focuses solely on the technical aspects of a robot, like how it should be able to move something to perform a task. Whereas the second definition includes aspect like the robot's ability to sense its surroundings and have intelligence. When I read the first definition I think of machines in a factory, and when I read the second, I think of a more complex robot, for example the Mars Exploration Rover.

I think the second definition is a more accurate reflection of my image of a robot, and I would therefore define a robot as: A robot is a physical machine capable of interaction with the environment autonomously through sensors and programmed actions in order to perform a task.

1.2.3 AI vs. robots

I think it is important to clarify that a robot is something physical that interacts with a physical environment. This is an important distinction that differentiates AI and chatbots from robots. Also, it needs to have a degree of autonomy, and not just be an extension of a human.

AI is computer technology that seeks to accomplish tasks previously thought to require human cognition to be achievable. This includes tasks that require learning and tasks that mimic human behavior.

When I look on my own definition of AI, I see a lot of overlap between AI and robots. But where the definitions differ is in the physical embodiment. AI refers to a program or algorithm that for instance can give a robot more dynamic and human behavior. A robot is a machine, whereas AI is a technology.

1.2.4 A contemporary robot

Boston Dynamics is a company that have been developing robots for a long time. One of their most recent developments is the robot dog called Spot. This is a human controlled robot with four legs. The four legs give it great stability and the ability to move relatively quickly. The robot uses 5 different camera arrays to scan and interact with its environment. And the view from all cameras is visible to the human on the controller unit, which looks kind of like a gaming controller. All the sensors give it the ability to walk without difficulty in many sorts of terrain and overcome obstacles. Spot is still in the research and development phase, but it is likely that a robot like this can be used in many different applications. For example, by the police or the military in search and rescue missions. It opens the possibility to accomplish tasks where it might be dangerous to send a human.

1.3 Universal Design and AI systems

1.3.1 Definition of Universal design

The definition of universal design from Digitaliseringsdirektoratet translated to English is:

“Universal design [of ICT] means that the users, regardless of their abilities, should in a good way be able to use both web pages and machines they encounter in everyday life. “
https://www.digdir.no/digitale-felleslosninger/universell-utforming-av-ikt/1499

What I like about this definition is the use of the word “abilities”, and not “disabilities”. I think a very useful perspective is to look at all people as unique, and with a different set of abilities. We live in an inclusive society. This means that we should design solutions for everyone, with inclusion on mind. This also does not mean to design specialized solutions for some people, but rather to strive towards one common design that works for all people.

1.3.2 The potential of AI

I think AI has huge potential for improving the inclusiveness of new technology in regard to human perception, movement and cognition. For example, an AI robot that uses sensors in combination with principles of animation to mirror a human can lead to more understandable interaction (Schulz et al., 2018). Another example where AI contributes to better inclusion is via the use of chatbots on webpages. This service can help people with difficulties navigating and finding information on a web page.

Sadly, the use of AI also has the risk of excluding more people. For instance, concerns have been raised connected to ethnic biases in facial recognition software. In one study they found that a facial recognition software had a much larger percentage of mismatches of women of African and Asian origin (BBC, n.d.). These kinds of issues raise important ethical questions about the use of AI, and this is important to be aware of as designers.

1.3.3 AI´s ability to understand

In the WCAG 2.1 principles it is stated that users with different abilities should be able to understand information and operations of the user interface, and also that different people understand differently. To understand something fully you also need to know the motivation or the meaning behind it, and I do not think this is something machines are capable of. For example, you could program a machine to divide a cake equally, but it would not “understand” that it should do so because we cherish equality in our society.

1.4 Guideline for Human-AI interaction

I have chosen guideline number 4 about showing contextually relevant information during interaction. This means that the information displayed should be relevant to the task the user is doing. Showing the user what options they have for interaction depending on the stage in a chatbot could be an example of this. Donald Norman´s HCI design guidelines of visibility, feedback, affordance, mapping, constraints and consistency have a lot in common with Microsoft´s guidelines (Norman, 2013). For example, visibility is about always showing the user what is going on, and giving them relevant information. This almost the same as Microsoft´s guideline number four. In general, I think that the Microsoft guidelines cover

more aspects that Norman´s. I find the guidelines from the "over time" section especially interesting, and more specific to AI. For instance that the AI should learn from user behavior.

2.1 Characteristics of AI-infused systems

When we are talking about AI infused systems we are mainly talking about artificial *narrow* intelligence. This describes AI systems with very specific abilities and functionality, like for instance self-driving cars. AI infused systems are 'systems that have features harnessing AI capabilities directly exposed to the end user' (Amershi et al., 2019). These systems can be ascribed four different attributes that differentiates them from "regular" systems.

Firstly, AI infused systems have the ability to learn. This means that they are dynamic and not static. The systems use data to learn from the users of the system over time. This is important because otherwise the system could not adapt and change with the user. (Amershi et al., 2019)

Secondly, AI infused systems tend to improve over time. This is closely connected to the previous point, the system has to learn and improve from new input over time. This is important because it is almost impossible to have all the necessary input data ready for the AI before release/use . Because of this, mistakes are inevitable, and there is a lot of uncertainty. AI infused systems need to improve based on this.

Thirdly, AI infused systems are black box systems. Black box systems are systems where the technical implementation is hidden from the user. The user can not know how input and feedback is handled by the system, or how the logic/algorithm of the system works. According to Kocielnik et al. (2019) this may pose some issues regarding transparency, and they try to come with solutions for how to more easily show and explain to the user how the AI system works.

And finally, AI infused systems are fueled by large data sets. The systems gather data from its environment and users through interaction. As mentioned previously, this is how AI systems are able to learn and improve. Yang et al. (2020) include this in their definition of AI witch says that 'AI refers to computational systems that interpret external data, learn from such data, and use those learnings to achieve specific goals and tasks through flexible adaptation.'

2.2.1 YouTube and AI

YouTube is an AI infused system that I use a lot. It is an online community-based video sharing platform. It is owned by the tech giant Google, who has focused a lot of resources on AI technology. I mostly notice YouTube´s AI capabilities through the recommendation

algorithm. If I visit YouTube after watching some videos on a specific subject, I always get recommendations for new videos on that subject. This signifies to me that the system learns from my use. You can also get suggestions based on what other people liked. Something like this: “People that viewed this video also liked… [recommendation]”. Based on this I can conclude that the AI uses large amounts of data from other users to achieve a desired result for my use. But, exactly how this happens YouTube does not say, so it is a black box.

2.2 Human-AI interaction design

2.2.1 Article review

Amerish et al. (2019) argue that the established design guidelines in HCI are not good enough for designing AI infused systems. This is because of the characteristics of AI like unpredictability and inconsistency. Through testing and validation they propose 18 design guidelines for AI divided into four categories based on different stages during interaction. These categories are “initially”, “during interaction”, “when wrong” and “over time”. With these guidelines Amerish et al. hope to enable the potential of AI by supporting better interaction between AI and humans.

The article by Kocielnik et al. (2019) also focus on the problems of varying user experience when interacting with AI infused systems. They argue that an important factor for shaping the user experience is related to user expectations. AI system´s behavior is hard for the user to predict, therefore creating a disconnect between expectations and lived experience. This results in lower acceptance of AI systems. Kocielnik et al. (2019) propose strategies related to transparency, so that the user can better understand what is going on. This, together with shaping the user expectations, they argue that it is possible achieve a better user experience.

Select two of the design guidelines in Amershi et al. (2019). Discuss how the AI-infused system you used as example in the previous task adheres to, or deviates from, these two design guidelines. Briefly discuss whether/how these two design guidelines could inspire improvements in the example system.

2.2.2 AI Guidelines in YouTube

G9: Support efficient correction. Make it easy to edit, refine, or recover when the AI system is wrong.

When YouTube suggest a video to the user there is an option to click “not interested”. By clicking this button, you tell the system/AI that you do not want content like this, and that is predicted wrong.

G13: Learn from user behavior. Personalize the user's experience by learning from their actions over time.
YouTube obviously learns from the user. Every time you watch a new video, or update the front page, the page is updated with new content based on what you have watched previously.

I think YouTube has pretty good alignment with the two guidelines I have mentioned. Still there might be room for improvement, for example in regard to G9: the system could give you the option to exclude certain categories of videos, or reset the suggestions if there are a lot of bad ones in general.

2.2.3 Article summary of Bender et al. (2021)
The article from Bender et al bring forward concerns regarding large language models. It argues that there are great risks related to environmental end financial costs. To power these language models great amounts of processing power is needed, which comes at a great cost. The article also argues that the gains from these language models only benefits certain groups, especially English speakers. Also, the data is maybe not as representative as one might think, because it is only retrieved from users of the internet. This, they argue, might lead to further marginalization in our society. To combat these risks Bender et al. suggests using considerable recourses on developing practices for data collection, as well as environmental and financial considerations.

2.3 Chatbots / conversational user interfaces

2.3.1 Key challenges in the design of chatbots
Chatbots are one type of AI infused systems. Many large companies and actors predict this technology to become a big part of the future, but there are still a lot of challenges related to chatbots.

Følstad and Brandtzæg (2017) addresses three main challenges going forward when designing chatbots. Firstly, we need to focus on the conversation as the object of design, and bring the focus away from graphical interfaces. Chatbots rely heavily on how words and sentences are used to achieve the user´s goal. This is someting designer have to focus on, at it is also very context dependent.

Secondly, Følstad and Brandtzæg argue that there need be a greater focus on service design. This is because chatbots are not, or at least will probably not, be connected to one device or touchpoint. It is likely that chatbots consist of different touchpoint where the technology is used within a larger service.

Thirdly, chatbots have the ability to interact with multiple humans and intelligent machine agent. These kinds of networks of interactions needs to be addressed when designing chatbot and conversational technologies.

In general, after being in the lectures and discussing the topic, I think an important part of chatbot design is related to the complexity and delicacy of human conversation. How humans talk and interact is very hard to replicate, and there are also ethical questions brought to light when designing conversational technology.

2.3.2 Guidelines G1 & G2

G1: Make clear what the system can do. Help the user understand what the AI system is capable of doing.

G2: Make clear how well the system can do what it can do. Help the user understand how often the AI system may make mistakes.

I think guideline G1 is important for making the user understand the limitations of the technology/system they are using. As mentioned previously, user´s expectations can have an impact on the user experience (Kocielnik et al., 2019). Giving the user information about what the system can, and cannot do, is a big part of that.

Guideline G2 is also related to user expectations. By making it clear for the user how well the system can do what it can do, it is much less likely that the user becomes disappointed or asks the chatbot to do something outside of it´s scope. For example, some customer service chatbots that I have used in the past are presented like personal assistants, but are actually more like glorified Q&As.

3. Appendix

Based on the feedback I received on iteration 1 I have changed the layout of the assignment. This includes adding informative subheading, numbering the headings and subheadings, and adding a table of contents. After feedback on iteration 2 I sorted my references. I was also informed where to find the Bender et al. (2021) article, so I added my summary from that.

4. 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-read y.pdf)

BBC. Facial recognition fails on race, government study says. Retreived from https://www.bbc.com/news/technology-50865437 Visited 09.09.21.

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. Retrieved from https://dl.acm.org/doi/pdf/10.1145/3442188.3445922

Bratteteig, T. & Verne, G. Does AI make PD obsolete? Exploring challenges from Artificial Intelligence to Participatory Design. Retrieved from https://dl.acm.org/doi/pdf/10.1145/3210604.3210646

Deloitte. Atificial Intelligence Defined. Retreived from https://www2.deloitte.com/se/sv/pages/technology/articles/part1-artificial-intelligence-defined.html Visited 07.09.21

Følstad, A., & Brandtzæg, P. B. (2017). Chatbots and the new world of HCI. interactions, 24(4), 38-42. (https://dl.acm.org/citation.cfm?id=3085558)

Grudin, J. (2009). AI and HCI: Two Fields Divided by a Common Focus. Retrieved from https://aaai.org/ojs/index.php/aimagazine/article/view/2271

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)

Norman, D. (2013). The Design of Everyday Things. Basic Books.

Schulz, T., Herstad, J., & Torresen, J. (2018). Classifying Human and Robot Movement at Home and Implementing Robot Movement Using the Slow In, Slow Out Animation Principle. International Journal on Advances in Intelligent Systems, 11, 234–244.

SNL. n.d. Kunstig Intelligens. Retrieved from https://snl.no/kunstig_intelligens Visited 07.09.21

Tesla. Artifical Intelligence and Autopilot. Retrieved from https://www.tesla.com/AI Visited 07.09.21.

Wired. The Wire Guide to Robots. Retrieved from https://www.wired.com/story/wired-guide-to-robots/ Visited 09.09.21.

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)