Third iteration

Individual Assignment – Module 1

1. Concepts, definition, and history of interaction with AI and autonomous systems

1.1 History of AI

Alan Turing, a famous British mathematician first discovered artificial intelligence (AI) during World War II when he cracked two German code machines, Enigma and Lorenz (Busch, 2018). These were encrypted messages sent by the Germans during the war. The term itself, artificial intelligence (AI), was first used and created in 1956 by the mathematician John McCarthy (Tidemann, 2020).

1.2 Definitions of AI

1. The IT company Kantega believes that there is no clear definition of the term but defines artificial intelligence as an interaction between adaptability and autonomy (Hovde, 2021). The definition reads as follows: *«Autonomi betyr at systemene kan utføre oppgaver uten kontinuerlig hjelp fra mennesker. Adaptiv er evne til å forbedre hvordan oppgavene utføres ved å lære av erfaringer.»*

2. Store Norske leksikon defines AI as *«[...] datamaskin som er i stand til å løse oppgaver uten å få instruksjoner fra et menneske på hvordan den skal gjøre det, har kunstig intelligens.»* (Tidemann, 2020).

3. A third definition is taken from SINTEF, which defines artificial intelligence as *«Kunstig intelligens er digitale systemer som utfører oppgaver som normalt sett krever menneskelig intelligens. Det gjør at oppgaver kan automatiseres og at mennesker kan få støtte til å ta gode beslutninger.»* (Sørensen, 2021).

My definition of the term artificial intelligence is machines that can perform tasks without the help of humans and can learn human behavior with the help of imitation. I would also like to add that I see artificial intelligence as algorithms that are trained by computers to become

intelligent and autonomous so that they can perform the same tasks as humans. My definition is inspired respectively on the definitions written above.

1.3 Similarities and differences between human autonomy and machine autonomy

There are differences between human and machine autonomy. One difference is that humans have feelings, which machines do not. On the other hand, there are also similarities, including "reasoning", i.e. thinking for humans and calculations for machines (Herstad, 2022). A similarity between the concepts automatic and autonomous is that both should reduce human effort. The concepts can be based on more efficiency and precision when performing a task.

1.4 Article from curriculum

1.5 Contemporary companies working with AI

Google is one of many companies working with AI. The company presents AI by talking about how they want everyone to have access to the information found on the platform, and that it should have a useful value for everyone and describes it as a product. Google explains that "AI is helping us do that in exciting new ways, solving problems for our users, our customers, and the world." (Dean, 2020). Google uses AI for many different purposes, including tackling problems within climate and energy.

1.6 Interaction between people and AI

The film "I, Robot" is based in the future in 2035 and is about robots tasked with keeping humans safe. Everything goes well at the start of the film, but the further into the storyline, you realize that the robots are out to take over the world and are a threat to humanity. In this film, AI is portrayed as a threat and something disturbing to humans.

2. Robots and AI systems

2.1 The word Robot

the word robot comes from the Czech word *robota* and means “forced labor”. The word was first introduced in the Karel Capek play, R.U.S in 1920 (Markel, 2011).

2.2 Definitions of robot

Store norske leksikon defines a robot as «*Robot er en maskin som kan utføre handlinger automatisk, og som ved hjelp av elektroniske «sanser» (sensorer) og «oppskrifter» på hva de skal gjøre (algoritmer) kan bearbeide inntrykk fra omgivelsene, som igjen påvirker hvordan disse oppgavene utføres.» (Søraa, 2022).*

This definition emphasizes the importance of algorithms and sensors in their definition of robot.

Britannica defines robot as *«**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.” (Moravec, n.d)*

This definition basically says that a robot does or can do everything a human being can do functionally but doesn’t appear like a human.

2.3 Relation between AI and Robots

Common for both definitions I described of the word robot is that it is a machine and this I would say is the biggest difference between AI and robot. The robot is defined as an object that can use AI in its technology.

2.4 Contemporary physical robot

“Sophia The Robot” from Hanson Robotics is a human-crafted science fiction character. She has, among other things, appeared on The tonight show with Jimmy Fallon, where she had a brief conversation followed by a duet with him.

3. Universal Design and AI systems

3.1 A definition of Universal Design

"Universal Design is the design and composition of an environment so that it can be accessed, understood and used to the greatest extent possible by all people regardless of their age, size, ability or disability."

I think this definition was a concise summary of what universal design entails. The term tells us something about the inclusion of everyone and the simplification of tasks encountered in everyday life

3.2 The potential of AI

AI has great potential in this field. It can contribute to the inclusion of people with various disabilities in everyday activities, such as "smart-home" devices in a home. Google Home and Alexa are good candidates in this category. Both technologies are voice-assisted. In this way, information can be made more accessible to someone with disabilities. You can, for example, set an alarm, play music and check the weather forecast just by using your voice. Another area of application is self-driving cars.

3.3 Describe the potential of AI for both including and excluding people.

Even though AI has great potential in universal design, it can still have some downsides. Facial recognition is a biometrics that is used by several actors. This type of identification has trouble with recognizing people with darker skin, because the AI system has only been fed with data on light-skinned test subject (Chou, m.m, 2017).

3.4 The concept "understand" and "understanding"

In order for something to be understood, I think it is important to have a certain degree of understanding. In the guidelines, it is emphasized that it is important to understand, in order to have inclusive and accessible designs for all kinds of users. As for the second question I think a machine understands by learning and feeding data.

4. Guideline for Human-AI interaction

4.1 Guidelines for human-AI interaction from Microsoft

13 Learn from user behavior: *Personalize the user's experience by learning from their actions over time*

An example of this is Spotify’s way of personalizing and suggesting new music based on the songs the user have listened to in the past.

4.2 Guidelines for human-AI interaction from Microsoft vs. HCI design guidelines

HCI Design guidelines: Nielsen's Ten Heuristic Principles are ten principles that helps to evaluate problems in an interface. In the guidelines for Human-AI it seems like the focus area is on the system itself as for the HCI guidelines its more around the value of a good user interface.

Feedback from the first iteration

The feedback I got was about a small improvement in section 2.1, about robots. I have therefore written more extensively about the definition.

Module 2

Characteristics of AI -infused systems

In this section of the module some of the key characteristics of AI-infused will be identified and described.

In the lecture with Asbjørn Følstad, he introduced four key characteristics of AI-infused systems. These were: **learning, improving, black box and that they are fuelled by large data sets**. By **learning** it means that the systems are adapting from human behavior, and from that they're dynamic. The systems learn from failure, and then **improve** their behavior. The systems are opaque and different, which means we don't know how the processing takes place. The systems are **fed with large data sets,** which means that they learn actively and passively through interaction with users.

AI-infused systems are *"systems that have features harnessing AI capabilities that are directly exposed to the end user"* (Amershi et al., 2019). This is the definition presented in the article. The three key characteristics that describes such systems are **uncertainty, inconsistency and lastly personalization.** Kocielnik et al. (2019 presents some characteristics similar to Følstad's lecture, such as transparency issues and the black box. Also, probabilistic behaviors are important here because the systems often change over time via learning. When designing for user experience Yang et al. (2020) points out the uncertainties that finds place, because there's yet to find out what the systems can do and will do.

An example of an AI-infused system that addresses some of the above-mentioned characteristics is Siri on Apple's products. Especially the iPhone. Siri is made of algorithms and large datasets and turns it into an AI-infused system that learns from the end-user. Siri is an AI that is directly visible to the ender-user. The AI has its imperfections and does not hide them. Whenever Siri misunderstand the users request, it pops up a error message with a voice that goes like "Sorry, I am not sure I got it correctly." . On some occasions, the end-user fails to give proper information or instructions to the AI, and therefore she plays the wrong song or gives the wrong forecast. Either way the AI does not cause a lot of damage to the user and adapts her suggestions fast based on the collected datasets.

Human-AI interaction design

To summarize what Amershi et al. (2019) and Kocielnik et al. (2019) discuss in their paper about interaction design in AI-infused systems, the main take-aways are how one can

incorporate the 18 guidelines for how to design AI. In the article it is pointed out that the scope of AI is extremely broad, and difficult to cover all the needs that must be considered. However, the guidelines are divided into 4 categories: initially, during interaction, when wrong and over time. These cover a broad specter and were tested several times by the authors and designers to test against AI. Kocielnik et al. (2019) study the expectations that's set on different types of AI systems in their article "Will You Accept an Imperfect AI? Exploring Designs for Adjusting End-user Expectations of AI Systems". In their work, they use a Scheduling Assistant - an AI system for automated meeting request detection in free-text email (Kocielnik et al., 2019, p. 1). To conclude they found that the results, driven by an algorithm, rather have higher number of relevant results than irrelevant.

In the category of "initially" there is one guideline (G2) that focuses on "making it clear how well the system can do what it can do". This means that the system must be built so that the user understands when the system makes a mistake or misinterprets a message. In our case Siri makes it pretty clear when she doesn't understand the message or the given task by saying something like "Sorry, I'm not sure I got the message", therefore I argue that Siri adhere to this guideline.

The second guideline is G4, "show contextually relevant information". This guideline tells us that the AI should display relevant information to the user's current task and environment (Amershi et al., 2019). If Siri is asked to show the weather forecast, and the user has allowed tracking of their position in the settings, the user will receive the weather forecast for his current location. I will therefore argue that Siri shows contextually relevant information and adheres to this guideline as well.

The article *"On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?"* By the authors Bender et al. (2021), discusses AI-infused systems that use large language models and the problematic around it. To summarize the article the authors, talk about the huge financial and environmental costs, as well as diversity in datasets appears to differ due to low to no access to internet in certain parts of the world.
In the article it is said that "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). Training such systems requires a lot of energy and computing power.

The datasets used to train these systems are not as diverse as one might think. This is something that creates an overrepresentation of certain categories. among these are white supremacy and misogyny. The outcome of this is a polarizing effect of already existing biases.
One recommendation from the authors to the problems stated above, is to carefully choose datasets and plan them thoroughly. It is also proposed to document the creation of a dataset, and only collect as much data as within the given budget.

Chatbots / conversational user interfaces

One of the key challenges in designing AI-infused chatbots is the way of thinking the interaction between design and what the chatbot is capable of doing. It can often be difficult to map out whether a chatbot understands a given task or not, and therefore difficult to create interactions based on that. “Designers frequently report that it is difficult to grasp what AI can or cannot do” (Yang et al., 2020). The result of this is poor interactivity between the chatbot and the user.

G1 “Make clear what the system can do” and G2 “Make clear how well the system can do what it can do” (Amershi et al., 2019). For these guidelines to apply to the challenge discussed above, one must make the systems capabilities clear to the designers. This can make the design process more cohesive and agile, as well as greater UX. One proposal for that to happen is to open a chat with an explanation of what’s happening. It could be as simple as how and what it can give answers to. In that order the user has more context of the situation.

Feedback Module 2:

The feedback was mainly about writing a bit more about the guidelines of Amershi et al. (2019), and reflect my own thoughts.

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/en-us/research/uploads/prod/2019/01/Guidelines-for-Human-AI-Interaction-camera-ready.pdf)

Busch, A.P (2018). Alan Turing. Alan Turing – Store norske leksikon (snl.no)

Hanson Robotic (n.d.). Sophia. https://www.hansonrobotics.com/sophia/

Hovde, E (2021, 3. februar) Hva er nå egentlig AI? [Blogginnlegg]. Hva er nå egentlig AI? | Kantega

Ibars, R., Murillo, O. & Chou, J. (2017, 26. September). How to Recognize Exclusion in AI. [Blog post] https://medium.com/microsoft-design/how-to-recognize-exclusion-in-ai-ec2d6d89f850

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)

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)

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)

Markel, H (2011). The origin of the word “Robot”. The Origin Of The Word 'Robot' (sciencefriday.com)Tidemann, A (2020). Kunstig Intelligens. kunstig intelligens – Store norske leksikon (snl.no)

Moravec, H.P. Robot. *Britannica*. *https://www.britannica.com/technology/robot-technology*

National Disability Authority (n.d.). What is Universal Design. *https://universaldesign.ie/what-is-universal-design/*

Søraa, R.A. (2022). Robot. *https://snl.no/robot*

Sørensen, R.S (2021). Hva er kunstig intelligens?. *SINTEF*. Hva er kunstig intelligens? (sintef.no)