Maria Nystuen Roheim

Second iteration

IN5620 Individual assignment fall 2022

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

The history of artificial intelligence

Grudin (2009) examines the history of artificial intelligence and the factors that have influenced its emergence. AI originated in mathematics and engineering and was developed by Alan Turing during the second world war. Turing claimed that in 1949 computers should enter the fields typically covered by human intelligence and eventually be able to compete on an equal level with humans. The term AI was first introduced in 1956 by the mathematician and logician John McCharty (Grudin, 2009, p. 49). AI has evolved from being rule-based to algorithms and mathematics that mimic human problem-solving. Today, computers derive their own rules based on the data it processes and "experience" instead of implementing given rules and following them (Bratteteig & Verne, 2018, p. 2).

Four definitions of AI

In the article "Does AI make PD obsolete?" Bratteteig and Verne (2018) define AI as:

> *"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, pp. 2-3).

Merriam-Webster defines AI as such:

> *"A branch of computer science dealing with the simulation of intelligent behavior in computers" and "the capability of a machine to imitate intelligent human behavior»* (Merriam-Webster).

SINTEF defines AI as:

> *"Digital systems that perform tasks that normally require human intelligence. This means that tasks can be automated and that people can be supported in making good decisions*" (SINTEF, 2021).

Both the Merriam-Webster, Bratteteig and Vernes and SINTEF definition emphasizes that AI must act as a human would, focusing on intelligent human behavior. SINTEF's definition also includes an explanation of what AI is used for. Based on these three definitions, I would say that:

"*Artificial Intelligence is an information technology that performs tasks by simulating and imitating intelligent human behavior. This means that the AI performs tasks requiring human traits and intelligence. I think the human perspective."*

Human autonomy and machine autonomy

Human autonomy is about a person's ability to determine their actions, and the reason for doing something is their own. Machine autonomy is similar in that it refers to a machine's ability to accommodate variations in its environment. For a machine to be autonomous, it must be capable of achieving something without human interventions (Ezenkwu & Starkey, 2019). Thrun (2004) claims that autonomy is the biggest difference between robots and other physical devices. The degree of autonomy is measured by relating the degree to which the environment can be varied to the mean time between failures and other factors indicative of robot performance (Thrun, 2004, p. 10). A robot's autonomy determines the tasks it can perform. The more complex the environment is, the more autonomy a machine needs. (Thrun, 2004, p. 14). That is a big difference from human autonomy, where we do not need to consider the environment the same way a machine does.

Similarities between the concept of automatic and autonomous

The concept of automatic is about things happening without the intervention of the will or humans. It is a self-regulating mechanism. This differs from autonomous, meaning we can make choices and do something independent of outside influences. The concepts are similar in that they are both about machines acting without human assistance.

Does AI make PD obsolete?

The article "Does AI makes PD obsolete" by Bratteteig and Verne (2018) looks at how AI challenges PD and which design decision users can take part in when we design AI technology. The authors question if AI accommodates all your needs and provides tailored

solutions, will it be necessary to design for user needs anymore? (Bratteteig & Verne, 2018, p. 1). They claim that AI challenges PD in several ways, including that if users are to be able to participate in the design process, they must have sufficient knowledge of the possibilities and challenges of the technology to be used. This could be difficult with AI. Furthermore, they explain that AI's behavior is unpredictable as it changes over time based on the data it receives, especially considering that it may be incomplete or biased. Another challenge is that it is difficult for users to evaluate an AI system as they will only be able to see the interaction mechanisms and the design of the system, while the statistics and data sets used will be too extensive to evaluate (Bratteteig & Verne, 2018, p. 3).

AI at Amazon Web Services

Amazon Web Services presents AI as a product and service they can deliver. They offer consumer and business-oriented AI products and services and advertise with the promise that their services do not require a machine learning experience. They explain that they can provide ready-made intelligence for consumers' applications and workflows. The services integrate with their consumer applications to improve safety and security, increase customer engagement, and create personalized recommendations. One example is the business version of the intelligent voice server Alexa, called Polly, which turns text into speech (Maguire, 2022).

Black Mirror – Hang the DJ

This episode of Black Mirror describes a future where an AI facilitates romance and finding a partner. The AI measures the love compatibility of individuals by running simulations based on dating experiences to find the perfect match. The more people you date, the closer you get to finding the one. The human interaction with the AI is portrayed as a relationship where they are friends. They trust the algorithm and talk to the system about everything they are feeling, asking them for advice, and they do as the AI says. Ultimately, the relationship changes when the participants begin to doubt the AI and decide to make their own choice instead.

1.2 Robots and AI systems

First, write a section about how the word Robot came about.

The word Robot was first used by Karel Capek in 1920 in a play that describes a factory that creates artificial, organic human copies called “robotics,” which rebel against and exterminates humanity. The word derives from the Czech word Robotnik, meaning “worker” and “slave” (Søraa, 2022). The concept of a robot is constantly changing, and what we consider a robot today will probably not be the same in fifty years (Dautenhahn, 2018, p. 2).

Two definitions of robot

Merriam-Webster defines a robot as

> *“a machine that resembles a living creature in being capable of moving independently (as by walking or rolling on wheels) and performing complex actions (such as grasping and moving objects)”* (Merriam-Webster).

Store Norske Leksikon defines a robot as:

> *A machine that can perform actions automatically and therefore consists of both software that instructs it in tasks, and hardware that gives it a "body" to move* (Søraa, 2022)

The Merriam-Webster definition underlines the concept of automatic as the robot needs to be able to move without human intervention, while the Store Norske Leksikon focuses on the robot requiring both software and hardware. Based on these two definitions, I define a robot as:

> *“A robot is a machine including both hardware and software, that can perform tasks automatically based on instructions.”*

The relation between AI and Robots.

From the Store Norske Leksikon definitions, one thing that separates AI and robots is that robots need to have a “body.”. From my definition, I would say that one thing that separates AI and robots is that AI has autonomy (at least some), which means that it does not rely as much on humans as robots do. A robot cannot think and make decisions for itself; as Bratteteig and Verne (2018) write about AI, a robot cannot, in the same way, do tasks that would require human intelligence. For example, a robot vacuum cleaner will not realize by itself that it needs to vacuum.

A robot vacuum cleaner

A typical robot that many people use nowadays is a robot vacuum cleaner. When interacting with the robot, we use an app to tell it when to start, where to move and when to stop. The interaction is also influenced by the fact that we often name the vacuum cleaner, and some people talk about it and to it as a person. For the robot to work correctly, it needs us to tell them how to behave and what to do. It also depends on us cleaning the area to be vacuumed before it starts, by moving wires, chairs, etc., for the robot to perform as we want and move around correctly.

1.3 Universal Design and AI systems

Universal design and inclusion

Universal design (UD) is about us designing for everyone. The National Disability Authority describes it as "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 age, size, ability or disability" (Authority). Therefore, we aim to meet the needs of those using our design. Herstad (2022, p. 2) explains that the main aim of UD is that when we create new systems, these should not create barriers for users, and the existing obstacles should be reduced. It can be seen as essential inclusion work and human rights where we design systems that are accessible to everyone, without us having to make different arrangements and facilitation for these people.

The potential of AI

AI has a huge potential to include humans with different perceptions, movements, and cognition. Bratteteig and Verne (2018) explain that AI provides the opportunity to receive personal services in a form that it can help answer your questions and make life easier. They describe that an AI will know it and "understand" what you need and therefore meet your needs. (Bratteteig & Verne, 2018, p. 1). For example, AI can be used to add text to internet content so that people with visual impairments can have content read out loud. Another example is that AI can help with moment disabilities by doing tasks like Siri can write an email for you.

AI including and excluding

Even though AI has the potential to include people by tailoring a system to a user's needs, it can also exclude people. One of the problems with AI is that it is usually based on statistical data, like average and median. This means that some people will not be within this data group. If data is excluded from the AI, this will also exclude people. A concrete example is facial recognition systems, which could cause racial bias and discrimination concerns.

Understand and understanding

The concepts “understand” and “understanding” are about comprehensible information and interfaces. I understand the terms to mean that we try to give meaning to things by perceiving and analyzing an input. Both humans and machines "understand," but in different ways. In the machine, we say that we receive an input, this is processed, and we give an output. It can be argued that the machine tries to "understand" the input during processing, but humans have made this link between input and output and what will happen. We can say that machines can understand in the sense that they can perceive a command or message.

1.4 Guideline for Human-AI interaction

I have selected the guideline for human-AI interaction, nr. 16,” Convey the consequences of user actions. This guideline shows and explains how the user's actions affect the AI and its effects for future use, for example, how the AI adopts user behaviors.

There are several similarities between Microsoft’s guidelines and HCI. One example is the nr.11, “Make clear why the system did what it did,” and in HCI, “offer Informative feedback» (Shneiderman). This guideline is about explaining to the user what happens and why. An example of differences between the two guidelines is where Microsoft focus on the AI, for example, “Learn from user behavior,” where the AI adopts the system to personalize the user’s experience by learning from their actions over time.

Feedback from the first iteration

From the first iteration, I got feedback on one thing I could improve, to be more specific in explaining similarities between the concept of automatic and autonomous. Therefore I have worked on explaining these concepts better.

MODULE 2

1.5 Characteristics of AI-infused systems

Key characteristics of AI-infused systems

Amershi et al. (2019) describe AI systems as "*(..) systems that have features harnessing AI capabilities that are directly exposed to the end user*". Asbjørn (2022) characteristics these systems as systems that are learning, improving, black box, and are fueled by large data sets

- **Learning**: An AI-infused system is dynamic because it learns from the user's behavior and changes over time. Amershi et al. (2019) describe that AI components can be inherently inconsistent because of this. This means that the system can behave differently from one user to the next, i.e., by returning different results in search engines based on personalization.
- **Improving**: One characteristic of an AI-infused system is that errors are common, and making mistakes is inevitable. Amershi et al. (2019) describe that we strive to constantly improve to create intuitive and effective systems and to design for uncertainty.
- **Black box** is a term used for explaining AI systems where the user cannot see the input and output. The systems are opaque, which makes it difficult for the user to understand and validate output, and we, therefore, need to design for explainability (Kocielnik et al., 2019)
- **Fueled by large data sets**: AI-infused systems are fueled by large data sets gathered through interaction (Asbjørn, 2022). Yang et al. (2020) describes that in the ideal case, the system will grow based on this data, and adapt to more varieties of users and contexts. Sometimes this new data also drives system performance in the wrong direction, deviating from what the system was developed to do.

Spotify and AI

Spotify is an AI-infused system that I use frequently. One of the characteristics of this system is that it learns based on my behavior. It collects data about the music and podcast I listen to and uses this to generate suggestions for things it think I might want. This makes the system

feel more personal because it creates playlists based on my taste in music and shows me things like “Your most popular mixes” and “Based on what you recently have been listening to.”. Another characteristic is that some parts of the system are fueled by large data sets, like charts. When people listen to music, Spotify saves this data to present top songs globally and for specific countries.

1.6 Human-AI interaction design

Interaction design for AI-infused systems

Amershi et al. (2019) claim that we need more study and innovation on AI technologies and their growing uses because AI systems can violate established usability guidelines of traditional user interface design. He claims that AI-infused systems may demonstrate unpredictable behaviors that can be disruptive, confusing, offensive, and even dangerous. Therefore, Amershi et al. (2019) present 18 guidelines for designing for human-AI interaction. These are based on multiple rounds of evaluation, including a user study with 49 design practitioners who tested the guidelines against 20 AI-infused systems. Amershi et al. (2019) believe these design guidelines can serve as a resource for people designing AI technologies.

Kocielnik et al. (2019) have a different angle on interaction design for AI-infused systems; instead of developing guidelines for design, they investigate how we can manage end-users' expectations of AI-infused systems for designing for a better user experience. They present three techniques for how we can do this. The first one, “Accuracy Indicator”, is about showing and explaining to the user how accurate the system is. The second is “Examples based explanation,” which aims to increase user understanding by explaining the basic principles of the system. The last one is the “Control Slider”, which allows the user to directly adjust the performance of the system (Kocielnik et al., 2019, p. 2)

Two guidelines for AI-infused systems – Spotify

G4 - Show contextually relevant information.

Spotify follows this guideline by recommending podcast episodes and music for me while in the explore/search/home section. It has a category containing “Episodes for you”. This information is not displayed when I am in my library, which means the AI considers my

current task and environment. This guideline can inspire improvements in the system by just showing the user what is necessary, preventing the user from getting confused or making errors. It can help display the relevant information for my task, for example, showing me all the albums from an artist if I search the artist's name.

G12 – Remember recent interactions.
Spotify remembers my previous interaction, i.e., "recent searches," so it is easy to find back things I have been looking for. This guideline can inspire improvements by remembering for the user. It could also autofill the rest of a sentence if you remember a line in the lyrics and not the artist or name of a song.

Critical discussion of AI-infused systems

Bender et al. (2021) argue that large datasets based on texts from the Internet overrepresent hegemonic viewpoints and encode biases, and that the size of the data material does not guarantee diversity. There are several reasons for this, one being that people with a hegemonic viewpoint are more likely to be retained and, therefore, overrepresented. Another reason is that Internet access is not evenly distributed, so the people who contribute to Internet text collections will be from a narrow group, for example, the data will be overrepresenting younger users from developed countries. This could potentially damage marginalized populations (Bender et al., 2021, p. 614). Another argument concerning the problematic aspects of textual content is changing social views. If a social movement is poorly documented or does not receive significant media attention, there will not be text about this, and the data becomes biased. The consequences could be that the data language models are based on will misrepresent social movements and regimes of power (Bender et al., 2021, p. 614)

Bender et al. (2021) explain the term "documentation debt", meaning that when we rely on large datasets, we risk that the datasets become both undocumented and too large to document post. Bender et al. (2021) argue that the solution to this problem is to budget for documentation as part of the planned costs of dataset creation. This is because, without documentation, we cannot understand the datasets and cannot learn and reduce mistakes.

1.8 Chatbots / conversational user interfaces

Key challenges in the design of chatbots and conversational user interface

We have several different chatbots which serve different purposes, like customer service, assistance, guides, and social relations (Asbjørn, 2022). When designing for usability, we are used to thinking of navigation through menus and links and interaction with graphical elements. We see design as a task of explaining to the user which content and features are available and which steps to take to reach the desired goal. This differs from the design for chatbots and CA, where we need to try to solve the task of understanding what the user needs and how he/she may best be served. One of the key challenges here is that we are designing hidden features for the user. The interaction and output depend more on the user's input. Therefore, we need to design suggestions for the user about what he/she may expect in the service and interpret their response. We need to see conversations as the object of design, which could be challenging. Another challenge is that we will not just by designing for a specific user interface. However, we must consider the entire service process across conversational touchpoints with the user (Følstad & Brandtzaeg, 2017, pp. 40-41).

When designing chatbots and conversational user interfaces, a challenge is that user expectations are usually far from the practical realities of use. This is reinforced by a lack of feedback regarding the system's capability and intelligence. This makes it hard to match the user's mental model, leading to disappointment and confusion. In relation to this, the anthropomorphism of the systems sets unrealistic expectations because it frames the user's perceptions of what the system can do (Luger, 2016, p. 5294). In addition, it is hard to design chatbots and CA that people trust. Luger (2016) explains that users are unwilling to use CA on complex or sensitive tasks, like making Siri call a person, because they cannot ascertain the level of system intelligence. This causes them to confirm the tasks visually, and the point of the CA is gone.

Guidelines G1 and G2

I think both guideline G1, “Make clear what the system can do,” and guideline G2, “Make clear how well the system can do what I can do what it can do,” could resolve some of the challenges in current chatbots/CA. As previously discussed, user expectations are usually far

from the truth of what a system can do. If we can help the user understand what the AI system can do, we can avoid the user getting disappointed, confused, or irritated. This will also be easier to achieve if we explain to the user how often AI systems make mistakes.

References

Amershi, S., Weld, D., Vorvoreanu, M., Fourney, A., Nushi, B., Collisson, P., Suh, J., Iqbal, S., Bennett, P., Inkpen, K., Teevan, J., Kikin-Gil, R., & Horvitz, E. (2019). Guidelines for Human-AI Interaction. *Conference on Human Factors in Computing Systems*, 1-13. https://doi.org/10.1145/3290605.3300233 (CHI '19)

Asbjørn, F. (2022). Interaction with AI and autonomous systems - Module 2.

Authority, N. D. What is Universal Design. https://universaldesign.ie/what-is-universal-design/

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? , 610-623. https://doi.org/10.1145/3442188.3445922

Bratteteig, T., & Verne, G. (2018, 2018). Does AI make PD obsolete?: exploring challenges from artificial intelligence to participatory design.*PDC '18* ACM International Conference Proceeding Series,

Dautenhahn, K. (2018). Some Brief Thoughts on the Past and Future of Human-Robot Interaction. *ACM transactions on human-robotic interaction*, *7*(1), 1-3. https://doi.org/10.1145/3209769

Ezenkwu, P., Chinedu, & Starkey, A. (2019). Machine Autonomy: Definition, Approaches, Challenges and Research Gaps. https://link.springer.com/chapter/10.1007/978-3-030-22871-2_24

Følstad, A., & Brandtzaeg, P. B. (2017). Chatbots and the new world of HCI. *Interactions (New York, N.Y.)*, *24*(4), 38-42. https://doi.org/10.1145/3085558

Grudin, J. (2009). AI and HCI: Two fields divided by a common focus. *AI Magazine*, *30*(4), 48-57. https://doi.org/10.1609/aimag.v30i4.2271

Herstad, J. (2022). IN5620 Lecture notes 31 August 2022 – Universal Design and AI. In.

Kocielnik, R., Amershi, S., & Bennett, P. (2019). Will You Accept an Imperfect AI?: Exploring Designs for Adjusting End-user Expectations of AI Systems. *Conference on Human Factors in Computing Systems*, 1-14. https://doi.org/10.1145/3290605.3300641 (CHI '19)

Luger, E. S., Abigail. (2016). "Like having a really bad pa": The gulf between user expectation and experience of conversational agents.

Maguire, J. (2022, March 30.). *Top Performing Artificial Intelligence Companies of 2022*. https://www.datamation.com/featured/ai-companies/

Merriam-Webster. Artificial intelligence. In https://www.merriam-webster.com/dictionary/artificial%20intelligence

Merriam-Webster. Robot. In https://www.merriam-webster.com/dictionary/robot

Shneiderman, B. *Guidelines in HCI* tutorialspoint.com. Retrieved September 12. from https://www.tutorialspoint.com/human_computer_interface/guidelines_in_hci.htm

SINTEF. (2021). Hva er kunstig intelligens? https://www.sintef.no/fagomrader/kunstig-intelligens/hva-er-kunstig-intelligens/

Søraa, R. A. (2022, May 12. 2022). *Robot*. Store Norske Leksikon. Retrieved September 11. from https://snl.no/robot

Thrun, S. (2004). Toward a framework for human-robot interaction. *Human-computer interaction*, *19*(1-2), 9-24. https://doi.org/10.1207/s15327051hci1901&2_2

Yang, Q., Steinfeld, A., Rosé, C., & Zimmerman, J. (2020). Re-examining Whether, Why, and How Human-AI Interaction Is Uniquely Difficult to Design. *Conference on Human Factors in Computing Systems*, 1-13. https://doi.org/10.1145/3313831.3376301 (CHI '20)