First module – iteration 1:

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

The origin of AI

Jonathan Grudin (2009) writes in his article about the origin of Artificial Intelligence (AI). During World War II, Alan Turing made a statement about how computers should be able to compete with the human mind, which led to a series of events before the term AI was first introduced in 1956 by John McCarthy (p. 49). Over the next decades, AI research would experience multiple ups and downs when it came to interest in the field (Grudin, 2009, p. 49-55).

Different definitions of AI

Bratteteign and Verne (2018) defines 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."* (p. 1-2). This definition was stated in an article from 2018 describing the relationship between AI and Participatory Design (PD), discussing whether PD will survive the evolution of AI.

IBM is a contemporary company that generally works with technology. On their website, IBM describes that *"Artificial intelligence leverages computers and machines to mimic the problem-solving and decision-making capabilities of the human mind"* (IBM Cloud Education, 2022).

The third definition is from the Norwegian governments' pages. They define AI as systems that execute either physical or digital actions in order to achieve a certain goal. These actions are based on what the systems learn from data (Kommunal- og distriksdepartementet, 2022).

The definitions presented are from three quite different sources. The first mentioned is from an article discussing it in relation to a specific field within design, the second from a private tech company, and the third from a public sector. What these three definitions have in common is the notion that AI simulate the human mind and decision making, and that they learn from massive amounts of data that is presented. My personal definition will therefore be that AI is systems that is capable of simulating or mimicking human capabilities. How well

they are able to do this is limited or supported by the amount, as well as the quality, of data that the systems consume in order to learn.

Differences between human autonomy and machine autonomy

Machine autonomy refers to how well the machine in question can react and adjust to different environmental factors (Thrun, 2004, p. 14). How advanced this autonomy is differs from machine to machine, and is often related to how fast it can adjust or fail to different settings (Thrun, 2004, p.14). While machines and robots often are characterized by the state of the autonomy (Thrun, 2004, p. 14) human autonomy is a general concept embedded in a person. Human autonomy does not describe how well a person will function within a set environment, as machine autonomy does.

Similarities between the concept automatic and the concept autonomous

The concept of something being automatic is often considered as something being executed without a second thought or routinely. Machine autonomy is quite like the concept automatic. "Good" or "strong" autonomy in machines might suggest that it adjusts properly and executes the tasks that it was set to do. One might see this as a machine´s automatic response to the environment and the tasks that it is supposed to do.

Brief review of article from the curriculum

The article "Intelligence on Tap: Artificial Intelligence as a New Design Material" (2017) discusses how the evolution of AI and the new approaches to learning, specifically *deep learning*, have increased the capability of the AI systems which again will challenge and affect how HCI research and design work (Holmquist, 2017). The author describes how intelligence and AI *"will become a new design material."* (Holmquist, 2017, p. 31) and presents throughout the article different challenges associated with this. I am curious to learn just how prominent AI will become in designing solutions, as it still can be quite unpredictable.

Contemporary company, and how they present AI on their web pages

To discuss how a contemporary company describes their work with AI, I decided to look at Apple´s approach to the subject. On their website, they present their products as a combination of a physical object as well as a service which will improve your life. The new apple watch has "Adventure awaits" and "A healthy leap ahead" (Apple, 2022a) as their

immediate taglines, implying how they will contribute to change your life. On their career site for Machine Learning and AI (Apple, 2022b) it is presented as an idea or experience.

Human interaction with AI in a fictional film

The fictional film "Avengers: Age of Ultron" (2015) portraits how scientists design and creates a fully conscious and functional humanlike robot who absorbs massive amounts of data it accesses from the internet, and creates its own perception of the world, along with an agenda to make the world "right". Human interaction with AI is portrayed as rather dangerous and reckless in this film. Although the scientist's idea and vision were a peacekeeping, smart and "good" AI robot, the results are catastrophic.

1.2 Robots and AI systems

The word "Robot" was first introduced in1920 in a Czech theatre performance (Søraa, 2022). However, the first major introduction of robots like todays were in the 1960´s, and in the 1970´s industrial robots were being used on assembly lines (Thrun, 2004, p. 11).

Different definitions of "robot"

Merriam-Webster defines "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, n.d.).

Further, it is described by Thurn as *"systems with multiple actuated elements, often arranged in chains."* (Thrun, 2004, p. 11). This description is usually used in the context of industrial robots. Thurn explains how a key characteristic in robots are that they are autonomous. This enables the robots to make choices based on assessments made in the situation, and are more applicable to more modern robots.

Based on these definitions, I would classify a robot as a system or machine, capable of executing tasks or movement either on its own, or by being controlled by a person. The reason for this is that I understand the term "robot" as something broad, which can apply to many different creations, and work either automatic or autonomously on different levels.

Relation between AI and Robots

Considering the definitions I described as my personal understanding, the main difference between AI and Robots is that AI do not in most instances have a physical "body". A Robot executes tasks they are set to do in physical environments, while AI are often found in systems. While AI learn from the data they are presented and evaluate this to continue their growth and decision making, robots are often programmed to do specific task without the possibility of learning as they go. However, I believe there are multiple robots with AI as well, who can be classified as more of a hybrid between the two.

Contemporary physical robot

A common contemporary physical robot is the robotic vacuum cleaner. This type of robot is often seen in many households and are a kind of "personal service robots" which aids the user in a domestic setting (Thrun, 2004, p. 12). The robotic vacuum cleaner often has a charging station that it goes back to, and the users usually controls the vacuum with an app to indicate the room layout of their home, which room it should clean, and during which days. If the threshold between rooms is too high, the robotic vacuum cleaner can experience some difficulties entering certain rooms, and the human user must manually lift it or place it in the correct room. When it is finished, or need to charge, the vacuum should ideally go back to its designated spot at the charging station.

1.3 Universal Design and AI systems

Universal Design and inclusion

Universal Design is the notion that everything should be equally usable and accessible no matter your age, situation, or whether you have a disability (uutilsynet, n.d.). This explanation describes how all services should be made to accommodate everyone, not matter what their situation is. Universal Design is an important concept, and by understanding how crucial it is regarding creating and maintaining services and solutions, will it benefit the society as a whole. Universal Design is about inclusion, even though it should be a standard mindset when developing. Everybody deserves equal opportunities no matter their situation.

AI systems and including "more" users

AI has potential regarding aiding people with different challenges or disabilities in their everyday life. With new technology such as voice recognition, text to speech, eye tracking and more, the user group of different technologies becomes broader. An example is Apple´s newest Siri remote, which gives you the possibility to use the buttons on the remote as originally intended, or to speak with Siri giving voice commands instead of touching the remote. Text to speech programs is also a great way of making written sources available for people that are blind, or generally people with problems connected to eyesight.

While AI has the potential for including people, it also has the potential of excluding people. Regarding how AI learn, people are being reduced to data and numbers, and not treated as individuals which will lead to systems disregarding more complex and personal circumstances (Hagendorff, 2020, p. 110). Multiple parking houses has implemented a system that automatically scans your numberplate when entering, and then register the number again when you drive out. The thought behind this might be to make the parking experience smoother and reduce the need of the user to pay on machines. However, to register properly within these facilities, you need to either download an app or use a new machine where you register your plate number. This new way of registration may exclude some user who potentially does not have a smartphone, or does not know how these new machines work, nevertheless being forced to use them when the old system is no longer in use.

WCAG 2.1, Human AI-interaction and the concept of "understanding"

The WCAG 2.1 principles includes the word "understandable". This concept is described as "Information and the operation of user interface must be understandable." (W3C, 2018). This definition of the word refers to how the user must be able to understand and interact with a solution. When discussing if machines "understand", I believe they do understand within their own limits and data. However, as Bratteteig and Verne writes *"Human cognition and actions take place in a context and cannot be represented by algorithms or in a machine."* (2018, p. 2). There might be social cues or restrictions that a machine does not comprehend in the same way.

1.4 Guideline for Human-AI interaction

Microsoft´s 13th guideline: *"Learn from user behavior"* (Amershi et al., 2019a) can be described by how Word alters their spellcheck and grammar system to adapt to the user. When the user often misspells the same word and accepts Words suggestion to what the user intended to write, it will begin to automatically change the word.

Jakob Nielsen´s HCI heuristics describes 10 guidelines for Human-Computer interaction design that will help to make a solutions or system user-friendly (2020). Microsoft´s guidelines for human-AI interaction and Nielsens HCI guidelines shares many similarities in the sense that the main goal is for the user to have a good and clear experience when interaction with the system. However, Nielsen´s 5th guideline "Error prevention" focuses on how the solution should avoid having flaws or errors, although IF an error occurs, the message describing it should be concise and clear (Nielsen, 2020). Microsoft´s 2nd guideline: *"Make clear how well the system can do what it can do"* (Amershi et al., 2019a) concentrate on making it clear to the user the limitations of the solution, and which errors might occur during the interaction. While the HCI guidelines strive for as little flaws as possible, the Human-AI interaction guidelines normalizes that there will be weaknesses in the system and focuses on informing the user of this early on.

Second module – iteration 2:

2.1 Characteristics of AI – infused systems

Identify and describe key characteristics of AI-infused systems

Asbjørn Følstad identifies four main characteristics of AI-infused systems in his lecture. These are learning, improving, black box and that they are fueled by larger datasets (2022).

Learning is a key characteristic when it comes to AI-infused systems. They continuously learn throughout being interacted with, and the system adapts its behavior accordingly. Amershi et al., (2019b) describes how many AI-infused systems are inconsistent since they continuously change by learning over time. This might result in users abandoning the systems over time, due to inconsistency and uncertainty.

Improving and learning can share similarities when it comes to the system adjusting and adapting to new information, but it also refers to how mistakes are inevitable, and the system will continuously improve their performance based on feedback and new data. Kocielnik et al., (2019) exemplifies improvement through an AI email scheduling assistant that is supposed to detect meetings and schedule this into the user's calendar. Initially, users are dissatisfied with the system making errors, but when the system is altered and improved to avoid making false positive errors, the users are more satisfied and more accepting.

Black box refers to how it is difficult to know and understand what exactly the AI-infused system does to generate the kind of output it gives (Følstad, 2022). Yang et al., (2020) describes how there are "challenges in understanding AI capabilities – designers frequently reports that it is difficult to grasp what AI can and cannot do." (p. 2). They further explain how in general, the AI capabilities often are uncertain, and how the boundaries of AI are continuously evolving (Yang et al., 2020). Amershi et al., (2019b) uses the example of automated filtering of information which can be a very nice feature, but at the same time can result in important or specific information being hidden from the users without their knowledge.

Fueled by larger datasets (data gathered through interaction) is a characteristic that explain how the AI-infused systems learn and improve, and that they are designed to use the data gathered when the system is being used to alter and evolve its behavior and understanding. "What might seem like a blue-sky AI design idea may suddenly become possible because of a newly available dataset." (Yang et al., 2020, p.4) This indicates AI-systems as continuously improving, learning, and changing.

Identify one AI-infused system which you know well, that exemplifies some of the above key characteristics. Discuss the implications of these characteristics for the example system, in particular how users are affected by these characteristics.

Instagram, which is a social platform where users share pictures and videos, is an example of an AI-infused system that I personally know well. Instagram has a feature which is calls "stories" where the users can post additional pictures and videos that will disappear after 24 hours if nothing else is determined. These "stories" is always displayed at the top of the app, while the regular pictures and videos that are posted and are not meant to disappear is

presented in the regular user feed and at the user's profiles. An example of how this system is ***learning*** is the order it displays the stories. It changes the order they appear after which stories you usually view and displays the most interesting ones first. Further, it has a "discover" page where it displays post from users you do not follow, but that they think you might enjoy based on your previous interactions in the app. I believe this site is relevant for the characteristic ***fueled by larger datasets*** since it alters what it displays based on information continuously gatherer by the user's behavior. These two examples might indicate how the users are affected in a way where Instagram makes decision as to what the user sees when interacting with the app.

2.2 Human-AI interaction design

"Will You Accept an Imperfect AI? Exploring Design for Adjusting End-user Expectations of AI Systems" – Kocielnik et al., 2019

As the title suggests, Kocielnik et al., (2019) presents in their article different methods to manage expectations from the users of a system, and in what capacity this might alter how satisfied or dissatisfied they are. They primarily focus on what kind of error the system should try to avoid, either False Positives or False Negatives, and how this might greatly affect the acceptance of an AI-infused system. As previously mentioned, the authors examine this through a Scheduling Assistant that is supposed to automatically detect meetings in the user's emails and schedule these appointments. Their findings show that the acceptance rate of False Positives is much higher than for False Negatives. The authors suggests that this might be that it is easier to recover from scheduling a meeting that is not supposed to be there, rather than missing a meeting that one was supposed to attend. They designed three adjustment techniques to adjust user's expectations: 1) Accuracy Indicator, 2) Example-based Explanation, 3) Control Slider (p. 5), which all proved to increase the user's satisfaction by adjusting the expectations with the AI-infused system.

"Guidelines for Human-AI Interaction" – Amershi et al., 2019b

Amershi et al., (2019b) discusses in their article the growing need for innovation and studies regarding human-AI interaction. They present 18 guidelines for design that have been examined through multiple evaluations and a user study. The aim with these guidelines is that they "can help with the design and evaluation of AI-infused systems that people can

understand, trust, and can engage with effectively." (p. 2) When examining the guidelines, the authors focused on whether the guidelines were relevant and if they were clear. After multiple iterations of the guidelines, they ended up with the 18 that are presented in the paper. Although they tried making a general set of guidelines, they do specify that they might not be applicable for all AI-infused systems. Overall, the aim of the guidelines is to provide help av guidance regarding human-AI design, and further the discussion and research regarding the subject.

Design guidelines and Instagram

To discuss Instagram, as previously described in the paper, regarding the guidelines presented in the paper by Amershi et al., (2019b), I have selected guideline G8 - Support efficient dismissal and G13 - Learn from user behavior (p. 3). These two are selected as I believe they represent how Instagram both adheres to and deviates from the design guidelines.

G8 - Support efficient dismissal: When you have reached the end of your feed and you have viewed all the pictures posted from the people you follow, Instagram automatically continuous the feed with "suggested posts". They do give you the opportunity to click "view older posts", but this is not continued in the general feed. Instead, you are directed to a new page. When clicking back, you are directed to the general feed with the suggested posts. I believe that supporting efficient dismissal, as the guideline say, would result in the older posts being presented directly in the users general feed, dismissing the suggested posts.

G13 – Learn from user behavior: This guideline's description is "Personalize the user's experience by learning from their actions over time." (Amershi et al., 2019b, p. 3) As described earlier in the paper, Instagram's ability to present stories that previously have been the most interesting for the user, as well as present the pictures and videos in the main feed in order after the people the user is most commonly interacting with, I believe adheres to guideline 13.

"On the Dangers of Stochastic Parrots: Can Language Models be Too big?" – Bender et al., 2021.

In their article, Bender et al., (2021) takes a critical viewpoint of AI-infused systems based on language models, specifically larger ones. They investigate when these language models become too big, and what risks appear alongside them. Environmental and financial concerns are raised, and how already marginalized communities will take the majority of the negative

consequences regarding the development using language models. Further, they discuss how the language models where the input is text from the internet consists of overrepresented and dominating viewpoints which again can cause substantial harm to marginalized communities.

2.3 Chatbots/conversational user interfaces

Key challenges in the design of chatbots/conversational user interfaces

Følstad & Brandtzaeg defines in their article chatbots as "machine agents serving as natural language user interfaces to data and service providers, typically in the context of messaging applications." (2017, p. 38). Luger & Sellen (2016) defines conversational agents (CA) as "the emergent form of dialogue system that is becoming increasingly embedded in personal technologies and devices." (p. 5287) They distinguish CA from chatbots which they present as something that does not possess any knowledge or memory, but rather imitate conversations. However, both chatbots and CA's face large design questions regarding user satisfaction and acceptance. Designing for chatbots and natural language user interfaces presents new challenges for the HCI field, as it shifts the focus from interaction and layout to designing conversations (Følstad & Brandtzaeg, 2017). Focusing more on what the user needs and how to achieve this, rather than explaining what features are possible to interact with (Følsdtad & Brandtzaeg, 2017). An extended focus on guidelines 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., 2019b) might help resolving some of the issues regarding chatbot and CA user interfaces. In their study, Luger & Sellen (2016) presented that the users in general did not understand the systems capabilities, limiting their experience with CA. An even more explicit definition of what the system has the possibility to achieve, and in what capacity, might result in higher user satisfaction and understanding.

References:

Amershi, S., Vorvoreanu, M., & Horvitz, E. (2019a, 1. February). *Guidelines for Human-AI interaction design.* Microsoft. https://www.microsoft.com/en-us/research/blog/guidelines-for-human-ai-interaction-design/

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. (2019b). Guidelines for Human-AI Interaction. *Conference on Human Factors in Computing Systems*, 1-13. https://doi.org/10.1145/3290605.3300233(CHI '19)

Apple. (2022a). Retrieved September 12, 2022. https://www.apple.com/

Apple. (2022b). *Machine Learning and AI. The work is innovative. The experience is magic.* Apple. Retrieved September 12, 2022. https://www.apple.com/careers/us/machine-learning-and-ai.html

Bender, E. M., Gebru, T., McMillan-Major, A., Shmitchell, S. (2021). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? In *Conference on Fairness, Accountability, and Transparency (FAccT'21), March 3-10, 2021, Virtual Event, Canada.* ACM, New York, USA, 610-623. https://doi.org/10.1145/3442188.3445922

Bratteteig, T. & Verne, G (2018). Does AI make PD obsolete? Exploring challenges from Artificial Intelligence to Participatory Design. https://doi.org/10.1145/3210604.3210646

Følstad, A. (2022). Module 2: Design of Interaction with AI (I). https://www.uio.no/studier/emner/matnat/ifi/IN5620/h22/interacting-with-ai-2022---module-2---session-1---handout.pdf

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

Grudin, J. (2009). AI and HCI: Two fields divided by a common focus. *The AI*

Magazine, *30*(4), 48–57. https://doi.org/10.1609/aimag.v30i4.2271

Hagendorff, T. (2020). The Ethics of AI Ethics: An Evaluation of Guidelines. *Minds & Machines* 30, 99–120. https://doi.org/10.1007/s11023-020-09517-8

Holmquist, L. (2017). Intelligence on tap. *Interactions (New York, N.Y.)*, *24*(4), 28–33. https://doi.org/10.1145/3085571

IBM Cloud Education. (2020, 03. June). *Artificial Intelligence (AI).* IBM. https://www.ibm.com/cloud/learn/what-is-artificial-intelligence

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 *CHI Conference on Human Factors in Computing Systems Proceedings (CHI 2019), May 4-9, 2019, Glasgow, Scotland Uk.* ACM, New York, NY, USA, 1-14. https://doi.org/10.1145/3290605.3300641

Kommunal- og distriksdepartementet. (n.d.). *Nasjonal Strategi for Kunstig Intelligens.* Regjeringen. Retrieved September 13, 2022. https://www.regjeringen.no/no/dokumenter/nasjonal-strategi-for-kunstig-intelligens/id2685594/?ch=3

Luger, E., Sellen, A. (2016). Like Having a Really bad PA: the Gulf between User Expectation and Experience of Conversational Agents. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, 5286-5297. https://doi.org/10.1145/2858036.2858288

Merriam-Webster. (n.d.) *Robot*. In Merriam-Webster dictionary. Retrieved September 12, 2022. https://www.merriam-webster.com/dictionary/robot

Nielsen, J. (2020, 15. November). *10 Usability Heuristics for User Interface Design.* nngroup. https://www.nngroup.com/articles/ten-usability-heuristics/

Søraa, R. A. (2022, 12. May). *Robot i Store norske leksikon.* snl. Retreived September 12,

2022. 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/10.1080/07370024.2004.9667338

Uutilsynet. (n.d.). *Kvifor universell utforming av IKT?* Uutilsynet. https://www.uutilsynet.no/veiledning/kvifor-universell-utforming-av-ikt/240

W3C. (2018, 05. June). *Introduction to understanding WCAG.* W3. https://www.w3.org/WAI/WCAG21/Understanding/intro#understanding-the-four-principles-of-accessibility

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 2019 CHI Conference on Human Factors in Computing Systems*. ACM, 1-13. https://doi.org/10.1145/3313831.3376301

Appendix:

The feedback I received for Module 1 was in general good. The “one wish” that I got was that the person reading my text missed a broader explanation regarding the definition of “robot” beyond just defining what it is. To accomplish this, I wrote a bit more surrounding Thrun’s perspective on robots.