IN5480 - Individual assignment fall 2021

Martine Leine Rafteseth

3. Iteration

Table of contents

1.1 Concepts, definition, history and interaction with AI

1.1.1 History of AI

According to Grudin (2009) the history of Artificial intelligence goes way back, starting with the established potential of computation during the second world war. After the war, Governments funded the building of computers at some universities. Not far after, Alan Turning was one of the first to comment on Artificial Intelligence. Turning was a British leading code breaker, mathematician and logician, and his statements in the London Times in 1956 created a sensation: *"I do not see why [the computer] should not enter any one of the fields normally covered by the human intellect, and eventually compete on equal terms. [...]"*

The actual term *Artificial intelligence* was first used at a workshop in 1956 by John McCarthy, an American mathematician and logician. Through the following decades, the funding, interest, and development of AI had its ups and downs. Today, ordinary computers can utilize AI and machine learning through browser access to servers, and the development of AI goes on (Grudin, 2009).

1.1.2 Definitions

The first definition is by Russell & Norvig, two researchers within Participatory design. It focuses on machines being able to perform tasks with intelligence similar to a human, rather than emphasizing technical perspectives and abilities.

"A subfield of computer science aimed at specifying and making computer systems that mimic human intelligence or express rational behaviour, in the sense that the task would require human intelligence if executed by a human" by Russell & Norvig in 2010 (Bratteteig & Verne, 2018, p. 1-2).

The second definition is also focused on the 'intelligence' of machines within the field. Most importantly this intelligence must adapt to the environment and different situations in order to function properly.

"Artificial intelligence is that activity devoted to making machines intelligent, and intelligence is that quality that enables an entity to function appropriately and with foresight in its environment" by *Stone et.al. (2016).*

For the last definition, which is the most recent one, the authors from the High-Level expert group on Artificial Intelligence (2019) are elaborating a bit more about where and how systems with artificial intelligence are and can be of use:

“Artificial intelligence (AI) refers to systems that display intelligent behavior by analyzing their environment and taking actions – with some degree of autonomy – to achieve specific goals. AI-based systems can be purely software-based, acting in the virtual world (e.g. voice assistants, image analysis software, search engines, speech and face recognition systems) or AI can be embedded in hardware devices (e.g. advanced robots, autonomous cars, drones or Internet of Things applications)”

My definition of Artificial Intelligence is mostly inspired by the one from the High-Level expert group on Artificial Intelligence (2019), because it was less abstract than the other, yet not as assertive, taking the different ‘forms’ of AI into consideration:

"Artificial Intelligence is a technological approach making machines analyze and adapt to environments and situations, take rational decisions and perform actions that normally require human intelligence."

1.1.3 Article review

For the article review, I chose the article by Dautenhahn (2018) concerning thoughts on the past and future of human-robot interaction. I found the article interesting because it addressed both the past and the future of HRI. The author expressed concern about how much robots could develop in 50 years’ time, taking the past development into consideration. To get to the point of having super-intelligent robots there is a lot of long-term and time-consuming research that needs to be done in ecologically valid environments. Dautenhahn addresses such work as anything but easy, yet necessary if we want HRI to have implications and social impact in the real world outside the labs.

1.1.4 A contemporary company that work with AI

For the contemporary company that uses AI I chose Apple. According to an article by Mark Allison (2021) Apple has implemented AI functions such as face, handwriting and song recognition, translation, sleep tracking and more into their software. On Apple's webpage they

say that AI and machine learning are building amazing experiences into every Apple product, as AI allows people to do what they never imagined. Apple is using AI while still focusing on improving user experiences and protecting user data (Apple, 2021).

1.1.5 A film that addresses the use of AI and interaction between people and AI

I watched the movie “Do you trust this computer?”, which discussed how the AI-industry has developed and will continue to develop further, reshaping every aspect of our lives. In the movie, self-driving cars, human replicates reading emotions, atom bombs and AI machines for surgery were some of the AI-technologies shown. A lot of positive sides of AI were mentioned, but the main characters expressed fear that machines with AI somewhere in the future could outsmart us. Self-learning machines have and will replace a lot of jobs, leading to unemployment. One of the questions I found the most interesting was expressed by one of the characters. He said that technology will always contain parts of us, since we created it - the question is just if I will be the good or the bad parts.

1.2 Robots and AI systems

1.2.1 How the term “robot” came about

The term "Robot" has its roots back to Prague in 1920, where it was introduced in a Czech play by Karel Capek called R.U.R (*Rossum’s Universal Robots).* The word itself derived from the Czech word "robota" which referred to a system of forced self labour (Love, D., 2020).

1.2.2 Definitions

According to Thrun (2004) Robotics is a broad term. He presents two definitions of a robot, where the first definition from The Robot Institute of America in 1979 is as follows:

“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.”

The second definition is from Webster dictionary, where a robot is defined as “An automatic device that performs functions normally ascribed to humans or a machine in the form of a human.”

The definitions differ in their focus. Where the first is concerned with Robots being manipulated to move certain things through programmed motions, the second focuses on it performing tasks as a human. I liked both definitions, but wanted to include that robots may have different levels of automation and that I consider them something physical. My definition of a Robot is:

"A physical machine that is programmed to perform a series of given tasks automatically. It can be fully autonomous or semi-autonomously, and may be built to replicate human behaviour."

1.2.3 The relation between AI and Robots

Some of the different definitions above describing AI and robots are using similar terms, which is natural as we today can see robots with AI together with expanding technological opportunities. Even though the fields may be connected through devices capturing both terms, there are some clear differences between the two.

In my opinion, they are different in the way that AI is not necessarily a physical object, but still needs to have some kind of "intelligence" making it able to adapt to the environment and perform rational actions. This intelligence is most commonly approaching human intelligence. A robot on the other hand *is* a physical object, and the tasks performed can be pre programmed without it being adaptive, intelligent or resembling human behaviour.

1.2.4 A contemporary physical robot

For this task I would like to describe the robot vacuum cleaner, which is more and more common to see in households today, saving people from the tedious task of vacuuming and cleaning. The robot is a lot smaller than a traditional vacuum cleaner, allowing it to easily clean under different types of furniture and stay located at a central charging station. Humans can easily interact with the robot and set it to clean at convenient times. The vacuum cleaner can also be moved between floors, and will adapt by itself to different obstacles.

1.3 Universal Design and AI systems

1.3.1 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."

The definition above by the National disability authority describes Universal design as the design and composition of an environment, where I understand the environment as a broad term also including products, services, or buildings within that environment. The main factor of universal design is that it should benefit all, by being accessible, usable, convenient, and pleasurable for people with different "needs". This way, universal design is *good design*.

1.3.2 The potential of AI

AI has great potential in respect to human perception, as well as emotion/cognition and movement, because AI has the ability to recognize and take action based on different situations. In the movie "Do you trust this computer?" AI was used on human replicated robots, where the AI responded to the humans talking to it based on their mood. Much research has been done with AI in robot pets (and robots) for instance to prevent elderly people from feeling lonely. Here, AI can be used to guide people feeling lost, as well as cheer them up. Also, within education, AI has great potential to increase the learning of people with disabilities. This kind of AI includes people with all needs, but there are also examples of AIs who exclude people, where for instance face-ID has been disputed.

1.3.3 Do machines understand?

I make sense of the words "understand" and "understanding" as being able to make sense of given information. When it comes to machines being able to understand, I believe that they can understand logical information they receive, because they are programmed to, but contrary to humans they are not able to understand certain empathic perspectives.

1.4 Guidelines for Human-AI interaction

1.4.1 Microsoft’s guidelines for human-AI interaction

I chose the fourth principle: “Show contextually relevant information. Display information relevant to the user’s current task and environment.”

This principle is based on the AI providing the user with relevant information to the user's environment and task. I also understand this principle as the AI being able to find relevant information where this is not necessarily natural. For instance, if a person is searching on Netflix for a movie, the AI should be able to suggest similar movies if the movie searched for is not available.

1.4.2 HCI principles vs Microsoft’s guidelines

For the set of HCI guidelines I chose Donald Norman’s six design principles: visibility, feedback, affordance, mapping, constraints, and consistency (Norman, 2013).

I believe that Microsoft’s guidelines for AI are divided into phases making them easier to apply directly, whereas Norman's design principles are more abstract. Yet, they are both similar in their use of principles such as *feedback and* visibility, which is essential regarding user-friendly interactions.

2.1 Characteristics of AI-infused Systems

2.1.1 Key characteristics

In the lecture of the second module of the course, we were presented with four key characteristics of AI systems.

Learning entails that the AI adapts dynamically through interacting with people. This way the system "changes via learning over time" (Amershi et al., 2019, p. 2), and allows for personalized user experiences.

Improving refers to AI's ability to gradually learn from user input, feedback, interaction and mistakes and thereby becoming a bit more accurate and intelligent each time.

Black Box is a term used to address the lack of insight to the AI-systems handling of given information and interaction. According to Kocelnik et al., "providing explanations will lead to higher perceptions of understanding how the AI system works" (Kocielnik, 2019, p. 4).

Fuelled by large data sets refers to the knowledge and data behind the AI system. The data is given to the AI by the designer of the system or/and collected from user input and interaction (Amershi, et al, 2019).

2.1.2 Zalando as an AI-infused system

The Zalando website is an example of an AI-infused system. Through the past years with the growth of online shopping, website functionalities have emerged. AI has been so smoothly implemented that the users now might perceive the functionality that derives from it as an obvious part of the system. At Zalando.com, when spelling a word wrong in the search bar, the website often still shows the wished products as it has learned from former mistakes and interaction.

The most prominent AI characteristic is the Ais personalized recommendations for the users. "Based on your favourites, maybe you will also like", "similar products" and "Recommended for you, we think you will like" are examples of how the AI personalizes the interaction, and at the same time reveals what the tips are based on. If you ask me, I would say that all these functions have a positive effect on the overall user experience.

2.2 Human-AI interaction

2.2.1 Main takeaways from Amershi et al. (2019) and Kocielnik et al. (2019)

Amershi et al. (2019) look at the design of user interfaces and discuss how AI can pose new challenges and opportunities. The article proposes 18 design guidelines for interaction between human and AI, categorized by when the interaction takes place: “initially”, “during interaction”, “when wrong” and “over time”. The guidelines were synthesized by guidance over the past 20 years about the design of human-AI interaction. Through four iterations, researchers, designers and usability practitioners tested, refined and reviewed the guidelines.

The article by Kocielnik et al. (2019) focuses on user expectations of AI systems and *(how this affects the overall experience of the interaction.)* acceptance of an imperfect AI. New functionalities have emerged with AI-infused systems, affecting the user experience and creating new expectations. Examples of such functionalities are natural language understanding, object recognition, web behavior prediction and sensor-based inference. Kocielnik et al. (2019) describe these functionalities as probabilistic, as they almost always operate at less than perfect accuracy. Considering the effect user expectations have on the user experience, the article explores three techniques that help form the user expectations towards AI-systems and discuss how this affects the user acceptance.

2.2.2 Discussing Zalando and two guidelines from Amershi et al. (2019)

G15 – Encourage granular feedback.

Zalando gives the user the opportunity to favorite items. This way the users can save the items for later purchase or gather different items and compare them before purchase. By favoriting the users can indicate their preference while interacting with the system.

G11 Make clear why the system did what it did.

Zalando gives the user recommendations based on recent search and the user’s favorites. The system tells the user what the recommendations are based on, and this way makes it clear why it did what it did.

2.2.3 Summarizing arguments inn Bender et al. (2021)

The article puts a critical view on large language models and outlines several concerns around a variety of risks and costs. They first and foremost express concerns related to the environmental impact and financial cost. A main argument is that language models punish marginalized communities, who are the most likely to be harmed by environmental consequences caused by the model's consumption. At the same time the ones affected are the least likely to benefit from the technology.

Just as the impact on the environment grows with the size of the LM's, so does the struggle to understand the data given to the systems (Bender et al., p. 610). With very large models it is hard to filter, categorize and process the amount of data, which often can result in bias, discrimination and errors in the text. Bender (2001) therefore argues that language models of a large size do not equal diversity, and suggests solutions involving careful data collection practices, financial and environmental costs consideration and documentation (Bender et al., 2001, p. 617-618).

2.3 Chatbots/conversational user interfaces

2.3.1 Key challenges in the design of chatbots / conversational user interfaces

One of the key challenges when designing chatbots is that they are relatively new to the HCI field, and new approaches are required in the design compared to traditional interaction design (Følstad & Brandtzaeg, 2017, p.40).When designing for usability within traditional interaction design, visual layouts in the shape of graphical interfaces and visual interaction mechanisms are in focus. However, since chatbots as a design object is conversational itself, Følstad & Brandtzaeg (2017) argues that we need to shift the focus from designing an object to designing a service (Følstad & Brandtzaeg, 2017, p.41). This means refocusing to understand the users, their needs and help them reach their goals. Users will ask questions in different ways, leaving one of the key challenges in the design of chatbots to make sure that it answers appropriately to the user input (Følstad & Brandtzaeg, 2017, p. 42). This will also be affected by the accuracy and calculations of the algorithms and the size of the dataset behind the chatbot.

2.3.2 Key challenges in chatbot design and guideline G1 and G2 in Amershi et al. (2019)

Guideline 1 - *Make clear what the system can do*

The first guideline in Amershi et al. (2019) should be applied initially in the interaction between a user and a chatbot, to help the user understand both what the chatbot can and can't help with, thus making the service more efficient. This could for instance be applied by explaining to the user what the chatbot can do or/and give suggestions for what to ask for, making it less likely that users perceive the chatbot as useless and reject it.

Guideline 2 - *Make clear how well the system can do what it can do*

The second guideline depends on the complexity of the chatbot, making it slightly more difficult to implement. In this guideline I believe the most important thing is to try and adjust the users' expectations towards the chatbot, as discussed in the article by Kocielnik et al (2019). This can be done by trying to clarify how complex the robot's interaction skills are and the accuracy of the answers. The chatbot may tell the user that it is "still learning" and be humble in the way it talks to the user. This way, being less determined in the conversation, the users may be more forgiven if the AI makes mistakes.

References

Allison, M., 2021. *The Use of Artificial Intelligence Technology in Apple Devices.* [Internett]

Available at: https://roboticsandautomationnews.com/2021/06/14/the-use-of-artificial-intelligence-technology-in-apple-devices/43866/

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.

Apple, 2021. *apple.* [Internett]
Available at: https://www.apple.com/careers/us/machine-learning-and-ai.html

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.

Dautenhahn, K., 2018. Some Brief Thoughts on the Past and Future of Human-Robot Interaction. ACM Trans. Hum.-Robot Interact. 7, 4:1–4:3.

Følstad, A., & Brandtzæg, P. B. (2017). Chatbots and the new world of HCI. interactions, 24(4), 38-42.

High-level Expert Group on Artificial Intelligence. "A definition of AI: Main Capabilities and Disciplines" (April 8, 2019). European Commission. Available from URL: https://itlaw.wikia.org/wiki/A_Definition_of_AI:_Main_Capabilities_and_Disciplines

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.

Love, D., 2020. *Where does the word 'robot' come from?.* [Internett] Available at: https://www.sciencefocus.com/future-technology/where-does-the-word-robot-come-from/ [Funnet 3 september 2021].

National disability authority, u.d. *What is Universal Design.* [Internett]
Available at: http://universaldesign.ie/What-is-Universal-Design/
[Funnet 2 september 2021].

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

Appendix

Feedback and implementations from the first iteration

After the first iteration I received feedback in the form of “two stars and a wish”. The “wish” was that I elaborated a bit more about how I came to my definitions of “AI” and “robot”. Hence I added more details to them before the second module.

Feedback and implementations from the second iteration

For the second iteration I recived a wish to sort out a few unfortunate punctuation errors in a section. I fixed that and also scanned through both iterations for the final delivery.