IN5480

Contents

Iteration 1

Concepts, definition and history of interaction with AI

First, write a section about how AI came about, the history of AI!. When, and by whom, was the term first used?

The term AI, or artificial intelligence, was first used in 1956 by John McCarthy, an American logician and mathematician (Grundin: 2009, p. 49). McCarthy used the word to describe machine simulation of learning and other characteristics of human intelligence. The world's first artificially intelligent program is called "Logic Theorist" and was written the same year (Bosch: 2018).

Then, find three different definitions of AI. Describe and explain these three definitions, for example by when it was defined, by whom and in what community. Based on these three definitions, make one definition yourself - and describe and explain your definition.

The first definition of AI that appears on Google when I search for "artificial intelligence" is *"The simulation of human intelligence processes by machines, especially computer systems. These processes include learning, reasoning and self-correction" (Rouse: 2018).*

The Oxford dictionary defines AI as *"the theory and development of computer systems able to perform tasks normally requiring human intelligence, such as visual perception, speech recognition, decision-making, and translation between languages (Lexico: 2019).*

Wikipedia says that AI often is *"used to describe machines (or computers) that mimic "cognitive" functions that humans associate with the human mind, such as "learning" and "problem solving"", which is a definition that was used by Stuart Russel and Peter Norvig (both computer scientists) in "Artificial Intelligence: A Modern Approach" from 2009* (Wikipedia: 2019).

Rouse defines AI as a type of simulation, the Oxford dictionary defines AI as a theoretical and practical framework behind it, and Russel and Norvig defines AI as a description of a certain type of machines. Both Rouse's and Russel and Norvig's definitions are from around 2009, while the date of Oxford dictionary's definition is hard to pinpoint due to the nature of the website. The two other definition both have named authors, Russel and Norvig being a

logician and mathematician and Rouse being a writer and manager of the online IT encyclopedia WhatIs.com.

My definition of AI is that "AI is the simulation of human cognitive processes in machines". I believe that AI is not truly intelligent, but is able to mimic certain processes by being made in a certain way. I do not believe that you can create something equal to human intelligence, but that algorithms and machine learning may create a something that may resemble it. The machine's intelligence is limited by the developer's choices when developing it, so the intelligence will be subpar to human intelligence.

Find one contemporary company that work with AI and describe how this company present AI on their web pages. In what way does this company talk about AI, as a product, as a service, framework or "idea"?

I wanted to find a Norwegian company that work with AI, so I decided to check the report "Artificial Intelligence in Europe: Norway. Outlook for 2019 and Beyond" commissioned by Microsoft, but only one of the six Norwegian companies mentioned had information about AI on their webpage: Telenor. Telenor calls AI "the most important technology in the 21st century" and predict that AI will be the momentum of all processes and services within the next 20 years. They talk about the effect AI will have on the financial development and stress the importance strengthening Norway's position by educating more people on AI.

Select one documentary or a fictional film, book or game that is about the use of AI systems. Describe with your own word how human interaction with AI is portrayed in this work

In the HBO and BBC series *Years and Years* by Russel T Davies we follow the fictional Lyons family who all consumers of AI technology. The story is set to the period 2019 to 2034, and the technology in the fictive 2019 is about as advanced as it is in the real 2019. All adults in the series (elderly too) have a personal assistant (not unlike Siri, Alexa, etc.), which they use to call each other, note down plans and organize their everyday lives. This type of artificial intelligence is portrayed as something ordinary and common, as it seems like most, if not all, either owns one or knows how to use one. There are several other examples of technologies that are portrayed as niche and uncommon, but personal assistant AIs are not among them.

Towards the end of the series (without spoiling too much), one of the characters are getting their memories downloaded to a new storage technology to become a fully digital human intelligence. In this way, the show blurs the lines between human intelligence and artificial intelligence. In the very last scene of season one, the character's consciousness is supposed to "inhabit" an old personal assistant hardware. In that way, the person can live on as a part of all their lives. The characters debate whether the character really will be conscious, or if the downloaded version only will be a shadow of who the person really was. The issue of whether emotions can or have to be uploaded to fully realize a human intelligence as digital is brought up, but we don't get the answer to that as the show ends before the digital human intelligence is supposed to answer whether or not it is, in fact, there. The show ends on an open note, which leaves the seers to ponder the future of artificial intelligence.

Human-Robot Interaction

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

The word "robot" is originally Czech (*robota*) and means "forced labour". It was the Czech playwriter K. Čapek who coined it as early as 1920 when we called his play "Rossum's Universal Robots" (R.U.R.) (Lexico, 2019). In the play, "robot" is used to describe humanlike machines made by the fictional scientist Rossum. Later in the play, the robots are made more human by a second scientist, and the robots become capable of feeling pain. The robots become more and more humanlike, and in the end the robots come to dominate humans (Kuiper, 2019).

Then, find two different definitions of robot. Describe and explain these definitions. Based on this definitions, make one definiton yourself, and describe and explain this definition.

Oxford Dictionary defines the word robot as "A machine capable of carrying out a complex series of actions automatically, especially one programmable by a computer" (Lexico: 2019). This definition is also used by Wikipedia.

The Robot Institute of America, as mentioned in Sebastian Thrun's paper on HRI, defines a robot 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" (Thrun, 2004).

As mentioned in the task about AI quotes, the dates of definitions by the Oxford dictionary are hard to pinpoint, but the other definition from as early as 1995.

My definition of a robot is that “a robot is a machine able to do complex tasks”. It is fairly simple, but I think that it is broad enough to cover all robots. It might be too broad, so that it includes other technologies other than robots, but I am of the impression that multiple of the definitions over are too, so I stand by my broad definition.

Discuss the relation between AI and Robots. Is “a robot” different from “an AI”? In what ways are they different and similar? Bring in the definitions that you described earlier about robots and AI for this discussion.

While AI is about simulating human intelligence, robots are more about doing complex tasks. AIs may be capable of doing complex tasks, and robots may simulate intelligence, but it is not given that they do that. An AI is often a software able to understand, reason or learn, while robots are physical artefacts able to move around either globally or locally.

Find one contemporary physical robot, either described in a research article - or a commercial robot, and describe how this robot moves and how a human user is interacting and using the robot in a specific situation.

The Boston Dynamics robot “Spot” is a semi-autonomous four-legged robot that can be controlled with a controller with a layout that resembles a hybrid between an Xbox controller and a Nintendo Switch. Spot can be steered in all directions, and due to the four cameras placed on either side of Spot, the operator can touch the screen to choose a waypoint that the robot should walk to (The Verge: 2019). Spot also has balance sensors to help with keeping it on all four legs while walking on uneven terrain. The robot resembles the robot from the episode “Metalhead” (episode 5) in season 4 of Black Mirror.

Universal Design and Interaction with AI

Please find and describe a definition of Universal Design. Explain this definition, how you understand what Universal Design is about with respect to inclusion.

The University of Washington state that “Universal design is the process of creating products that are accessible to people with a wide range of abilities, disabilities, and other characteristics”.

Here, the UoW indicates that universal design is not an end state, but the act of making something accessible. The definition implies that the end goal is not to have products that are “universally designed”, but to make products that are accessible. The distinction between the passive phrasing of “having” and the active phrasing of “making” indicates that accessibility is something we actively have to make sure is there, instead of expecting it to already be there.

Describe the potential of AI with respect to human perception, human movement and human cognition/emotions. You are encouraged to use examples.

AI is already being used to mimic human movement in so-called “deepfakes”. The deepfake technique is to use AI to make one person’s movement look like a different person’s movements. An example of this is the video “Friends ross Nicholas Cage faceswap piano” by YouTube user 9gag videos” where the faces of all the characters of the show “Friends” are switched out with the face of Nicholas Cage. While this can be seen as a trivial and fun thing, it has already been used to fake videos of political leaders. It is already becoming hard to differentiate between real and deepfake videos, so it is plausible that deepfakes can become a real issue in upcoming elections, as well as for people’s personal reputations. The technology can be used to do a lot of harm, such as making fake videos of war declarations, revenge pornography, false official statements, etc. While it is wonderful that technology keeps on evolving and new inventions do exist, it is important to question whether the technology should exist.

(9gag videos, YouTube).

Describe the potential of AI for including and excluding people. You are encouraged to use examples.

With the rising focus on voice-controlled AI, hearing-impaired users will be excluded from this handsfree technological experience. Those of us who are heavily hearing-impaired will rely on their eyes or touch to receive the output and touch (i.e. keyboard) or eyes (i.e. eye-tracking) to give input. This makes it harder to multitask when you i.e. are driving a car, making dinner, carrying a lot of grocery bags or in any other way are temporarily busy with your hands. On the other hand, speech-based AI allows for users with strong vision-impairments or missing limbs to interact with technology more freely compared to vision or touch-based technologies like computers, smartphones and smart watches.

There are other AI technologies that allow for use of smartphones while vision impaired. Object recognition apps (either separately downloaded or already integrated in the phone) allows users with reduced sight to “see” through their phone’s camera. The Huawei P30 Pro has an integrated app called HiVision in the camera application that can recognise QR-codes, translate text, scan regular objects and find related products for sale, recognize art pieces, and count calories in food. These apps can help both able-bodied and disabled users to access information without typing. This might be more efficient for some, but for users with rheumatic diseases and lower fine motor skills this might be harder to use. Luckily, most smartphones today support most of these artificial technologies, so each user can choose to download the applications that work for them.

Iteration 2

Characteristics of AI-infused systems

AI-infused systems are ' systems that have features harnessing AI capabilities that are directly exposed to the end user' (Amershi et al., 2019). Drawing on the first lecture of Module 2, **identify** and **describe** key characteristics of AI-infused systems. Also **read** Amershi et al. (2019) and Kocielnik et al. (2019) to possibly expand on this set of key characteristics

When we talk about AI, we differentiate between three distinct types or depths of AI: Artificial super intelligence, artificial general intelligence and artificial narrow intelligence. When talking about interaction with AI, we refer to interaction with artificial narrow intelligence. Artificial super intelligence refers to AI doing something beyond human capabilities, while artificial general intelligence refers to AI mimicking general human intelligence (like the AI robot Sophia (Hanson Robotics)). The systems that we are referring to when talking about AI-infused systems are therefore the systems infused with artificial narrow intelligence.

In the first lecture of module two in this course, Følstad mentioned these four key characteristics of AI-infused systems: Learning, improving, black box and fuelled by large data sets.

Learning refers to the system being dynamic and designed for change. When talking to a learning AI, tow identical messages or interactions will not give identical responses. This is because the AI systems “change via learning over time” (Ameshi, 2019: 2) through the interactions they have with people.

Improving refers to the AI systems’ ability to become better over time. When interacting with people, the AI learns from its mistakes which, as a result of this, makes it a bit more intelligent for each time. The fact that AI systems can improve also indicates that they are not perfect from the beginning, and mistakes will happen. Through feedback from users of the system, the AI will gradually learn more and more and become more accurate. The concepts of learning and improving are therefore closely related.

Black box refers to the view of AI systems as black boxes. The term black box refers to “a system or process that uses information to produce a particular set of results, but that works in a way that is secret or difficult to understand” (Cambridge University Press). When using the term to talk about AI systems, we often refer to the lack of understanding or insight of

what happens between the input is given and the output is presented. It is hard to understand how the AI presents the data it presents, and it is hard to validate it. It is therefore desirable to make the system less of a black box to design the system for more explainability. Kocielnik et al confirms this by providing support for their hypothesis that says passive AI systems that provide explanations "will lead to higher perceptions of understanding how the AI system works" (Kocielnik, 2019: 4).

Fuelled by large data sets refers to the input of AI systems. In order for AI systems to learn, they need a lot of data, and this data is collected from users. This can happen either actively by e.g. chatting with AI systems, or passively by sharing location data. Through this data, AI systems can learn and improve to become even better at their task.

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.

I choose the AI-infused keyboard I used to use on my old iPhone called Swift key. The keyboard allowed me to write words by swiping my finger between different letters instead of lifting it each time and individually press each letter. Usually, the keyboard would be able to recognize what words I was trying to write, but sometimes it did not. In those cases, I would have to type out the word manually, and the keyboard would ask me if I would like to add the word to its dictionary. This is an example of the system improving by asking for feedback on its suggestion. When I would try to swipe out that word the next time, the keyboard would have learned the word and suggest it.

The development of the keyboard must have required a lot of big data sets in order to be able to make suggestions. The keyboard supported multiple languages, and each language has up to several hundred thousand words. They keyboard had to be able to recognize and suggest these words for every supported language. Extreme accuracy was not needed in order to use the keyboard; it would suggest words based on what word "pattern" it resembled the most. By letting users actively give feedback to suggested words, I felt that I was contributing to making the keyboard smarter and better. While I do not know whether it changed on only a micro level vs. a macro level, it did improve my writing experience.

Human-AI interaction design

Amershi et al. (2019) and Kocielnik et al. (2019) **discuss** interaction design for AI-infused systems. **Summarize** main take-aways from the two papers.

Amershi at al. present 18 guidelines for human-AI interaction design and when to apply them. Their goal is for the guidelines to result in "better, more human-centric AI-infused systems" (Amershi, 2019: 12). Since the use and expansion of AI is ever increasing, they are of the impression that clear guidelines are significant for the field. Amershi et al. also stress the importance of further development and refining of these guidelines.

Kocielnik at al. "explore techniques for shaping end-user expectations of AI-powered technologies prior to use and study how that shaping impacts user acceptance of those technologies" (Kocielnik, 2019: 2). They also investigate the impact different types of AI imperfections have on these techniques and conclude that the techniques in fact do have an impact on key aspects of user expectations of AI-powered technologies. They present five hypotheses by which one is rejected, three are supported and one is partially supported.

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.

I have chosen two guidelines from the "over time" category of guidelines: G13 and G16. I argue that Swift keyboard adheres to both of these guidelines.

G13	**Learn from user behavior.** Personalize the user's experience by learning from their actions over time.

(Amershi et al, 2019: 3)

This guideline says that the system should learn over time, and in my experience the Swift keyboard did exactly that. Every time it did not know what I meant and I spelled out the correct word, the system learned that new words (if I wanted it to) and personalized my typing experience for each time.

G16	**Convey the consequences of user actions.** Immediately update or convey how user actions will impact future behaviors of the AI system.

(Amershi et al, 2019: 3)

This guideline is related to the guideline over. When I spelled out a word, the system asked me if I wanted to add the word to the dictionary. If I chose yes, the system would give me feedback telling me that the word was added. In this way, the system conveyed the consequence of my choice to add the word to its dictionary (which I assume might have been my personalized dictionary).

However, it has come to my attention that the Swift keyboard has changed somewhat since I downloaded and used it on my iPhone. I recently bought a Huawei smartphone, and a few days ago I realized that the phone has the Swift keyboard as default. When I type words now, the keyboard does not ask me whether I would like to add the word to the Swipe dictionary or not. Now, if I write the word enough times and/or click the word in the word suggestion after I have spelled it out completely, the keyboard will automatically suggest it for me. Personally, I like the new update because it makes the writing experience more seamless, but it does frustrate me in cases where I misspell the same word multiple times and the keyboard saves the misspelled version.

Chatbots / conversational user interfaces

Chatbots are one type of AI-infused systems. **Read** Følstad & Brandtzaeg (2017) and Luger & Sellen (2016) and **discuss** key challenges in the design of chatbots / conversational user interfaces.

One of the key challenges in the design of chatbots and conversational user interfaces as presented by Følstad and Brandtzaeg is combatting diversity among users and be open and inclusive technology. The chatbots need to be able to communicate naturally regardless of gender, age, language and preferences (Følstad, 2017: 4), but it is hard to prevent bias. It is especially hard to prevent bias surrounding language and tech knowledge. Users of younger age, users with learning disabilities, users with a different native language, and users with less tech knowledge than the assumed target user might have issues understanding the language used by the chatbot, making the interaction hard for them. When those of us with these kinds of challenges talk to a human, their conversation partner will be able to modify

their language according to the recipient's questions or choice of words. A chatbot, unless specifically programmed to simplify its language, will not do the same. To make chatbots that fit all, this will have to be addressed.

Luger and Sellen mentions that most of their participants have issues with the conversational agents' feedback and transparency. The participants had a hard time figuring out what the system could do or not, which lead to them either "feeling overwhelmed by the unknown potential, or led them to assume that the tasks they could accomplished were highly limited" (Luger, 2016: 5291). The design of the conversational agents does not make it easy to know exactly what they are capable of doing, and in multiple cases participants experience that the agents are not capable of doing what they want them to. Here, I am referencing to their experience with asking follow-up questions or related questions to their conversational agents. Luger and Sellen write that participants reported unsatisfactory results when trying to do so: "I don't ask for more information from it. It tends not to be very good at that. […] Asking it to do sub-tasks, to follow up or to give you more information about something you've just asked it, it tends to be really bad at." (Luger, 2016: 5289).

Revisit Guidelines G1 and G2 in Amershi et al. (2019). **Discuss** how adherence to these could possibly resolve some of the challenges in current chatbots / conversational user interfaces.

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

(Amershi et al, 2019: 3)

If the conversational agents mentioned in the task over adhered to these two guidelines, the majority of the participants would have had a much more positive experience with using them. If the system made clear what it can (and cannot) do, the participants would not have to use time to figure it out by themselves. When using a new technology, figuring out the extent of functionalities can be a very long process, no matter how tech-savvy you are. Personally, I learned something new about my phone's screenshotting abilities today. Since I am new to Android and Huawei, I did not know about a lot of the shortcuts you can do to access split screen, screen recordings and picture taking. I have only found out about them by accident,

and I imagine that there are tens of other shortcuts that I have yet to discover. This is true for conversational agents, as well as most other tech, as well. It is hard to find something if you do not even know what to look for.

As for the second guideline, the same example with conversational agents would benefit greatly from this too. The majority of the participants had trouble when trying to follow up on topics they had just talked with their conversational agents about, and the consensus was that the Cas did not really know how to do it. The fact that the majority had this experience tells us that most of the participants tried to follow up on something, and that this probably is a functionality they would like their CA to have. If the CAs do not have this seemingly highly demanded functionality, it should tell its users that it is not available to prevent frustration.

Iteration 3

Collaboration and levels of automation

Example 1 of human-robot collaboration - Spot

As mentioned in iteration one, Spot is a semi-automatic robot developed by Boston Dynamics. Its physical form resembles that of a dog, with four legs, a long body and a type of head in the front. The head of the robot is where the camera sit, and it is used to show what Spot sees to its operator.

(Spot)

As of September 2019, Spot could walk in rough terrain, climb up and down stairs and avoid crashing into walls, but the plan is to implement more and more modules to add to it. This includes an arm with a claw, an arm with a camera, and more specialized modules made for specific uses.

To control spot, the operator uses a handheld controller with a touchscreen, joystick, arrows and other buttons. Its layout resembles the layout of an Xbox controller.

The operator can control both Spot and the camera in Spot's head. Spot can be given directions by tapping a location on the touchscreen. Spot will then walk over to the marked area on the screen. The goal of Boston Dynamics is to make Spot fully automatic and not need an operator at all.

In the figure of levels of automation in the book Designing for Situation Awareness by Endsley, we are presented with twelve levels of automation, ranging from manual control to full automation (Endsley, 2011: 185). To easier talk about the different levels of automation presented in this table, I will include a simplified version of said table and give each level a number.

Level number	Level name
1	Manual control
2	Information cueing
3	SA support
4	Action support / tele-operation
5	Batch processing
6	Shared control
7	Decision support
8	Blended decision making (management by consent)
9	Rigid system
10	Automated decision making
11	Supervisory control (management by exception)
12	Full automation

I believe that Spot is at level 5 now. The description of batch processing is as following: "Computer completely carries out singular or sets of tasks commanded by human." As I have

understood Spot, it is not capable of making decisions itself, but I does not need human intervention after it is given a task. It seems like the goal for Boston Dynamics is to keep this level of automation future modules as well. When these modules are fully developed, some of the tasks Spots will be doing are metal detection, gas detection and 3D-modelling.

If Spot was to have any higher level of automation that level 5, Spot would be able to generate its own decision options. This may include walking without being navigated by an operator, deciding where to go, where to check for gas, what to grab and how hard to grab a given object. In order to do that, the robot will need a higher level of artificial intelligence.

On the one hand, giving the robot a higher level of artificial intelligence will greatly reduce the amount of manpower. Since the robot can walk and do tasks on its own accord, no operator will be needed to constantly control it. The robot's task will therefore not be affected by human factors such as fatigue and distractions of the operator. Endsley states that "when automation aids in or task over task implementation, overall performance improves" (Endsley, 2011: 184).

On the other hand, making the robot more automated may affect the operator's concentration. Endsley also states that in research done on situation awareness, "people were faster to respond to system failure when operating under intermediate levels of control than when operating under full automation" (Endsley, 2011: 184). This means that it is important to find the right level between over-automation and under-automation of a system.

Example 2 of human-robot collaboration – Sophia

The robot Sophia is developed by Hanson Robotics in 2017. Sophia is a physical robot that resembles a human being. She can move her limbs, make facial expressions and show emotions, and understand and convey meaningful language. She even has her own Twitter account where she tweets.


(Sophia's Twitter image)

As Hanson Robotics write on their webpage, Sophia can recognize human faces, see emotional expressions, recognize various hand gestures,

estimate feelings during a conversation and try to find ways to achieve goals for the person she is talking to (Hanson Robotics, 2019). This is all possible due to Sophia's artificial intelligence.

Since Sophia is quite a complex robot, I will focus mainly on her AI dialogue system. In Sophia's Twitter bio, she states that the account is "run in collaboration with my AI dialogue system and my human social media team" (@RealSophiaRobot, 2019). Based on this, I believe that Sophia is at either level 7 or 8 of automation, or decision support or blended decision making.

If Sophia's level of automation is changed to level 11 or 12, several things might happen. On the one hand, I believe that Sophia (depending on her algorithm) would develop to become even more complex and, arguably, more humanlike. As of now, Sophia's Twitter mainly consists of videos, articles, tweets and pictures of robots, technology-related issues or herself. Her Twitter seems overly narcissistic. If Sophia had more control over her own Twitter page, her feed might look more diverse, and she might want to address issues that does not concern robots, technology or herself. I argue that this would make her more human.

On the other hand, she might have ended up like Microsoft's artificial intelligence chatter bot Tay who also had her own Twitter account. In 2016, Microsoft launched Tay and let her tweet for 16 hours before they had to shut her down. Tay was fed racist and sexually charged tweets, learned from them and started tweeting with the same intentions. When the developers realized what Tay was tweeting about, she was shut down.

(Tay, Microsoft's Twitter bot).

Since Sophia is not a task-doing or work-related robot, the advantages and disadvantages of changing her level of automation will probably mostly affect her research team. If Sophia the physical robot, not only her dialogue system, is fully automated, she might be able to maintain herself; just as we humans go to sleep, Sophia will have to charge herself. An advantage of this would be that the research team could do more thorough research on general artificial intelligence over time since Sophia would be able to live a next-to-normal human life. However, she might encounter several obstacles if this is the case, such as the

uncanny valley effect. While human tend to prefer interacting with robots that do have some resemblance to known living creatures (like Paro, the seal-like emotional comfort robot(Phillips et al, 2016: 106)), robots that are too close to human resemblance without actually being human are likely to make humans distrusting of and creeped out of it, and the distance between the robot and humans will grow.

Another obstacle of Sophia being level 12 of automation is that it would be impossible for her developers to intervene without having to physically and forcefully stop her in some way. As we have already experienced, Sophia might end up believing or doing things we would like to avoid:

> “In March of 2016, Sophia's creator, David Hanson of Hanson Robotics, asked Sophia during a live demonstration at the SXSW festival, "Do you want to destroy humans?...Please say 'no.'" With a blank expression, Sophia responded, "OK. I will destroy humans."”
>
> (Weller, 2019).

It would be catastrophic if Sophia ended up planning and executing something to harm humans. Therefore, changing the level of automation of a robot or AI system is something that should only be done if one can make sure to find ways to prevent or handle possible negative consequences.

References

Articles

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-ready.pdf)

Endsley. M. (2011). Designing for Situation Awareness: An Approach to Use-Centered Design. CRC Press.

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)

Phillips, E., Ososky, S., Swigert, B. & Jentsch, F. (2016). Human-Animal Teams as an Analog for Future Human-Robot Teams. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 56, no. 1, 100-125.

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)

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 (pp. 5286-5297). ACM. (https://www.microsoft.com/en-us/research/wpcontent/uploads/2016/08/p5286-luger.pdf)

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.

Thrun, S., 2004. Toward a Framework for Human-robot Interaction. Hum.-Comput. Interact. https://www.tandfonline.com/doi/pdf/10.1080/07370024.2004.9667338

Other references

9gag videos: Friends ross Nicholas Cage faceswap piano. YouTube. Accessed 27.09.19. https://www.youtube.com/watch?v=m0VNeufVPKY

@RealSophiaRobot (2019), Sophia the Robot, Twitter.com.

https://twitter.com/realsophiarobot?lang=en

BLACK BOX | Meaning in the Cambridge English Dictionary. Accessed 24 October 2019. https://dictionary.cambridge.org/dictionary/english/black-box.

Bosch: The history of artificial intelligence. Accessed 24.09.19. https://www.bosch.com/stories/history-of-artificial-intelligence/

Følstad, Asbjørn (2019). Interaction with AI – Module 2. Accessed 17.10.19. https://www.uio.no/studier/emner/matnat/ifi/IN5480/h19/undervisningsmateriale/interacting-with-ai-2019---module-2---session-1---handout.pdf

Grudin, Jonathan. AI and HCI: Two Fields Divided by a Common Focus. AI magazine 30, no 4 (September 18, 2009).

Hanson Robotics: Sophia. Accessed 17.10.19. https://www.hansonrobotics.com/sophia/

Kuiper, K.: R.U.R. Britannica.com. Accessed 24.09.19. https://www.britannica.com/topic/RUR

Oxford Dictionary: artificial intelligence. Lexico.com. Accessed 24.09.19. https://www.lexico.com/en/definition/artificial_intelligence

Oxford Dictionary: robot. Lexico. com. Accessed 24.09.19. https://www.lexico.com/en/definition/robot

Rouse, M.: AI (artificial intelligence). Tech Target. Accessed 24.09.19. https://searchenterpriseai.techtarget.com/definition/AI-Artificial-Intelligence

Russell, Stuart J.; Norvig, Peter (2009). Artificial Intelligence: A Modern Approach (3rd ed.).

The Verge: Boston Dynamics Spot hands-on: new dog, new tricks. Posted 24.09.19. Accessed 24.09.19. https://www.youtube.com/watch?v=bmNaLtC6vkU

Weller, C. (2019). The First 'Robot Citizen' is the World Once Said She wants to 'Destroy Humans'. Inc.com. Accessed 13.11.19 https://www.inc.com/business-insider/sophia-humanoid-first-robot-citizen-of-the-world-saudi-arabia-2017.html

Wikipedia (2019). Tay (bot). Accessed 13.11.19 https://en.wikipedia.org/wiki/Tay_(bot)

Appendix

Feedback iteration 1

The feedback I got from my peer on this paper was quite helpful. I had one unfinished sentence in the middle of my paper, which I did not notice myself. I ended up finishing that sentence and adding a few lines to that paragraph as well. I had stopped in the middle of an argument, so it was great to be notified about the absence of the actual argument.

I also got feedback on a paragraph where I wrote that I did not feel like I truly understood the topic enough to make my own definition of it. I felt like it was a bit too simple, but I ended up keeping my original definition and rewriting my explanation of it.

Feedback iteration 2

I got good feedback on iteration two. I had some strange formulations that my peer picked up on, so I ended up changing most of them. I also had some arguments where I did not complete my trail of thought, and luckily my peer picked up on that too.