INDIVIDUAL ASSIGNMENT

IN5480

HØST 2020

henreld

MODULE 1

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

Write a section about how AI came about, the history of AI.

Thoughts on the possibility of machinery showing intelligent behaviour first came around after Turing's code breaking in World War II. The term Artificial Intelligence (AI) was first used in a workshop in 1956 by American mathematician and logician John McCarthy (Grudin, 2009, p. 49). The Cold War and Soviet launch of the Sputnik satellite brought new attention to AI and how it could be put to use by computers. Several research projects received financial support which lead to AI researchers getting financial independence, resulting in AI being established as a field (Grudin, 2009, p. 50). The ambitions about AI and its future were high in the 1960s, and AI was viewed as the "renaissance machine" that would be capable of doing any work a man can do within 20 years. Mid 1970s to early 1980s resulted in an AI winter. It became clear that AI has been oversold, and it got a lot of criticism which made it difficult to obtain funding for AI projects, resulting in the U.S and British government cutting off exploratory research in AI. (Grudin, 2009, p. 52). In the 1980s AI projects restored funding due to the threat of Japan's launch of the "Fifth Generation" AI effort. The invention of the LISP machines also contributed to the "AI bloom" (Grudin, 2009, p. 53). In the 1990s came a new winter with the collapse of the LISP machine (Grudin, 2009, p. 54). In the current decade AI came to a new bloom because of the increase of computational power and emphasis on solving specific problems. The chess-playing system called Deep Blue beat the reigning world chess champion, which also contributed to the bloom (Grudin, 2009, p. 55).

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 I chose is by Tone Bratteteig and Guri Verne (2018):

> *"AI is 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 intelligence if executed by a human"*

Bratteteig and Verne work at the department of informatics at UiO and have a background in participatory design, design of information systems, and computer supported cooperative work. The definition is centered around the computer being able to think humanly. When a computer thinks like a human, they can be able to perform different tasks that would require intelligence from a human to accomplish, like for example driving a car.

The second definition from the International Dictionary of Artificial Intelligence written by W. J. Raynor (1999):

> ´"*Generally, Artificial Intelligence is the field concerned with developing techniques to allow computers to act in a manner that seems like an intelligent organism, such as a human would."*

Raynor earned a Ph.D. in Biostatistics from the University of North Carolina at Chapel Hill and is currently a Senior Research Fellow at Kimberly-Clark Corp. This definition is not too far from the first one, and also states that artificial intelligence has to do with computers acting like an intelligent organism - such as humans and therefore does not exclude other intelligent beings. In comparison to the first definition, which focuses on the computer *mimicking* human intelligence, this definition focuses on computers acting in a way which makes it *seem* like an intelligent organism.

The third definition is taken from Britannica written by B.J Copeland for Britannica.com:

> *"Artificial intelligence (AI), the ability of a digital computer or computer-controlled robot to perform tasks commonly associated with intelligent beings."*

Copeland is a Professor of Philosophy and Director of the Turing Archive for the History of Computing, University of Canterbury. Copeland's definition differs from the others because he focuses on the computers being able to perform tasks *associated* with intelligent beings. Also, he mentions computer-controlled robots.

This is my attempt on a definition of AI:

> *" Artificial intelligence is techniques the allows computers to perform tasks that would require some sort of intelligence as if it were accomplished by a human"*

Of the three definitions I agree with Bratteteig and Verne's the most, and therefore based my definition from theirs.

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

I chose a company called Spacemaker AI, which delivers a service for real estate developers that uses AI to simulate and modulates building projects in the early stages. This way one can come up with more proposals than what an architect would have done, because of the quickness and efficiency of the AI technology. On their website we are greeted with a headline stating: *"Early stage planning. Re-imagined*". Below this headline there is a subtext stating: *"Spacemaker AI is a cloud-based AI software, empowering teams to collaborate, analyze and design real estate sites. Lower risk. Faster projects. Better homes"* (Spacemaker AI, n.d.). Spacemaker AI describes AI as a service and also a product, especially for real-estate developers. The company also introduces the idea that AI will strengthen collaboration and create better homes. Their focus on AI is aimed towards its qualities like it being able to perform tasks fast and effective.

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

The movie "Her" (2013) has a futuristic approach to AI and revolves around the relationship between the two protagonists Samantha, which is the computer program, and Theodore, which is the user of the program. Samantha goes from being Theodores virtual assistant to a companion whom he develops a connection with. Samantha uses voice queries and natural-language user interface, similar to today's Siri from Apple. In the movie, the AI is portrayed very human-like, and shows human qualities like emotional intelligence, and shows understanding and reflection of its surroundings. This is especially shown in the movie where Samantha asks questions like "What's it like to be alive in that room right now?" and "How do you share your life with someone?". "Her" gives a portrayal of AI's

potential for the future where it can serve as a replacement for human interaction and connection.

1.2 Robots and AI systems

How the word Robot came about.

According to sciencefriday.com, the term "robot" was first introduced by a Czech Karel Čapek in one of his plays in 1920 Rossum's Universal Robots (Sciencefriday). The play tells a story about a company that mass-produced workers that "lack nothing but a soul", which would do all the work humans didn't want to do. "Robot" is a translation of the word "robota" which means "servitude", "forced labor" or "drudgery".

Two different definitions of "robot". Describe and explain these definitions. Based on these definitions, make one definition yourself, and describe and explain this definition.

The first definition is from 1979 retrieved from the Robot Institute of America:

> *"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"* (Russell&Norvig, 1995).

This definition sounds to me like it is very machine-focused and leans more toward Thrun's (2004) categorization of an industrial robot as it mentions that the robots' purpose is to move parts and tools for different tasks and does not mention any interactivity with humans.

The second definition is retrieved from Merriam-Webster's dictionary

> *"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)" (Bowker, n. d.)*

This definition mentions the robot resembling living creatures, which differs from the first definition. It also states that the robot moves independently, in likeness with living creatures, whereas in the first definition it focuses more on the robot's movement being programmed and planned.

My attempt at a definition of "robot" would be:

"a programmed machine that can move independently and perform several tasks".

It is a quite vague definition, but I feel that it includes all three categories of robots stated by Thrun (2020). I wanted to include that robots can interact with people in some way, but came to the conclusion that it would exclude certain types of robots used in industrial environments.

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.

The AI definitions focus on intelligence and intelligent behaviour, which can in some sorts be implemented into robots. The definition of robots focuses on the ability to move, which I feel is an attribute specific for robots and is where the difference lay between AI and robots. AI can be implemented into robots, but a robot doesn't necessarily need AI to function. The definition from Britannica mentioned above includes computer-controlled robots in their definition of AI.

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.

I chose the robot "Pepper", which is advertised as "the world's first humanoid robot able to recognize faces and basic human emotions" and is designed to interact with humans (Softbank Robotics, n.d.). Peppers has a body with a head, arms with joints, hands and fingers and a moveable waist - which overall makes its movements very human-like. Pepper is set on wheels and can move 360 degrees. The video-clip of the robot shows it "dancing" to music by moving its arms, head and upper body. It is a robot designed for human interaction, so the humanly movements and friendly interface makes sense in order to encourage people to communicate with it as if it were a human.

1.3 Universal Design and AI systems

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

Definition taken from National Disability Authority webpage (2020) states that:

> *"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"*.

This definition describes that one should take into account people with certain limited abilities, for example visually impaired, when designing a product so that no one gets excluded from using the design. Universal design concerns including *all* people in a way that they can use a design without difficulty related to their abilities. When these matters are not taken into consideration, some people will be excluded from accessing the design without any reasoning other than their limited abilities.

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

Some of the concerns people can have related to AI, is the increasing loss of future jobs and workforce that might be replaced by AI systems. As mentioned above the company Spacemaker AI already has technology that can perform a job that architects usually would, a lot quicker and more cost-efficient. On the other hand, AI has great potential considering universal design, where AI implementation can help include people with limited abilities, for example blind or visually impaired. An example of this is Vision, which is a smartphone app that uses AI to help blind and visually impaired people to access visual information around them. AI can also have great potential concerning human emotions when used for social interaction, briefly discussed in the section about the movie "Her" above.

In WCAG 2.1 principles and in the Human AI-Interaction guidelines the concept "understand"and "understanding" is used. Explain briefly in what way you make sense of the concept "understand" and "understanding". Then address the question: Do machines understand?

"Understand" has meaning to me in the sense that you perceive the meaning of something in a way that it was intended, or interpret something in a particular way so it gives meaning to you. I interpret "understanding" as the *ability* to understand. I do not think that machines

really understand, because then they would have to be conscious of their own existence, and have the ability to reflect upon matters independently.

1.4 Guideline for Human-AI interaction

Please select one of the 18 guidelines from Microsoft, and describe this guideline with a different example than what is given by Microsoft.

I chose the guideline G13 “learn from user behavior” where Netflix is a great example of this. As you watch movies and TV shows, Netflix personalizes the content available and recommends new movies and TV shows based on what you’ve previously have been watching.

Search, and find one set of HCI design guidelines. Discuss briefly similarities and differences
between the HCI design guidelines and the Human-AI interaction guidelines.

I chose Donald Norman’s seven principles for design (Preece, 2012) as an example of HCI design guidelines. I found it difficult to find any differences between Norman’s principles and the Human-AI Interaction guidelines other than that the Human-AI guidelines are more specific and detailed, whereas the design principles are more broad. This makes sense since the principles are more like general rules of thumb. One could almost categorize the 18 guidelines into Norman’s seven principles.

MODULE 2

2.1 Characteristics of AI-infused systems

Identify and describe key characteristics of AI-infused systems.

The key characteristics that were mentioned in the second lecture of module two are *learning*, *improving*, *black box*, and *fuelled by large data sets*.

That the AI system is learning means that it has a dynamic character. Amershi et al. (2019) highlights that the system is constantly changing each time the user interacts with the system. AI's ability to learn therefore leads to the system being able to improve itself based on the patterns and behaviour from the users interactions. AI systems can be seen as a "black box" which is the sense that the user usually does not have any insight towards how the AI system operates and how it works. The user can input something to the system and get output/feedback, but it does not show how it came about and why the user received that particular feedback. This may cause problems for the user-interaction with AI-infused systems because of the lack of understanding on how the system works, how well it works, and the expectations towards it. Therefore it is desirable to take the design out of "the black box" and design for a more explainable system, which is highlighted in the article by Kocielnik (2019). The last characteristic is that AI systems are fuelled by large data sets. The system is dependent on having access to large data sets to be able to perform better at their task, and are also the foundation for the systems learning and improving capabilities.

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 chose the AI-infused system Google Assistant in the form of a Google Home speaker. It shows the learning capabilities through for example recognizing my voice, where it was more difficult for it to pick up what I was saying in the beginning, but after using it for a while it learned how I pronounced the word Google and so on, which also is an example on how the system learns and then improves accordingly. A Google Home smart speaker also demonstrates the black box aspect of AI-infused systems, where I find it difficult sometimes

to understand why I got that specific response, where it found that certain information or why it misinterpreted me. As it has no visual interface this makes it complicated to get feedback on. Also, it sometimes responds to something said in a specific commercial when the TV is on which I am curious of what and why that happened. Google Assistant also works from a large set of data from the apps and accounts that are connected to it, and can use data from Google search to perform better. It also records the voice interactions for better task performance as briefly mentioned above.

2.2 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.

Some of the main take-aways from Amershi et al. (2019) is that they state that we need to rethink the way we design human-AI interaction because of the growing uses of human-facing applications with AI technology. They develop a set of 18 principles for interaction with AI, which they validate by multiple rounds of evaluation with design practitioners that testet them on several popular AI-infused products. The goal with these guidelines is to give practitioners a tool for making more human-centric AI systems, and highlights that the principles are not set in stone and it is important to constantly test and improve these.

Kocielnik at al. (2019) discusses the importance of setting the right user expectations within AI-infused systems, and investigates how to shape these expectations in order for greater acceptance of the system prior to use. They look at the design of an AI-powered scheduling assistant for emails, and how it should identify emails and ways to set up these to ease the work for the user. The authors highlight the importance of the user understanding of how the system works for better usability, and show them how likely it is that the system is right. They identify three ways where expectations are formed which resulted in suggesting three techniques for adjusting expectations; Accuracy Indicator, Examples based Explanation, and Performance Control.

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.

Guideline G7: support efficient invocation.

To initiate the Google Home one has to use the command “Hei Google” in Norwegain, which I understand is the only way. This works well at times, but I have experienced it having problems initiating with people who pronounce the word “google” differently, which can be quite common because some pronounce it the english way and others with their Norwegian accent like “goggel” or “gogl”. To assist the invocation there is a feedback sound that lets you know when you have successfully woken it up, which is really helpful to avoid users continuing with their request without the system even being initiated.

Guideline G4: show contextually relevant information.

When you ask Google Home about the weather or temperature it is based out of my location without having to specify it. Although it does not show it this visually, it portrays this by saying it; “the temperature in your location Oslo is..”. To improve this I think it would require a visual interface to compliment the speaker.

2.3 Chatbots / conversational user interfaces

Chatbots are one type of AI-infused systems. Based on the lectures, and the mandatory articles, discuss key challenges in the design of chatbots / conversational user interfaces.

When designing a chatbot there is no longer the graphical aspect that is the most important material you work with, but it is the design of the conversation. To design for a dialog involves interesting implications and challenges.

When designing a conversation for a chatbot it can be challenging to combat the diversity amongst users, which is one of the challenges Følstad and Brandtzaeg (2017) discusses. The chatbot should be able to communicate naturally with the user no matter what their gender, race, age, language or potential disabilities are.

The fact that chatbots often do not meet with the users expectations, suggests another challenge. Kocielnik et al. (2019) highlights the importance of designing a AI system in a way that it sets the appropriate expectations with the user before the initial use of the system. If the chatbot isn't responding in a way the end-user would expect, it could negatively affect their acceptance of the system.

Liao et al. (2020) discusses the explainability of AI systems which can be applied to chatbots, where the challenge is to design the chatbot in a way so the user understands how it works and why it answered the way it did. If the chatbot manages to communicate this well, it could help with the users perception and acceptance of the system.

Another challenge is that the interaction with a chatbot can be different from user to user, and knowing in advance how the conversation plays out for each user is challenging. Yang et al. (2020) points out this issue with AI-based systems, where it is challenging to design in advance for a system that is learning and dependent on data partly because one is not sure in advance what the system is capable of doing.

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.

Amerishi et al. (2019) guideline G1 states as follows: *Make clear what the system can do*. Adherence to this guideline can help solve some issues concerning the users expectations towards the chatbot. If the chatbot communicates clearly what it can offer and help with to the user at the very start of the interaction, it can alter the users expectations and have a positive impact on their perception.

Guideline G2 says: *Make clear how well the system can do what it can do.* This means that it also is just as important to make clear what the system *cannot* do. If a chatbot communicates this well throughout the user interaction it can resolve some challenges concerning its explainability. If the chatbot can provide the user with information on how well it can do what it can do, or cannot do, the user can then alter its inputs respectively and make way for a more fruitful interaction.

MODULE 3

3.1 Human AI collaboration

Philips at al. (2016) give a taxonomy and examples of human-robots collaboration. Choose 2- 3 examples, describe their levels of autonomy as described in Shneiderman (2020) and reflect on advantages and disadvantages if we decrease/increase their current level of autonomy.

Philips et al. (2016) writes about how robots should be seen as a team member rather than a tool. The authors describe how human-animal teams can serve as an analog for developing human-robot teams, and give examples and classifications of different human-robots collaboration. The examples I have chosen from the paper are the Big Dog robot and the Paro robot. To describe their levels of autonomy I will use Sheridan and Verplanks's 10 level autonomy scale as described in Shneiderman (2020), where 10 would be the highest level of autonomy and 1 would be the lowest level.

Level	Description
High	10. The computer decides everything and acts autonomously, ignoring the human.
	9. The computer informs the human only if it, the computer, decides to.
	8. The computer informs the human only if asked, or
	7. The computer executes automatically, then necessarily informs the human, and
	6. The computer allows the human a restricted time to veto before automatic execution, or
	5. The computer executes that suggestion if the human approves, or
	4. The computer suggests one alternative, or
	3. The computer narrows the selection down to a few, or
	2. The computer offers a complete set of decision/action alternatives, or
Low	1. The computer offers no assistance; the human must take all decisions and actions.

Example 1: Big Dog robot

The Big Dog robot is a robot designed by Boston Dynamics for the military, where it's purpose is to carry and transport cargo to reduce the soldiers load in terrain that's too rough for conventional vehicles (Philips et al., 2016, p. 104). As the name implies it is designed to resemble the structure of a big dog or a mule, and can walk in rough or uncertain terrain by using its four robotic legs and a variety of sensors. The robot has an on-board computer that uses data from the sensors to keep its balance, locomote through different terrains, and for navigating (Darpa, n.d.). The Big Dog operates by using its many sensors which gives it a high level of autonomy, where it can make decisions independently on where to place its legs in order to keep balance and move forward. However, It is dependent on having a remote human to operate it which communicates with the on-board computer (Raibert et al., 2008). Based on this I would place The Big Dog around level 4 to 5 on the autonomy scale because the autonomous behavior of the robot is its ability to move by itself and analyse different terrains in order to make the right movements to keep balance and stay on track, but is dependent on having a human to steer its course. The human makes the decision on where the robot should go, how fast, and when to stop but the robot carries these tasks out itself by being able to maneuver its legs by calculating the surface it walks on.

If one were to increase The Bog Dog's autonomy to the level where it would move entirely on its own without having a human to operate it would probably not fit in a combat situation, where communication related to where to move and when is essential in that

context with a group of people. If one were to decrease the automation it would require more support and attention from the soldiers and one would likely have to add extra resources to monitor the robot, which then goes against its purpose to relieve the soldiers load.

Example 2: Paro robot

The Paro robot is designed to look like a seal and is meant to be a therapeutic tool to give the same benefits as in animal therapy. It has been used for elderly people with diagnosed dementia to fight depression and stimulate interaction (Parorobots, 2014). Paro imitates the voice of a baby seal, and reacts and moves in different ways depending on how the human interacts with it. The robot does this by using five different kinds of sensors. For instance it recalls the last interaction with a human as something positive of the human were to pet it afterwards and tries to recall this and repeat it more often, and if a human were to slap the robot after an interaction it will try not to do the same thing in the future. It can therefore learn to behave in a way that the user prefers. Because Paro's actions are determined by how the human interacts with it would mean that the user is partly in control of their decision making, where Paro offers suggestions as in different kinds of responses of interaction and the user chooses a suggestion by how they decide to interact with it, which Paro remembers for further interactions. I found it quite difficult to place this on a certain level, but my understanding is that Paros behavior is automatic, but the behaviour is determined by a human's interactions which then makes the human partly in control of it. But, the human can not command the robot to do a specific action. Therefore I would consider placing Paro around level 7 on the scale of autonomy.

If one decreases Paro's current level of autonomy it may lose its function as a believable pet for the users and become more of a toy or stuffed animal, because it would decrease its mutual interaction with the user. If one were to increase its level of autonomy, I would imagine that Paro would behave more animal-like and move around seeking attention. This could either lead to giving the users a higher therapeutic value, or it could be too much to handle or intimidating for a potential elderly patient.

Reflect on their current and needed explainability (Hagras, 2018; Smith-Renner et al. 2020).

Hagras (2018) discusses how explainability is a necessity for AI to become a trusted tool and integrated into society and the importance of moving towards “explainable AI” abbreviated as XAI (p. 29). XAI is described as an AI where its actions and decisions can easily be understood by humans, in other words humans are able to understand and analyze how it works. The first example, The Big Dog, considering the level of autonomy it is at doesn’t require too much explainability as it is being remotely controlled by a human. Maybe if it failed to execute a certain behavior it would be beneficial for the operator to understand why this happened in relation to error prevention. If the Big Dog were to have a higher level of autonomy and move around by itself, its explainability would be essential in order to trust it being safe for operating in a military setting and have confidence that it will make the right decisions. For the Paro robot, it seems to me that the elderly patients have very little knowledge on how it actually works, and nor do I think they need to as they see it more as a pet. But the caregivers, that are responsible for the patients, need to have a certain understanding of the robot to feel safe leaving them alone with their patients. Paro does not really display why it did a certain behavior and the AI and its sensors are hidden inside a stuffed animal which at first sight could seem as a regular toy. So for humans to understand what it does and why, they need to be told by someone or read about it in a manual. But, can this give them enough understanding to feel safe that the robot does not repeat any unwanted behaviour, or that the AI system has not learned a biased view of the world?

APPENDIX

Feedback on iteration 1 and 2

I only received positive feedback on iteration 1, so I decided not to change anything from the first iteration. I made adjustments on the layout and added a table of contents. On the second iteration I got some feedback on part 2 where I did some slight adjustments.

REFERENCES

Module 1

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

Thrun, S. (2004) Toward a Framework for Human-robot Interaction, *Human-Computer Interaction,* 19:1-2, 9-24. doi: 10.1080/07370024.2004.9667338.

Bratteteig, T., and Verne, G. (2018) Does AI Make PD Obsolete?: Exploring Challenges from Artificial Intelligence to Participatory Design, *In Proceedings of PDC 2018*, Belgium, August 2018, 5 pages. doi: 10.1145/3210604.3210646

Raynor, W. J. (1999) *The International Dictionary of Artificial Intelligence*. The Glenlake Publishing Company.

Russell, S., and Norvig, P. (1995). *Artificial intelligence: A modern approach*. Englewood Cliffs, NJ: Prentice Hall.

Preece, Jennifer., Rogers, Yvonne. & Sharp, Helen (2012) *Interaction Design: beyond human computer interaction.* New York: John Wiley & Sons, Inc.

Bowker, M. (n.d.). Robot, In *Merriam-Webster.com dictionary*. Viewed: September 10 2020, <https://www.merriam-webster.com/dictionary/robot>.

Softbank Robotics (n.d.) Pepper the humanoid and programmable robot. Viewed: September 10 2020, <https://www.softbankrobotics.com/emea/en/pepper>.

Science Friday. (n.d.) The Origins Of The Word 'Robot'. Viewed: September 10 2020, <https://www.sciencefriday.com/segments/the-origin-of-the-word-robot/>.

Copeland, B.J. (2020) Artificial intelligence, in *Britannica.* Viewed: September 10 2020, <https://www.britannica.com/technology/artificial-intelligence>.

National Disability Authority. (2020) What is Universal Design. Viewed: September 10, 2020. <http://universaldesign.ie/What-is-Universal-Design/>

Spacemaker AI (n.d.) Viewed: September 10, 2020, <https://www.spacemakerai.com>.

Module 2

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)

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)

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. pd f)

Liao, Q. V., Gruen, D., & Miller, S. (2020, April). Questioning the AI: Informing Design Practices for Explainable AI User Experiences. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (paper no. 463). ACM.

Yang, Q., Steinfeld, A., Rosé, C., & Zimmerman, J. (2020, April). Re-examining Whether, Why, and How Human-AI Interaction Is Uniquely Difficult to Design. In Proceedings of the 2020 chi conference on human factors in computing systems (Paper no. 164).

Module 3

Darpa (n.d) Big Dog. Viewed: 11. november 2020, <https://www.darpa.mil/about-us/timeline/big-dog>

Shneiderman, B. (February 23, 2020), Human-Centered Artificial Intelligence: Reliable, Safe & Trustworthy, arXiv.org. https://arxiv.org/abs/2002.04087v1 (Extract from forthcoming book by the same title)

Raibert, M., Blankespoor, K., Nelson, G., Playter, R. (2008) BigDog, the Rough-Terrain Quadruped Robot, IFAC Proceedings Volumes, Volume 41, Issue 2, 2008, Pages 10822-10825, https://doi.org/10.3182/20080706-5-KR-1001.01833. (http://www.sciencedirect.com/science/article/pii/S1474667016407020)

Parorobots (2014) PARO the therapeutic robot, Viewed: 11. november 2020, <http://www.parorobots.com>

Phillips, E. K., Jentsch, F., Billings, D. R., Schaefer K. E., Hancock, P. A. (2015) Human-animal Teams as an Analog for Future Human-Robot Teams: Influencing Design and Fostering Trust. In Journal of Human-Robot interaction, September 2015. DOI: 10.5898/JHRI.5.1.Phillips

Hagras, H. (2018) Toward Human-Understandable, Explainable AI, Computer, 51, 9, 2018, 28- 36 https://ieeexplore.ieee.org/document/8481251