Individual assignment, third iteration

for Irene Solberg - irensolb@mail.uio.no

1. Search and find three definitions of AI, describe these briefly.

(1) The first definition I choose is John McCarthys definition. McCarthy was the first person to coin the term "artificial intelligence" in 1956, however this definition is from an article from 1998.

> AI is the science and engineering of making intelligent machines especially intelligent computer programs. It is related to the similar task of using computers to understand human intelligence, but AI does not have to confine itself to methods that are biologically observable. (McCarthy, 1998)

(2) The second definition I would like to put forward is wikipedia's definition. I chose this definition because wikipedia is a place people go to quickly research different things. I believe that the wikipedia definitions often are the "most common" and most "available" definitions to regular people. It reads as follows:

> Artificial intelligence (AI), sometimes called machine intelligence, is intelligence demonstrated by machines, in contrast to the natural intelligence displayed by humans and other animals. Colloquially, the term "artificial intelligence" is applied when a machine mimics "cognitive" functions that humans associate with other human minds, such as "learning" and "problem solving" ("Artificial intelligence," n.d.).

(3) The third definition is not exactly a definition but more a statement about artificial intelligence. I would like to highlight this, as it can be something to think about. Elon Musk says that;

"AI is capable of vastly more than almost anyone knows and the rate of improvement is exponential." (Clifford, 2018)

and

"..mark my words, AI is far more dangerous than nukes. Far. So why do we have no regulatory oversight? This is insane" (Clifford, 2018)

2. Search and find three definitions of Robotics, describe these briefly.

Nasa has the very brief definition on robotics (1) "Robotics is the study of robots" (Nasa, 2009). They say that robotics is only the field of study focusing on robots. They do not include the "use" of robots in their definition, nor do they specify the potential focus areas of the study of robots.

Again I would also like to use the wikipedia definition, for the same reason as in assignment 1. This definition is more detailed than the previous, and include the focus areas of field of study.

(2) Robotics is an interdisciplinary branch of engineering and science that includes mechanical engineering, electronics engineering, computer science, and others. Robotics deals with the design, construction, operation, and use of robots, as well as computer systems for their control, sensory feedback, and information processing ("Robotics", n.d.).

The third definition I chose is from the Encyclopædia Britannica. In this definition, robotics seems like it is more focused on how the robots can achieve human functions, than in the previous definitions.

Robotics is Design, construction, and use of machines (robots) to perform tasks done traditionally by human beings. Many aspects of robotics involve artificial intelligence; robots may be equipped with the equivalent of human senses such as vision, touch,

and the ability to sense temperature. Some are even capable of simple decision making, and current robotics research is geared toward devising robots with a degree of self-sufficiency that will permit mobility and decision-making in an unstructured environment. Today's industrial robots do not resemble human beings; a robot in human form is called an android (KILDE).

3. Search and find three definitions of Machine Learning, describe these briefly.

(1) Machine learning is an evolving branch of computational algorithms that are designed to emulate human intelligence by learning from the surrounding environment. They are considered the working horse in the new era of the so-called big data (Naqa & Murphy, 2015).

(2) Machine learning algorithms can figure out how to perform important tasks by generalizing from examples. This is often feasible and cost-effective where manual programming is not. As more data becomes available, more ambitious problems can be tackled (Domingos, 2012).

(3) Machine learning is a field of computer science that uses statistical techniques to give computer systems the ability to "learn" (e.g., progressively improve performance on a specific task) with data, without being explicitly programmed (Machine learning, n.d.).

These definitions are all talking about how machine learning is about how a machine can train itself to become more intelligent based on data In other words, the system have not been programmed to do different tasks, but is learning to do them based on large amounts of data. We have discussed that when an AI is made with machine learning, it has the potential to become more autonomous than if it were not. An ethical issue with machine learning is, in my opinion, that there is a chance that the data the system is using is incorrect or biased, which can lead to ethical problems. An example of this can be systems such as predictive policing or other systems that are based on human generated data.

**4. Write in three to five sentences the relationship between AI and Robotics as you understand this.**

The difference between AI and robotics is that AI is concerned with the development of intelligent systems, which not necessarily has to be a physical object. Robotics is concerned with the development of robots to do different tasks. A robot is physical, and include hardware, sensors, mobility etc. It can be programmed to do different tasks, or it can be implemented with an AI to be able to achieve more challenging tasks.

5. Make a text to describe your own definition of AI. Explain briefly this definition.

My definition of AI would be that there is no clear consensus on what AI actually is, what it can be, or how it should be developed. I think that the popular understanding of AI, is that it is some sort of computer system that can make decisions without commands from humans, learn from experience, solve problems, and perceive its surroundings and be autonomous to some degree, in many ways similar to the human cognition.

An AI can be abstract, such as a system that runs in the background, but also tangible, such as Amazon Alexa or Pepper the robot.

I understand the relation between AI and machine learning as AI being the broader field, whereas machine learning is a concept within, or a way of making, AI. In other words, machine learning is a sub-concept of artificial intelligence.

6. Make a drawing of an interaction with an AI - something that you imagine.
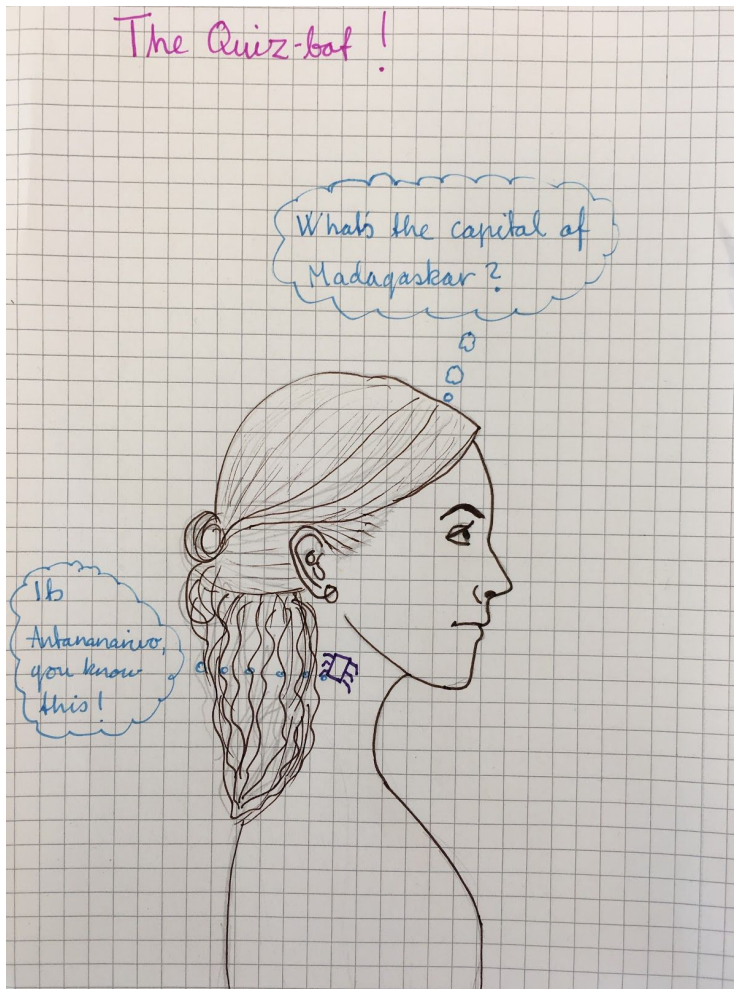
This AI is designed to merge with the human mind. This AI will enhance the human brain, giving the person that has it abilities that it didn't have before. At the same time, the AI can not exist without the human, so it will be safer for humans because the AI can't kill us all!

Key characteristics of interaction design for AI-based systems
1. Conversation is the object of design
2. Speech acts - we speak to perform something
3. Be as informative as required
4. Speak what you believe is the truth.
5. Be relevant
6. Be clear and unambiguous
7. The persona must be consistent
8. Conversation repair: prevent error, provide help or alternatives, if fail: do it well

Tentative design principles
Learn
- Design for the systems to be changing (it is learning)
- Explain what the program does, and that change is going to happen, dynamic character
- show capabilities: make clear to the user what you can do
- be clear on limitations: controle expectations

Improve
- mistakes inevitable - design for uncertainty
    - inform about level of certainty - are the AI uncertain (this looks like glasses?)
    - make recovery easy
    - learn from mistakes (Was this useful for you? Yes/no)

Fuelled by large data sets
- design for data capture
- make users benefit from data

This is our chatbot from the group assignment. It shows its capabilities (what it can do) and is describing its limitations, by saying it is a bot (and not a human) and to call emergency care if you need it.
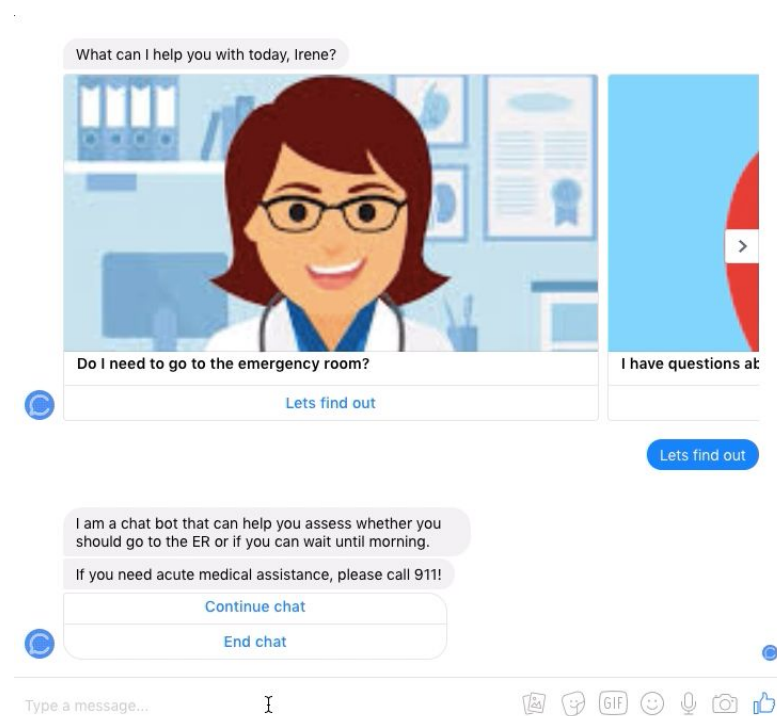
7. Read the article: "On the Subject of Objects: Four Views on Object Perception and Tool Use" by Tarja Susi / Tom Ziemke. Write in your own words one page about the different perspectives on the human relationship with tools.

In this article, the authors is trying to address the relationship between an agent and its environment. They does this by comparing four different perspectives on the relationship between subject and object. The four perspectives are Functional tone, equipment, affordance, and entry point. The authors points out that these perspectives have often been used interchangeably, because of their similarities, however there are also important differences between them.

In the first perspective, functional tone, the relationship between the subject and the object is that the object is neutral until the subject ascribe a meaning to it. In this perspective, the subject, or the person, has its own subjective universe, and what the subject perceive is the subjects own reality. When the subject (a person) form a relationship with the object, the subject ascribe meaning to the object. This can change according to the subjects mindset, as a pillow can be a chair, something to lay your head on, a object for pillowfight, a decorative piece etc., depending on the subjects current mindset.

The second perspective, equipment, is somewhat similar to functional tone. However, Heidegger means that we can not separate subject and object. Subjects are in the world, and can not be considered as separate from their surroundings. An object is not objective or context free, and should be considered as involved with other objects, and not in isolation.

The third perspective, Affordance, is also opposing the separation of subject and object, and emphasize the reciprocal relationship between them. They talk about how each subject lives in its own niche, and that information about the environment is made available in the perceived pattern of light that is reflected from surfaces. In this sense, the environment (affordance) is objective, as it is always there to be perceived, and even though a user's needs change, the affordance does not. They also say that affordance is always in relation to the subject, and thus the subjects movements influences the perceptual activity.

The fourth perspective, Entry point, is about how the subject make use of the environmental structures to achieve various tasks. They say that people actively structure their environments to create sets of entry points to scaffold their work. The structures provided by entry points reduce cognitive load and improve performance. Entry points may be subjective or objective.

8. Select one of the perspectives from the article, and go into detail when you describe it.

The perspective, equipment, means that we can not separate subject and object. Subjects are in the world, and can not be considered as separate from their surroundings. An object is not objective or context free, and should be considered as involved with other objects, and not in isolation. A tool (object) and its possible uses, is perceived according to the subjects ongoing activity. An object is only what it is when it is within a meaningful context of an activity. In addition, Heidegger says that what is perceived by a subject is influenced by other subjects.

9. Select one other article from module 1, and write with your own words what this article is about.

In the article "What we talk about when we talk about context", Dourish criticize the many different understandings of context. He believe that context is not just the physical space where an action takes place, but that both the context and the action makes up the context. He says that context is an emergent property of occasions of interaction, rather than stable and objective set of features. He proposes a solution to the problem of context in developing systems by suggesting something that looks like "design in action", that is, a system that is "designed" by the actual users through the ways it is incorporated into practice.

10. Select one documentary or a fictional film, book or game: describe with your own word how interaction with AI is portrayed in this work.

Recently I have been watching a series called The 100. The series is about a group of people that grew up on a space station, because their ancestors escaped a nuclear war. The space station is running out of air, so 100 young adults is being sent to earth to see if it is habitable. For the first three seasons, the group is in constant war with the "grounders", people that had survived the radiations following the nuclear war. The grounders believes in reincarnation, however as the series continues, the group discovers that what the grounders believe is reincarnation, actually is an AI that is living in a symbiosis with the human mind (as a chip in the human brain). They then discover that there actually is two different AIs, developed by the same woman. The first one, called ALLIE, is actually the one that launched all the nuclear bombs a 100 years ago, and is now trying to kill the ones that survived, because she was not programmed to exist in symbiosis with humans. The second AI was built to stop the first one from destroying the human kind. This is how far I have watched in the series.

I think this is an exiting perspective on AIs, and also something that Elon Musk has stated about AIs. He has said that humans should merge with AIs, to still be relevant and able to control the AIs

"Some high-bandwidth interface to the brain will be something that helps achieve a symbiosis between human and machine intelligence and maybe solves the control problem and the usefulness problem," (Kharpal, 2017)

11. Describe what you understand by autonomy; both human autonomy and machine autonomy.

I understand human autonomy as having free will and the ability to make your own decisions. However, it is debatable whether humans have full autonomy or not. We often have a set of options that we can choose from, however there might be options that we are not even presented with. This is especially interesting when it comes to technology. Do humans really have the option not to adapt to the changing technologies, such as online banking, surveillance etc.?

I understand machine autonomy differently than human autonomy. I think that machine autonomy is a machines ability to carry out tasks and making decisions without any human involvement. This would thus require an AI.

12. When was the term "AI" first coined? Please make a reference.

The term AI was, as mentioned earlier, first coined by John McCarthy in 1956, before the first conference devoted to this subject (AAAI, n.d.).

13. Articulate one question for the article "What we talk about when we talk about context" by Paul Dourish in the curriculum.

In this article, Dourish talks about context in relation to interactive systems in everyday human life. How will this relate to the development of artificial intelligent systems? Can one say that the AI has to negotiate its own context in interaction with humans?

14. Articulate one question for any other article in the curriculum.

In Donald Normans "The problem of automation: Inappropriate feedback and interaction, not over-automation", he talks about how automation often is blamed for increasing the chance of human errors. He says that the irony of automation is not over-automation, but lack of feedback to the human actors (Norman, 1990). If an AI learn to "think like a human", would it then give the appropriate feedback to its actions? If for example humans have formal and rigid procedures of giving feedback they have to follow that can seem annoying to others, how can we design a system that is not like this? Maybe it is necessary in certain types of work, such as flying an airplane.

15. Read the article: "Like Having a Really Bad PA" by Luger & Sellen. Summarize in your own words key lessons learnt for interaction design with dialogue systems. Discuss the relevance of these lessons learnt for interaction with AI-based systems in general

In the article "Like Having a Really Bad PA", Luger and Sellen (2016) talks about conversational agents in everyday life. They argue that there is a gulf between the users

expectations and their experience with conversational agents. They investigate this by interviewing 14 participants with experience with CAs. They found that the users have higher expectations of the CAs intelligence based on how they are portrayed, than their experience when they interact with the CA. They then suggest some key takeaways from their research. The first is that developers of CAs should find ways to reveal the systems intelligens. When the systems were too human-like, the users over estimated the systems capabilities. The second takeaway is to reconsider the interactional promise made by humor. When the CA portray humor, this can give the users higher expectations on the CAs social intelligence. The third takeaway is to consider new ways of conveying CA capabilities through interaction. When the interaction is conversation, it's hard to convey the systems limitations and capabilities. The last takeaway is to rethink system feedback and goals in light of the dominant use case. They found that the primary use of CA was when the users were doing other activities that required both hands and visual attention. The CA should thus not give screen-based responses on verbal questions (Luger & Sellen, 2016).

I Think that the three first points the authors make are relevant for other types of AI systems that interacts with humans. What Luger and Sellen talks about is also similar to what Norman (1990) talks about in *The 'problem' with automation*. He says that automation and inappropriate feedback and interaction can lead to unwanted situations, which is similar to at least the first and third lesson from Luger and Sellen.

16. Describe with your own words what you understand by different levels of automation? What are the advantages/disadvantages related to higher/lower levels of automation? (1/2 – 1 page).

The ten levels of automation was developed by Sheridan & Verplank (1978). I understand the different levels of automation as a scale that describes to what degree a system is automated. This scale range from "The computer offers no assistance: human must take all decision and actions" to "The computer decides everything and acts autonomously, ignoring the human.". From level seven to ten, the system makes decisions on its own, and execute them, independent from a human. This can be a problem, especially if the system is not 100% correct in its decision. An example that is relevant for this is automated weapons. If for example a drone could lokalise, identify and eliminate a threat, all on its own, this

would raise some ethical issues. The system could potentially kill innocent people. In situations where AI have tried to identify threats, they often identify regular civilians hunting, or doing something else "irregular". In situations like this, it's important that a human is in control of an execution of an action. A system like this should not be on level seven or higher (in my opinion a system like this should not even exist).

Some tasks is not possible for humans to do without computers, and some systems could possibly reduce the occurence of human error. However, the level of automation should always be evaluated in regard to the potential risk. As we saw in Norman (1990), the autopilot system made decisions and acted without the pilot's knowledge. The system should probably still act without "approval", as it may not be time for the pilot to act, and the system should also notify the pilot that this decision was made, so that following actions the pilot had to make would be informed by the knowledge of the state of the plane.

When it comes to less critical tasks, like organizing a calendar, or send out invitations to a meeting etc. a higher level of automation could possibly be useful for humans. This would perhaps eliminate repetitive tasks for the user, and would not be a high risk if something went wrong. An interesting question when it comes to AI is also if the machines would reach a level where they can reproduce themselves and take control over humans. This is also something I have discussed in previous tasks in this assignment.

Reference list

AAAI. (n.d.). A Brief History of AI. Retrieved from https://aitopics.org/misc/brief-history

Artificial intelligence. (n.d.). In *Wikipedia*. Retrieved September 3rd, 2018, from

      https://en.wikipedia.org/wiki/Artificial_intelligence

Clifford, C. (2018, March 13). Elon Musk: 'Mark my words — A.I. is far more dangerous

      than nukes'. CNBC. Retrieved from

      https://www.cnbc.com/2018/03/13/elon-musk-at-sxsw-a-i-is-more-dangerous-than-n

      uclear-weapons.html

Domingos, P. (2012). A few useful things to know about machine learning.

      Communications of the ACM, *55*(10), 78-87.

Norman, D. A. (1990). The 'problem'with automation: inappropriate feedback and

      interaction, not 'over-automation'. *Phil. Trans. R. Soc. Lond. B*, *327*(1241),

585-593.

El Naqa, I., & Murphy, M. J. (2015). What is machine learning?. In *Machine Learning in*

      Radiation Oncology (pp. 3-11). Springer, Cham.

Kharpal, A. (2017, February 13th). Elon Musk: Humans must merge with machines or

      become irrelevant in AI age. *CNBC*. Retrieved from

      https://www.cnbc.com/2017/02/13/elon-musk-humans-merge-machines-cyborg-artif

      icial-intelligence-robots.html

Luger, E., & Sellen, A. (2016, May). 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.

Machine learning. (n.d.). in *Wikipedia.* Retrieved September 14th, 2018, from

      https://en.wikipedia.org/wiki/Machine_learning

McCarthy, J. (1998). What is artificial intelligence?. retrieved from

      http://cogprints.org/412/2/whatisai.ps

Nasa. (2009, November 9th). What is robotics. Retrieved from

      https://www.nasa.gov/audience/forstudents/5-8/features/nasa-knows/what_is_roboti

      cs_58.html

Robotics. (n.d.). In Encyclopædia Britannica. Retrieved September 26th, 2018, from

      https://www.britannica.com/technology/robotics

Robotics. (n.d.). In *Wikipedia*. Retrieved September 14th, 2018, from

      https://en.wikipedia.org/wiki/Robotics

Sheridan, T. B. & Verplank, W. (1978). Human and Computer Control of Undersea

Teleoperators. Man-Machine Systems Laboratory. Department of Mechanical

Engineering. Massachusetts Institute of Technology, Cambridge