Individual assignments - third iteration

IN5480 – Spesialisering innen forskning i design av IT - H20

nielsgt@uio.no

Table of contents

First module

Concepts, definition and history of AI and interaction with AI

The history of AI

The concept of AI is based on the idea of building machines capable of thinking, acting, and learning like humans.

The idea of artificial intelligence (AI) predates the technological advances that led to the invention of computers mid 20´s century. By the 1950s, the concept had already been presented to the world through science fiction films like Wizard of Oz (1934), and often in the shape of robots with human-like intelligence.

Throughout history, many prominent people have touched upon the theme. Alan Turing, sometimes referred to as the inventor of computer science, was also one of the first to suggest that computers eventually could solve problems and compete with humans on equal terms (Grudin, 2009).

However, the term artificial intelligence first appeared in the call for participation in a 1956 workshop written by American mathematician and logician John McCarthy (Grudin, 2009)

Definitions of AI

By its increasing popularity, many attempts to define what artificial intelligence actually is and the term is frequently, but inconsistently used today. The term Artificial intelligence is used to describe research fields and technologies, both tangible and fictional. Defining the

term is not restricted to researchers and institutions, so a simple web search returns a variety of explanations, maybe contributing to the mystification of AI.

Oxford's English dictionary describes 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."

Another definition, drawn from a randomly picked online technology magazine states that:

> "Artificial intelligence (AI), also known as machine intelligence, is a branch of computer science that aims to imbue software with the ability to analyze its environment using either predetermined rules and search algorithms, or pattern recognizing machine learning models, and then make decisions based on those analyses."(Techopedia,2020)

In the financial sector, one actor explains to its readers that:

> " Artificial intelligence (AI) refers to the simulation of human intelligence in machines that are programmed to think like humans and mimic their actions. The term may also be applied to any machine that exhibits traits associated with a human mind such as learning and problem-solving." (Frankenfield, 2020)

Combining these three definitions, we could also consider AI being:

> "... a branch of computer science that aims to imbue software with the ability to perform tasks by mimicking human intelligence."

Interpretation and understanding AI

PwC, a international consultant firm, present AI as a central component in "the fourth industrial revolution" and introduces it to its customers as both a broad and specialized technique:

> “AI is a technique used to provide computers and computer programs with the most intelligent response possible. The concept is broad, and we can divide it into weak AI and general AI. Weak AI includes specific tasks a computer can train to do very well, such as playing chess with you. General AI, on the other hand, is more advanced systems that can be trained for almost anything.”

Surprisingly for some, on their website, PwC has provided an overview of what must be inplace before a seemingly costly decision to implement AI in an organization is recommended, thus presenting AI as a driver for organizational change rather than a stand alone product.

AI in fiction

As AI is used in storytelling to generate sales for different businesses, artistic interpretations of AI is a common driver for storytelling in science fiction, film and literature.

An artistic representation of AI, perhaps deviated from the idea of machine learning’s ability to “predict” possible futures, is for example found in the science fiction series *Westworld’s* third season. Throughout the season we are familiarized with a society that is built by humans with optimized skill sets they have gained in accordance with an AI’s predictions.
The AI’s predictions are challenged when the main characters are able to deviate from the predicted path.
This touches upon ethical questions regarding the use of predictive models and machine learning.

Robots and AI systems

Robota

The word robot is drawn from an old slavonic word, robota, meaning forced labor or servitude. This word was "*a product of the central European system of serfdom by which a tenant's rent was paid for in forced labor or service.*" (Science Friday, 2011)

Definitions of robot

Thrun (2004) explains that in 1979, the Robot Institute of America (RIA) defined 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" (Russell & Norvig, 1995).

In science fiction, a robot is often depicted as a machine resembling a human being and able to replicate certain human movements and functions automatically.
This closely resembles another early definition of a robot, found in Merriam Webster's collegiate dictionary (1993) who defines a robot as "An automatic device that performs functions normally ascribed to humans or a machine in the form of a human."(Thrun, 2004)

Another definition of a robot could be "a multifunctional reprogrammable manipulator with a familiar physical appearance."

Relation between AI and Robots

An AI is rarely associated with a physical form, but rather with software. The automated processes in the software could of course resemble those of a programmed robot.
On the other hand, Thrun (2004) points out that. from a technological perspective, robotics integrates ideas from information technology with physical embodiment.
Modern robotics enables AI to be a part of a robot and manipulate the environment.

Robot movement example

Boston Dynamics are known for building several animal-like robots. Names like Big dog and Cheetah, and Spot (a typical name for a dog) further strengthen this image.

Boston Dynamics is funded by the Defense Advanced Research Projects Agency (DARPA) and does research into technology that also has military applications. Their robots are often built with four "legs" supporting movement that strongly resembles and mimics movement seen in the animal kingdom. This makes a robot capable of "jumping" and crawling, to lie down or traverse rough terrain in contrast to the limited capacity of wheeled robots.

These robots are built to support a human operator, and thus fits RIAs description of a robot, but since the robot also has the shape of a four-legged animal the definition from science fiction is also somewhat accurate.

Universal Design and AI systems

Universal Design

"Universal Design is the design and composition of an environment so that it can be accessed, understood and used to the greatest extent possible by all people regardless of their age, size, ability or disability. An environment (or any building, product, or service in that environment) should be designed to meet the needs of all people who wish to use it." (Universaldesign.ie)

As this definition suggests, it takes a considerable effort to address all users in all environments. To me, universal design is about democratisation of the design of process, to ensure inclusion.

AI and human perception, movement and cognition/emotions

AI has the potential to enhance human perception, support movement or cognition.

Image recognition algorithms can be applied in assistive technology as a form of "seeing", an

AI-powered exoskeleton could better the capabilities of the physically impared, or an AI could be used in remote therapy sessions.

Inclusion or exclusion

A fairly recent example of "discimination by AI" is when candidates for job interviews where based on an machine learning algorithm (mangler referanse)
Unfortunately, the "AI" picked applicants based on historic data of employees. This led to a heavy bias toward applicants with for instance a "western sounding" name.

Do machines understand?

I associate understanding with the assessment, utilization and operationalisation of information and context.
These are also processes that can be simulated, to some extent, by machines today, but the technology is far from perfect yet.

Guideline for Human-AI interaction

Microsoft has suggested a set of guidelines for designing for Human-Ai interaction.
Some of these are discussed below.

> Guideline 1: Make clear what the system can do.
> - Help the user understand what the AI system is capable of doing.

Alternative explanation: Explain the limitations and possibilities within the system.
The Microsoft AI guidelines are mainly concerned with context and the work to be done. This is in contrast to most HCI guidelines. These guidelines are mainly concerned with the configuration, orientation or visibility of elements in the system without explicitly mentioning context. This does not mean that those who work within the field of HCI do not take context into consideration. In Nielsen's Ten Heuristic Principles you could argue that both guidelines "Match between system and real world" and " "Consistency and standards." deals with context.

Second module

Characteristics of AI-infused systems

AI-infused systems

The term AI-infused systems is mainly used to describe systems utilizing artificial narrow intelligence. These are highly specialized systems with a very focused set of functionality , like facial recognition.

The key characteristic of these AI - infused systems are that they have the ability to learn and improve over time and that they are fueled by large data sets, and do not reveal how it actually works (black box).

In this case, the system is described as learning is when it is dynamic and designed for change. Being designed for change is not enough, it should also be designed to improve. As AI-infused systems are inherently flawed, mistakes are inevitable. The AI should therefore be able to improve based on feedback from, and interaction with the users.
This is also closely connected to the fact that these AI infused systems are based on large datasets which further is heavily reliant on input gathered from interaction with large user groups.

However, the technical implementation of how the feedback is handled and how the interaction alters the results is not exposed to the user. Thus AI-infused systems are viewed as black boxes. We do not know how they actually work or the functionality is hidden or secret.

Example of AI-infused system

The google owned online video-sharing platform youtube has a typical example of an AI -infused system - the youtube recommender system.

By using the service you give youtube new data points to improve their predictions about what to recommend to you (Learning / Improving)
Even though Google gives you an overview over what type of data they collect and in general terms explain how they rely on it to give you a personalized experience, you as a user is not really aware/knowledgeable about the inner workings of the system (Black box).
Your personalized experience on the platform is not only based on your recent movements on the platform, but it is also based on the patterns of thousands of others that have watched the same videos as you etcetera (Fueled by large datasets).

Implications for the user

One concern regarding the youtube recommender system is what type of content it presents as a recommendation to you.
Youtube is designed in such a way that it will continue to feed you closely related or even identical content to keep you on the platform as long as possible (forskning.no). Youtube will for instance avoid exposing you to new content in other genres. This is in contrast to other services , like the music streaming service Spotify who use their recommender system to encurage exposing yourself to new impressions.
An amusing thought is that the strategy google has chosen for youtube is based on an extreme interpretation of Amerisa (2019) notion that “Inconsistent and unpredictable behaviors can confuse users, erode their confidence, and lead to abandonment of AI technology”.

Human-AI interaction design

.

The main takeaway from Amershi et al. (2019) is their 18 proposed guidelines for interaction with AI. They argue these guidelines will contribute to a “better, more human-centric AI-infused systems” (Amershi, 2019: 12).

The main takeaway from Kocielnik et al. (2019) is their three design expectation adjustment techniques that prepare users for AI imperfections and result in a significant increase in acceptance.
Their findings open the way to shaping expectations as an effective way of improving user

acceptance of AI technologies.

Design guidelines

G 11 - Make clear why the system did what it did.
Youtube give you options to look into why you see certain ads but not why you see a certain video. Ads are also delivered using machine learning techniques, so If you define youtube as an advertisng platform you could argue that it adhere to this guidline, but if you look at the recommender system as an isolated service - it does not adhere to this guideline.

G4 - Show contextually relevant information. Display information relevant to the user's current task and environment.

The ads delivering service within the youtube platform is designed to provide contextual relevant information, leaving the definition of "relevant" to advertisers and the interpretation of data about the user.

How design guidelines could inspire improvement
Regarding G4 - its really hard to pinpoint what contextual relevant information is. I would argue that a feature-dense, but focused application like youtube would not benefit from displaying (even more) additional info.

Regarding G11 - I think both the user and the service owner would benefit from considering this guideline. If the system can communicate why it did what it did, the user might also be invited to suggest changes to the assumptions about the user, made by the system. This would give the system owner more detailed data about the user and the user might get a presentet with a broader selection of content.

Chatbots / conversational user interfaces

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

One definition of chatbots is that chatbots are machine agents that serve as natural language user interfaces to data and services through text or voice. (Ref Slides, 22 sept)

The perceived quality of a chatbot is closely related to its natural language processing capabilities.These capabilities determine the AI´s ability to autonomously interpret the context of an input correctly and return a meaningful output. Generative chatbots like this are experimental at best, and the potential of natural language processes are overshadowed by errors.

The chatbots we are familiarized with through customer support etc are either scripted or "trained" by adding connections between keywords and the likely intents of the user.

This coupeling of keywords with intents, to use conversation as material, is a tedious process and stays at the core of the the key challenges in the design of chatbots. Følstad and Brantzæg (2017) argue that we have to move from seeing design as a explanatory task to an interpretational task to minimize errors and faulty behaviour.

Challenges in current chatbots / conversational user interfaces

It is still early days in natural language processing and finding ways to handle the users expectations to a service might be wise.

Amershi et al. (2019) guideline 1 and 2 stresses the importance of managing users expectations:.

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.

I imagine adhering to these guidelines would result in more focused requests to the service from the user.

Third module

Human-AI collaboration

The integration of robotic technologies has led to the reimagining of robots as assets that more closely resemble interactive companions. This has led to the need for a transition of the robot's role from a tool to an interactive team member.
(Philips et. al. 2016, s 101) However, the development of human teammate-like competence (e.g., cognitive architectures, theory of mind, ability to react appropriately in dynamic team environments) is severely limited by the progress of the current state of the art in robotic technologies. (Philips et al 2016, s 101)
These contemporary and presumably temporary constructs imply that, in the near term, robots will not fully replicate their human counterparts in functionality or intelligence capacities. Yet, to meet the needs for near-term development, robots will need to possess a subset of skills that can be leveraged to perform work, not unlike the ways that working animals are utilized (Philips et al 2016, s 101)
Philips et al. (2016) give insight into the near-future stages of human-robot teaming through an examination of the different human-animal team types.
Drawing from how animal-human teams have been cultivated throughout history, robots in a robot-human team are designed to either replace or multiply (human) physical capabilities, augment and extend natural human functions to give the human physical, emotional or cognitive benefits.

An example of a robot that replaces physical capabilities is the four legged robot BigDog. BigDog is a robot designed to function essentially as a pack mule and traverse terrain not accessible by wheeled or tracked vehicles. (Philips et al 2016, s 102)

This robot is inspired by both physical form of an animal and its role in human-animal companionship.
Robotic designs inspired by nature (biomimetic designs) have brought about tremendous changes and improvements, especially concerning the dexterous and unique mobility of robots (Hancock, 2015).

As animals can provide humans with additional sensory information that can then be used by the team to make better, more informed decisions (cognition), so can a robot-companion. An example of a robot that has these capabilities are small bird-like nano-robots such as the Black Hornet, used by British soldiers in Afghanistan to replace sensing capabilities needed for scouting and reconnaissance. (Phillips et al 2016, s 112)

In the next section I will 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.

Levels of autonomy
Shneiderman (2020) introduces a two-dimensional framework of Human Centered Artificial Intelligence (HCAI) that separates levels of automation/autonomy from levels of human control. This is in contrast to Sheridan-Verplanks 1-dimensional levels, ranging from low to high (where low level of automation is when a human takes all the decisions, and high level of automation is when the computer ignores the human and acts autonomously.)
Shneiderman (2020) argues that high levels of human control *and* high levels of automation, are more likely to produce computer applications that are Trusted, Reliable & Safe (TRS).

BigDog can both receive commands and solve tasks autonomously, and be manually controlled by an operator. Much like the BigBig dog, an operator can use the control unit to pilot the drone directly or input a set of GPS coordinates for it to follow (Finsher, 2013)
Based on this (limited) information of functionality both BigDog and Black Hornet can be placed in the upper middle of the quadrants in (see figure 1). They are both capable of wayfinding on their own, but can be manually overridden at any time.

Thererfore both BigDog and the Black Hornet present similar levels of automation, whereas the application area for the robot capabilities differs (Figure 1.)

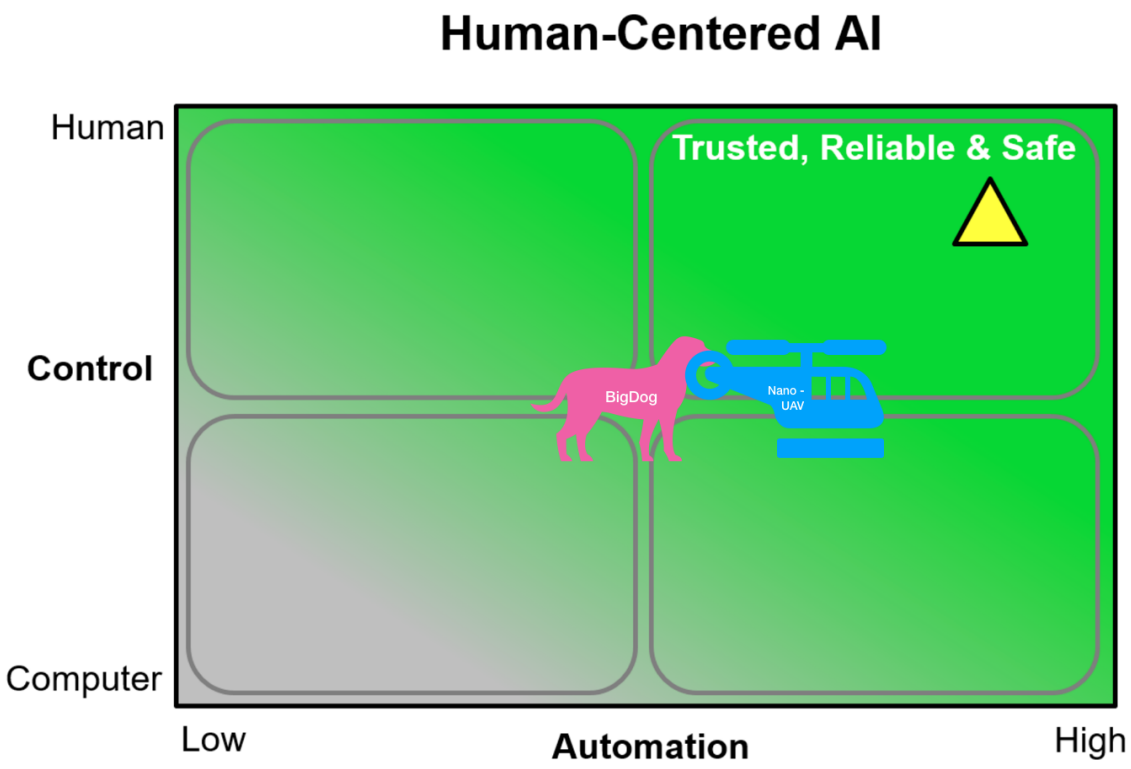

Figure 1 - suggested placement of two robots in Shneidermans 2-dimensional framework for HCAI.

(Dis)advantages

Shneiderman (2020) argues that well-designed automation preserves human control where appropriate, thereby increasing performance and enabling creative improvements.

When is human control appropriate for BigDog and the Black Hornet? This could be a question of trust. Does the operator trust the Black Hornet to find the most effective route to the target geolocation, maybe even while flying undetected? Does the Big Dog operator trust the robot to calculate that an approaching bridge can hold the weight of both the robot and its packing?

Shneiderman (2020) argues that increased autonomy should still give human operators a clear understanding of the machine state and their choices, guided by concerns such as the consequences and reversibility of mistakes.

Explainability

In both of the examples robotes used as examples above are subject to human control. The Ai component in Big Dog is to autonomously traverse rough terrain, and the Black Hornet can navigate to a fixed geoposition.

Hagras (2018) argues that users want (and benefit from) greater transparency through explainable AI (XAI) systems. An XAI or transparent AI or interpretable AI is an AI in which the actions can be easily understood and analyzed by humans. (Hagras 2018, s 29)

I recon both Black Hornet and Big Dog require little needed explainability.

Since their decisions are mostly based around how to handle/avoid physical obstacles, and modelled on animal capabilities, it is easy to imagine that visual feedback would be sufficient confirmation on decisions made by the robot.

I think this is helped, mainly by the fact that the movement closely resembles animals we are already familiarized with and their physical (mechanical) limitations.

Feedback

I received similar feedback on both module 1 and module 2. The feedback addressed missing references /inconsistency in formatting and few themes that they missed some elaboration /further explanation.

For this final delivery I have corrected the discrepancies regarding formatting of the references, focused on consistency in the layout and had another look at some of the argumentation or lack thereof.

Literature

Amershi, S., Weld, D., Vorvoreanu, M., Fourney, A., Nushi, B., Collisson, P., ... & Teevan, J. (2019). Guidelines for human-AI interaction : *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (paper no. 3)*. ACM.
https://doi.org/10.1145/3290605.3300233

Fincher, J. (2013, February). Handheld black hornet nano drones issued to U.K. soldiers. *Gizmag*. Retrieved from http://www.gizmag.com/black-hornet-nano-uav/26118/

Følstad, A., & Brandtzæg, P. B. (2017). Chatbots and the new world of HCI. interactions, 24(4), 38-42.
https://doi.org/10.1145/3085558

Grudin, Jonathan. AI and HCI: Two Fields Divided by a Common Focus. AI magazine 30, no 4 (September 18, 2009).
https://doi.org/10.1609/aimag.v30i4.2271

Hagras, H (2018) *Toward Human-Understandable, Explainable AI* : Computer 51, 28-36
https://doi.org/10.1109/MC.2018.3620965

Frankenfield, J (2020, Mars) Investopedia, Artificial Intelligence (AI) retrieved from: https://www.investopedia.com/terms/a/artificial-intelligence-ai.asp

Kaste, M. A (2020, May) Slik får Youtube deg til å se en video til (forskning.no) retrieved from : https://forskning.no/media/slik-far-youtube-deg-til-a-se-en-video-til-og-en-til/1677988

Kocielnik, R., Amershi, S., & Bennett, P. N. (2019). Will You Accept an Imperfect AI?: Exploring Designs for Adjusting End-user Expectations of AI Systems. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (paper no. 411)*. ACM.
https://doi.org/10.1145/3290605.3300641

Phillips, E., Ososky, S., Swigert, B. and Jentsch, F. (2016) Human-animal teams as an analog for future human-robot teams: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting,* Vol 56, Issue 1, pp. 1553 – 1557
https://doi.org/10.1177/1071181312561309

Pwc (2020) Fra strategi til AI retrieved from :
https://www.pwc.no/no/pwc-aktuelt/fra-strategi-til-ai.html

Science Friday (2020) The Origin Of The Word 'Robot' retrieved from:
https://www.sciencefriday.com/segments/the-origin-of-the-word-robot/

Shneiderman, B., (2020) Human-Centered Artificial Intelligence: Reliable, Safe & Trustworthy *arXiv.org*
https://doi.org/10.1080/10447318.2020.1741118

Techopedia (2020) Artificial Intelligence (AI) hentet fra :
https://www.techopedia.com/definition/190/artificial-intelligence-ai

Thrun, S. (2004) Toward a Framework for Human-robot Interaction: *Hum.-Comput. Interact. 19, 9–24.*
https://doi.org/10.1080/07370024.2004.9667338

Universaldesign.ie (2020) *What is universal design*, retrieved from:
http:/universaldesign.ie/What-is-Universal-Design/

Yang, Q., Steinfeld, A., Rosé, C., & Zimmerman, J. (2020). Re-examining Whether, Why, and How Human-AI Interaction Is Uniquely Difficult to Design: *Proceedings of the 2020 chi conference on human factors in computing systems* (Paper no. 164).
https://doi.org/10.1145/3313831.3376301