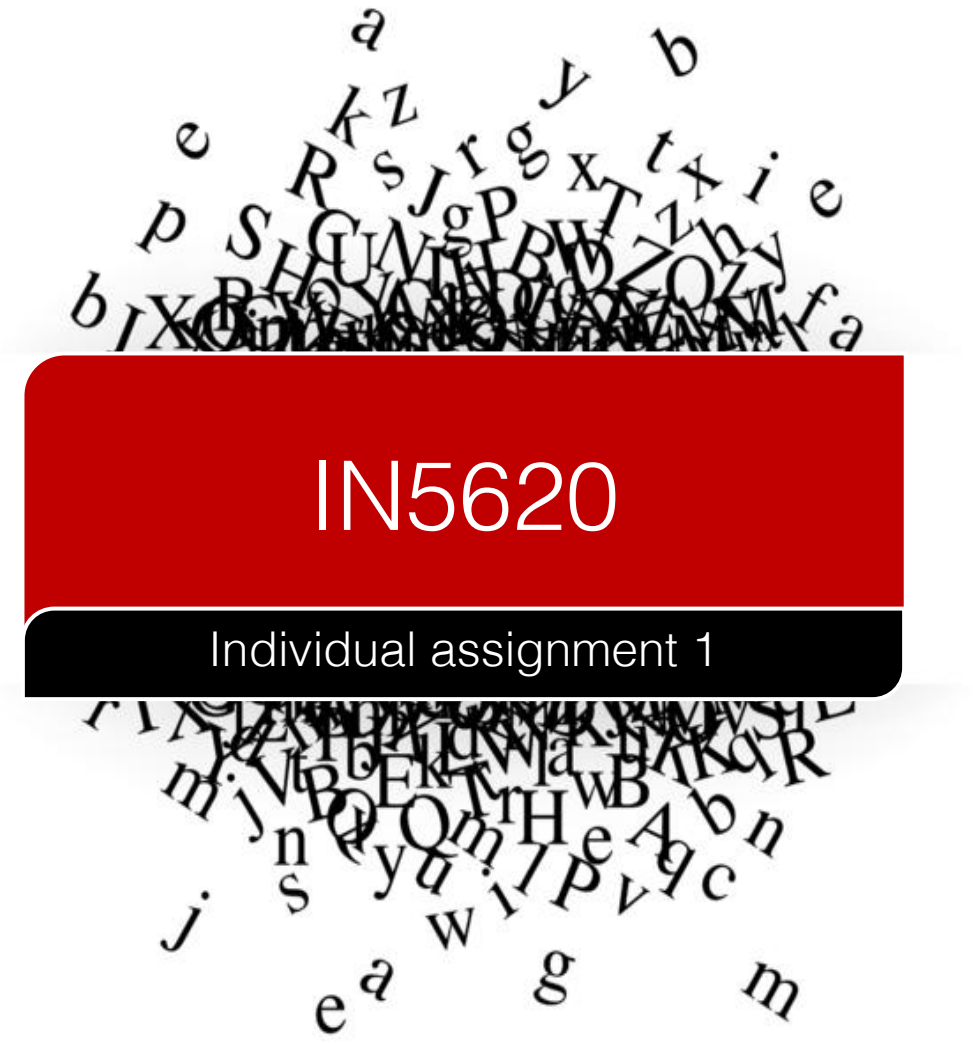


IN5620

Individual assignment 1

University of Oslo – Department of Informatics

Course: IN5620
Student: ratungla
Lecturer: Jo Herstad

Delivery: 14.09.2022

Pages: 6

Table of contents

1 Concepts, definitions and history of AI 3

1.1 History of AI 3

1.2 Definitions of AI 3

1.3 Human vs. machine autonomy 4

1.4 Article | Holmquist (2007) – Intelligence on Tap 4

1.5 Contemporary company working with AI 4

1.6 AI in popular culture 5

2 Robots and AI systems 5

2.1 How the term robot came about 5

2.2 Different definitions of the word robot 5

2.3 The relationship between AI and robot 6

2.4 Describe the interaction between human and a contemporary robot 6

3 Universal design and AI systems 6

3.1 Universal design defined 6

3.2 AI potential with respect to human perception, human movement and human cognition/emotions 7

3.3 AI - Exclusion or inclusion? 7

3.4 Do machines understand? 7

4 Guidelines for Human-AI interaction 8

4.1 Description of a Microsoft Human-AI interaction guideline 8

4.2 HCI design guidelines vs. Human-AI interaction guidelines 8

Litterature 9

1 Concepts, definitions and history of AI

1.1 History of AI

Although the concept of what we now define as artificial intelligence (AI) has been explored through the world of science fiction, popular media and through pioneers like Alan Turing during the first half of the 20th century, the term AI was first introduced at the Dartmouth Summer Research Project on Artificial Intelligence (DSRPAI) in 1956, by John McCarthy (Anyoha, 2017). The conference was held in order to further explore and discuss the possibilities of AI as a research-field (Anyoha, 2017).

1.2 Definitions of AI

Hearing the term AI, thoughts about autonomous robot armies attacking humanity or Siri never being able to understand what you try to say, might arise. As the scope and possibilities of AI and information technology continuously expands, finding one definitive all-encompassing definition could pose a challenge. Let's look at three different definitions found in a 2021 Medium.com article (Ribeiro, 2021):

"AI is the science and engineering of making intelligent machines" -John McCarthy

"AI is getting computers to do tasks that would normally require human intelligence" -Deloitte

"Artificial intelligence is a type of technology that mimics the human thought, empowering machines to act on their own and to perform functions similar to human intelligence, such as the ability to perceive, learn, reason, and act." -Jair Ribeiro

The abovementioned definitions have origins from three different domains. Whereas McCarthy's definition has a scientific basis, Deloitte is a private company with independent goals for growth and profit. Lastly, Jair Ribeiro (2021) represent a media-/news-outlet.

In all three definitions, intelligence is a common denominator. They differ when describing which systems are encompassed, ranging from machines to computers. They also define different scopes, types of systems and what those systems do. While McCarthy's definition describes machines just being intelligent, Deloitte describe computers doing tasks that normally would require human levels of intelligence. Ribeiro (2021) expands upon this, describing systems mimicking human thought and problem-solving patterns through perception, learning, reasoning and acting.

Based on these definitions, I've made an AI-definition that encapsules both AI as a scientific field and a type of technology. The new definition also expands upon AI and its ability to imitate human behavior, not just explicitly intelligence:

AI encompasses the science, engineering and technology of systems exhibiting or imitating human-like behavior and intelligence, in order to perceive, learn, reason and act.

1.3 Human vs. machine autonomy

Autonomy can be considered as a condition of thinking, and as a distinct characteristic of intelligence (Martinez, 2019). Humans and animals alike can, based on this, be characterized as autonomous. They're free to think and act upon these thoughts. As machines mostly just do what they're told, autonomy for machines can be seen as an ability of decision-making without any outside input (Martinez, 2019). This will still be by design and its original programming, but with the ability to learn and provide some output that was not originally programmed. In the same way as humans are "programmed" through genomes and environmental factors, autonomous machines have a less complex programming giving it wanted attributes and abilities to somewhat act outside of just direct control.

1.4 Article I Holmquist (2007) – Intelligence on Tap

Do you remember all those CAPTCHA-tasks you completed to prove you were not a robot? Well now, thanks to *your* impeccable capabilities in identifying traffic lights and fire hydrants, the robots have learned enough and are ready to replace you. But only at *certain, limited* tasks. In the article "Intelligence on Tap: Artificial Intelligence as a New Design Material", Holmquist (2017) highlights recent advances in machine learning through neural networks, due to increased processing power and huge amounts of available data (which may or may not have been ethically collected). Like training a dog with treats, neural networks learn through reinforcing/weighting nodes in its network when they succeed at pre-defined tasks.

This can make AI capable of solving complex problems and the recent advances is already making AI-technology and machine learning through neural networks more accessible without being an expert in the field. As a designer, "intelligence on tap" can be used in the future, but not without limitations. AIs can still not form an actual understanding of the world. Designers or whoever uses AI-generated solutions as design basis/material therefore needs to tread lightly, manage expectations and understand the current limitations of the technology for the technology to be useful (Holmquist, 2017).

1.5 Contemporary company working with AI

I've had a look at Levity.ai, a company delivering what they describe as an "out-of-the-box AI technology" for classifying documents based on set parameters (Kleinings, 2022). Clients can train neural networks with their own datasets and be provided with the necessary service in order to do said tasks. Both the service and the infrastructure is hosted by Levity for customers to subscribe to for a monthly fee. They therefore describe their services as AIaaS (AI as a service).

1.6 AI in popular culture

The 2009 science fiction movie Moon is set in space, where an astronaut is living and working in complete isolation, only accompanied by the AI-robot, GERTY. GERTY is portrayed as being able to understand and communicate natural language, while also being able to recognize and show emotion (using emojis). The emotional response from GERTY seems to be limited to reflecting the current mood of the astronaut through sympathizing. GERTY has its own pre-defined (programmed) goals, that differs from the astronaut.

Although GERTY is used to stifle the astronaut's jarring loneliness, the longing for actual human-to-human interaction is evident through the depiction of deteriorating mental health. The movie highlights the importance of human-to-human/sentient interaction and illustrates how AI companions might not suffice when trying to cope with loneliness.

2 Robots and AI systems

2.1 How the term robot came about

Looking at the etymology of the word "Robot" we can find the first English citation of the word back in 1839, with a meaning derived from the Czech word *robota*, meaning forced labour/slavery (Interesting Literature, 2022). The current understanding of the word robot however, was first seen in the Czeck play *Rossum's Universal Robots,* featuring a factory making "artificial people" and having them serve and work for humans (Interesting Literature, 2022). Nobody enjoys being slaved around and the robots uproared, as you can imagine.

2.2 Different definitions of the word robot

In the same way AI has a multitude of different definitions, there can be found a plethora of different definitions of the word "Robot". Let's have a closer look at two such definitions:

"Actuated mechanism programmable in two or more axes with a degree of autonomy, moving within its environment, to perform intended tasks." (ISO, 2012)

"Any automatically operated machine that replaces human effort, though it may not resemble human beings in appearance or perform functions in a humanlike manner." (Moravec, 2022)

The ISO 8373:2012-definition (ISO, 2012) defines robots as a programmable and movable *mechanism*, with the intent of solving tasks. Encyclopedia Britannia (Moravec, 2022) has a slightly different definition, defining robots as automatically *machines* that replaces human efforts in order to solve tasks, although not necessarily with human-like features and characteristics. The latter definition does not specify movement as an

ability. I've created my own definition that might be a bit vague, trying to encompass all types of programmable and physically movable mechanisms that are able to perform taskes based on input:

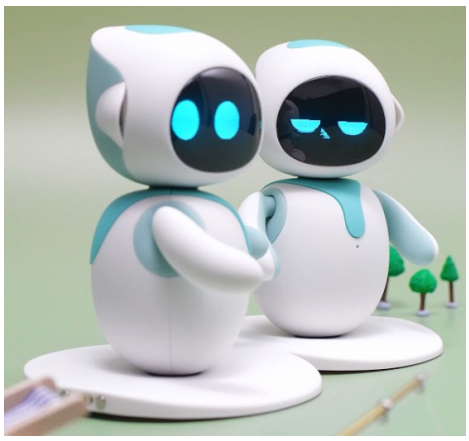
Robots are programmable, sensory and movable physical mechanisms, designed to perform intended tasks within a given environment.

2.3 The relationship between AI and robot

AI technology is defined as using human-like intelligence in order to solve tasks, while robots are defined as mechanisms performing programmed tasks using "human effort". Robots act through direct programming and sometimes with human-like movement, instead of autonomous actions or intelligence. A robot *can* have AI, but it doesn't *have* to in order to still be classified as a robot. Based on previously mentioned popular culture depicting humanities doom due to robot uprisings, huAIs with legs might not be the greatest idea?

2.4 Describe the interaction between human and a contemporary robot

Meet Eilik, described as "Tech with heart" by it's creators (Energize Lab, n.d.). This robot is minted for the consumer market as a social companion, responding to touch, movement, sound and its own position within a physical space. It mimics human emotion through digital "emoji-like" facial expressions, and through it's movement. It can spin 360 degrees and move it's head and arms.



Eilik seems to have interchanging moods, inviting users to interact by using a language that are recognizable to us all, visible emotion. If it's sad you can pick it up, pat it on the head or similar and it will respond to the sensory input and output a response.

3 Universal design and AI systems

3.1 Universal design defined

The Norwegian Digitalisation Agency (Digdir) defines universal design as:

"Designing, or accommodating, the main solution with regards to physical conditions, so that the solution may be used by as many people as possible, regardless of disability" (UUtilsynet, n.d.)

Universal design as per Digdirs definition (n.d.) is about designing for inclusion and availability regardless of individual uniqueness, variability or physical/cognitive abilities. The goal is to make systems that are accessible and usable for as many individuals as possible and not hindering use, which in sum should benefit all users, not just the minority.

3.2 AI potential with respect to human perception, human movement and human cognition/emotions

Every single human being is unique regarding capabilities, interests, culture, emotion, wants and needs. The current "one size fits all"-paradigm of software development might become a thing of the past, due to the future of AI-capabilities. Designing systems for every single unique user, with tailored user interfaces and functions is not feasible or possible today. There is however huge potential for a new genre of "generic AI-software", adapting to you, your capabilities, emotions and needs, with dynamically changing UIs and functions.

For students, maybe the "Focus"-mode of current cell phones/PCs could identify patterns of what triggers them to be distracted and then remove/hide these UI-elements/functions when a deadline is closing in, making it easier to work efficiently. Or maybe it could even understand how someone learn most efficiently and tailor/present content in a way that works for them. Maybe a boring text could be made richer by automatically generating or adding relevant images, videos or similar, in the same way social media algorithms are designed to keep you continuously engaged.

3.3 AI - Exclusion or inclusion?

As Holmquist (2017) described the future of AI as a design material (seen in task 1.4), he also highlighted some of the risks associated with neural networks and machine learning. These networks learn from existing user data, which will most likely also contain current biases, posing a risk of possibly reinforcing existing exclusionary tendencies, bias and prejudice in society.

There're already multiple examples of inherited AI-bias in real-world. US hospitals used an algorithm on more than 200 million patients, trying to identify which patients would most likely need extra medical care. In 2019 it was identified that the algorithm heavily favored white patients over black patients, not because of parameters separating by skin color, but due to the groups healthcare cost history (Shin, 2020). To counter inherited bias and increase inclusiveness, AI-systems and neural networks could also be designed to identify imbalance or possible bias and flag these in datasets.

3.4 Do machines understand?

In the context of the WCAG 2.1 principles (W3C, 2018) and the guidelines for Human-AI-interaction, I would argue that the words "understand" and "understanding"

represent the ability to not just perceive or know something (like an output from an interface), but rather being able to categorize and associate whatever is perceived as part a context, connected to a larger system of knowledge and as part of a complex subjective understanding of how the world works.

Based on this I'll argue that machines and AI in their current state (and in their walled-off garden) does not currently have the capability to understand something, but rather holds the ability to sense and perceive and know something about what was perceived. I do believe this is subject to change in the future.

4 Guidelines for Human-AI interaction

4.1 Description of a Microsoft Human-AI interaction guideline

Guideline 16: Convey the consequences of user actions (Amershi et al., 2019).

Make sure to convey the user how and if an interaction with the AI-system might impact system behavior in the future. If the AI-system detects behavior or data that may have been unintentional or not representative of the target it learns from or interacts with, an option should be given to delete or invalidate the data in order to not cause unintended unwanted behavior.

4.2 HCI design guidelines vs. Human-AI interaction guidelines

When taking a look at Jakob Nielsen and Rolf Molich's Ten User Interface Guidelines (building upon Ben Shneiderman's "eight golden rules"), one can see that these heuristics does have some similarities to the Human-AI interaction guidelines from Microsoft (Wong, 2020).

While HCI design guidelines are more generalized and applicable to the design of every type of interface, the Human-AI guidelines are specialized for this specific type of technology. This is specialization is highlighted by the Human-AI guidelines being divided into different categories or "stages", such as "initially", "during interaction", "when wrong" or more the more general "over time".

One of the ten user interface guidelines from Nielsen & Molich, "Consistency & Standards", recommends designs that implement a recognizable and established design language, similar to other contemporary interfaces (Wong, 2020). To me, this translates to following current norms within design and interaction. The human-AI interaction guideline "Match relevant social norms", is similar in that it recommends an experience that is delivered in a way users would expect, given their context and societal norms (Amershi et al., 2019).

Litterature

Amershi, S., Vorvoreanu, M., & Horvitz, E. (2019). *Guidelines for human-AI interaction design*. Microsoft. Retrieved September 13, 2022 from https://www.microsoft.com/en-us/research/blog/guidelines-for-human-ai-interaction-design/

Anyoha, R. (2017). *The History of Artificial Intelligence*. Harvard University. Retrieved September 12, 2022 from https://sitn.hms.harvard.edu/flash/2017/history-artificial-intelligence/

Energize Lab. (n.d.). *Eilik*. Engerize Lab. Retrieved September 13, 2022 from https://energizelab.com/consumerview/eilik

Holmquist, L. (2017). Intelligence on tap: artificial intelligence as a new design material. *Interactions (New York, N.Y.)*, *24*(4), 28-33. https://doi.org/10.1145/3085571

Interesting Literature. (2022). *The Curious Origin of the Word 'Robot'*. Interesting Literature. Retrieved September 13, 2022 from https://interestingliterature.com/2016/03/the-curious-origin-of-the-word-robot/

ISO. (2012). *ISO 8373:2012(en)*. ISO. Retrieved September 13, 2022 from https://www.iso.org/obp/ui/#iso:std:iso:8373:ed-2:v1:en

Kleinings, H. (2022). *What is AIaaS? Your Guide to AI as a Service*. Levity. Retrieved September 12, 2022 from https://levity.ai/blog/aiaas-guide

Martinez, R. (2019). Artificial intelligence: Distinguishing between types & definitions. *Nevada Law Journal*, *19*(3), 9. https://scholars.law.unlv.edu/cgi/viewcontent.cgi?article=1799&context=nlj

Moravec, H. P. (2022). robot. In *Encyclopedia Britannica*.

Ribeiro, J. (2021). *These Are the Best Definitions of Artificial Intelligence You Can Read Today.* Medium.com. Retrieved September 12, 2022 from https://medium.com/swlh/these-are-the-best-definitions-of-artificial-intelligence-you-can-read-today-7c53c0e38584

Shin, T. (2020). *Real-life Examples of Discriminating Artificial Intelligence*. Towards Data Science. Retrieved September 13, 2022 from https://towardsdatascience.com/real-life-examples-of-discriminating-artificial-intelligence-cae395a90070

UUtilsynet. (n.d.). *Regulations*. UUtilsynet. Retrieved September 13, 2022 from https://www.uutilsynet.no/english/regulations/278

W3C. (2018). *Web Content Accessibility Guidelines (WCAG) 2.1*. W3C. Retrieved September 13, 2022 from https://www.w3.org/TR/WCAG21/

Wong, E. (2020). *User Interface Design Guidelines: 10 Rules of Thumb*. Interaction Design Foundation. Retrieved September 13, 2022 from https://www.interaction-design.org/literature/article/user-interface-design-guidelines-10-rules-of-thumb