Individual Report – IN5620



Rifat Naim Islam

1.1. History of AI

The term “Artificial intelligence” was first defined by John McCarthy in 1956. This development took place through a workshop at Stanford University in the same year. The basis of the claim was that the central property of human beings, intelligence can be so precisely described that that it can be simulated by a machine (McCarthy , 2006). However, as the years progressed, we began to see the emergence of other definitions.

John McCarthy defined AI as “the science and engineering of making intelligent machines “. In his definition, there can be seen a focus on machines that are able to simulate human intelligence (Grudin, 2008). However, Bratteteig and Verne (2018) define it as “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”

Grewal defines it as Artificial Intelligence is the mechanical simulation system of collecting knowledge and information and processing intelligence of universe: (collating and interpreting) and disseminating it to the eligible in the form of actionable intelligence.

Thereby, my own definition of artificial intelligence would be “Artificial intelligence refers to systems that are able to process knowledge and information in a manner which simulates human intelligence”

The article “Some Brief Thoughts on the Past and Future of Human-Robot Interaction” (Dautenhahn, 2018) is a fascinating article revolving around Human Robot interaction (HRI) where the writer talks about the things she is excited and optimistic about in this field, while also mentioning parts of it she is worried about

The article writes in detail about how HRI is not only about robots and how they affect humans, but how much a\of a interdisciplinary field it is. Taking into account fields such as psychology and anthropology, there are a lot of developments happening gradually. This development raises concern for Dautenhahn in that she is sceptical of how this may seem over time. However, my question to this article would be if Dautenhahn could refer to a specific example that raises her concern for developments inside the HRI field?

A company that works with AI is c3.ai, which present AI as a service they provide. The company deals with providing enterprise solutions to consolidate data, build AI models and generate predictions. Therefore, while they do closely work with AI , it is catered to customers in way as more of a developmental platform and adjusting the type of AI services provided based on the requirements of the customers.

A video game that tackles the concept of interaction with AI is Detroit : Become Human. The game is set in a world where AI have advanced in a manner where they are able to mimic human behaviour, emotions and appearances. While human beings and AI are seen to be co-existing in this world, the dynamic between is that of a slave and a master. Despite being highly intelligent machines who have shown to have the ability to develop real feelings, human beings mostly mistreat their machines. The game also shows panic can rise amidst the population once the robots become self-aware.

1.2 Robots and AI

The first instance of the word robot being used can be seen is through the Czech author Karol Capek. The word was used in a play and was used with the meaning “forced labor” (MIT Press reader, 2021). The original meaning of the word is in line with how industrial robots are used in the current world, where robots are used to automate work activities which were previously done by human beings.

The insititute of electric and electronical engineers defines robot as “as an autonomous machine capable of sensing its environment, carrying out computations to make decisions, and performing actions in the real world.” (IEEE, 2020)

On the other hand, Brittanica dictionary defines robot as “any automatically operated machine that replaces human effort, though it may not resemble human beings in appearance or perform functions in a humanlike manner.” (Britannica, n.d.). Thereby, my own definition would be that robot refers to a machine that is able perform actions and simulate human effort without a will of its own.

From these definitions, we can notice that there lie some differences between robots and AI. The key difference would be that of intelligence, as the definitions of AI indicates that they have a higher executive function than robots. This would also indicate that AI would be able to carry out functions that a robot may not although are made to simulate human activity,

An example of a contemporary robot would be the Roomba vacuum developed by iRobot. The robot uses its sensors to autonomously detect dirty spots on the floor, and thus carrying out the function of cleaning and vacuuming the house. Moreover, the sensors allow to have a sense of its surroundings and presence of obstacles and are thus able to navigate throughout the house. The Roomba replaces human effort, while it is able sense its environment and carry out computations to make decisions as the IEE definition indicates.

1.3 Universal Design

Universal design refers to the design of products and environments to be usable by all people, to the greatest extent possible, without the need for adaptation or specialized design (Mace, 1985). Universal design ensures that the designed product can be accessible by any person, rather than developing customizing products to meet the needs of individual user-groups. Therefore, the spectre of inclusion is much broader than other design forms.

There lies a lot of potential in using AI to improve the quality of life for people with disabilities or cognitive impairments. This can be especially in our daily lives in the present world, where AI assistants such as Siri and Cortana have changed quite a lot on people interact with a phone. A blind person is able to navigate through their phone and make calls

simply by calling out Siri. Similar technology is also implemented in entire homes such as Google Home, which also allows a person with locomotive disabilities navigate and control elements of their home through the use of AI. Both technologies make it possible that more users are able to use smartphone and it also makes human navigation inside houses easier,

Through these examples, we see how AI can be has high potential for including more users. However, it also has the potential of excluding people which can be seen in a study by Hundt et al (2022) where stereotypes, bias and discrimination can manifest in AI through machine learning. The robots were programmed with a popular AI algorithm and were given different images to sort through and respond questions related to the pictures. These virtual robots responded to associating the word “criminal” with black men while words such as “homemaker” were associated with black and Latina women.

“Understand” and “understanding” are two words which may sound similar, but bear different meanings. While “understand” means being able to perceive and comprehend a situation while understanding is more related to the concept of reasoning. While machines heave shown that they can “understand” as displayed by a Roomba when it navigates according to its surroundings, its degree of understanding is still not at the level of human beings. Machines are not quite able to process human emotion and gestures in a level that would warrant for it to labelled understanding.

1.4 Guideline for Human AI interaction and HCI

One of Microsoft’s guidelines for human-AI interaction are guideline 13 “Learn from user behavior” which revolves around how the user experience can be personalized through learning from the actionsof the user. This can be widely seen in the Youtube and Google search bar which can increasingly be able to predict quite accurately what you are about to search based upon your use. This is also how a YouTube feed is able to display videos that aligns more with one’s interests.

There are several similarities between HCI design guidelines and Human-Interaction guidelines, one of these being that both guidelines focus on inclusion and having the user in its focus. The comparison is done by looking Microsoft's 18 guidelines for human AI interaction and Nielsen and Molich's 10 HCI design guidelines (Nielsen, 2020). These guidelines also have a significant difference in that HCI has a larger focus on the users and designing solutions with user's needs in mind while Microsoft's guidelines focus more on the behavior of the system over a period of time

2.1 Characteristics of AI infused systems

"Systems that have features harnessing AI capabilities that are directly exposed to the end user" are defined as AI-infused systems (Amershi et al, 2019). While this is an area where the user receives a higher degree of assistance from the system, it comes with its own set of challenges. The systems may provide the user a "false positive" or a "false negative" as a result of the challenges, as Amershi et al (2019) highlights. These challenges can stem from the AI's automatic nature of functioning, which makes the systems difficult to control and predict. Therefore, Amershi et al (2019) identifies traits such uncertainty, inconsistency and automated personalization which collectively leads to unpredictable behaviour for the users (Amershi et al, 2019; Yang et al, 2020)

However, as mentioned in the first lecture of module 2, AI-infused systems have several key characteristics, such as black box, improving, learning, which are fueled by large data sets (Følstad).

Kocielnik et al (2019) points out the probabilistic nature of these systems allows them to process natural language and possess the ability to predict behaviour. The accuracy itself of these services may not always be optimal due to it being based off probability. The systems do not necessarily have to solely be based off of probabilistic operations. Yang et al. (2020) classifies AI infused systems into four different types, namely, 1) probabilistic systems , 2) adaptive systems, 3) evolving probabilistic systems and 4) evolving adaptive system. These classifications have their own challenges, which are intertwined with the key characteristics earlier identified.

Voice assistants such as Google Assistant, Siri, Cortana are all examples of AI-infused systems that are widely adopted by the users. These assistants can process natural speech or text and react to the requests made. Through the communication between the users and the assistants, they are also able to learn details about the users for example setting a nickname. In addition to displaying the learning trait, it also uses the improving trait as they are able to provide personalised suggestions based upon the data it has gathered about the user, I

2.2 Human-AI interaction design

Amarshi et al. (2019) evaluated past studies through a number of design recommendations based on AI, and developed 18 design guidelines for human-AI interaction based on them. These guidelines are developed with the purpose of addressing the challenges that users face in terms of interacting with AI. While the guidelines are not to be taken as rules set in stone, they can function as a standard for the development and maintenance of AI. Thr goal being the guidelines ultimately leading to AI that is more user-friendly and superior.

Kocielnik et al (2019) studies the impact of user expectation on the general impression of AI systems. The study revealed a user's perception of how well functioning a system is impacted by their preconceived expectations of the system. User experience was observed to be improved upon the users being informed on the capabilities of the system. The conclusion being the managed expectations allows for a better user-experience, as the users are then more likely to accept the limits of the system

Smartphone assistants
Amershi et al (2019) presents 18 AI design guidelines, of which two of them are :

G4: Show contextually relevant information
G18: Notify users about changes

The identified smartphone assistants of Cortana, Siri and Google Assistant all adhere to the guideline 4. The AI assistants process the requests and attempt to provide results that match the user's request and task. This can be seen when asking for restaurant recommendations,

the results display the restaurants near your current location. Thus, the interaction remains contextually relevant for the user and providing a proper standard of user experience. However, the assistants deviate from the guideline 18 as it provides no update to the user about its improved capabilities. In Cortana's earlier days, the AI assistant could not formulate a proper response to the prompt "Tell me a joke". However, as the yuears passed by the assistants started having the ability to process more complex sentences from the users, as well as being able to crack a joke here and there. This development has happened gradually and it has been up to the user to find the development, in most instances.

Summary of Bender et al. (2021) argument

The paper discusses the question on the necessity of larger language models in comparison to the associated costs and risks. The authors of the paper examine the environmental costs, economic costs and opportunity cost. Potential dangers of using large training data which contain abusive language are raised, as this may lead to systems that use existing stereotypes and spread extremist ideology . Systems such as these bear a large potential of impacting individuals and society as a whole. Therefore, the paper concludes with solutions which are more cautious with the given training data, in an attempt to mitigate the earlier identified risks.

2.3 Chatbots

There are several challenges that arise with the design of chatbots. One of the risks that are highlighted by Kocielnik et al (2019) and Amaersi et al (2019) are the expectations of an AI-infused system affecting the user experience. This applies to chatbots in the sense that user exepectations may be of a high functioning system that is able to respond to whatever the user says. However, the complexity of natural language processing leads them to not be able to interpret the user's request accurately at all instances. This can stem from factors such as language accent, environmental noise or the speech of speech itself, and thus user expectation not being met. If a chatbot is designed with speech from a certain part of the world in mind, and then used in another context it has not been trained for – the possibility of user expectations not being met remains high.

Adherence to guideline G1 and G2

Amershi et al (2019) defines the guidelines G1 and G2 as follows:

G1: Make clear what the system can do
G2: Make clear how well the system can do what it can do

Adhering to these guidelines come with a high possibility of resolving the problems with current chatbots. If it is clear what a system can do, then the users know how to interact with it, and manage their expectations accordingly. The G2 guideline allows the users to comprehend that chatbots come with their own restrictions, and it may fail to execute some of the tasks requested. For example, a user receiving a generic answer to a complex question would understand that there are limits to what the chatbot's technology.

Endringer

Upon receiving feedback I have made some of the formulations more precise and used the syllabus in a bit more effective manner. This has helped shape the overall report in a better way.

References

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/en-us/research/uploads/prod/2019/01/Guidelines-for-Human-AI-Interaction-camera-ready.pdf

Bender, E. M., Gebru, T., McMillan-Major, A., & Mitchell, M. (2021). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?. In Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (pp. 610-623). ACM. https://dl.acm.org/doi/pdf/10.1145/3442188.3445922

Bratteteig, T., & Verne, G. (2018). Does ai make PD obsolete? *Proceedings of the 15th Participatory Design Conference: Short Papers, Situated Actions, Workshops and Tutorial - Volume 2*. https://doi.org/10.1145/3210604.321064

Encyclopædia Britannica, inc. (n.d.). *Robot*. Encyclopædia Britannica. Retrieved September 20, 2022, from https://www.britannica.com/technology/robot-technology

Enterprise AI. C3 AI. (2022, August 16). Retrieved September 20, 2022, from https://c3.ai/

Dautenhahn, K. (2018). Some brief thoughts on the past and future of human-robot interaction. *ACM Transactions on Human-Robot Interaction*, *7*(1), 1–3. https://doi.org/10.1145/3209769

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/en-us/research/uploads/prod/2019/01/chi19_kocielnik_et_al.pdf)

Mace, R. (1985). Center for universal design NCSU - home. Retrieved September 20, 2022, from https://projects.ncsu.edu/ncsu/design/cud/

McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (2006). A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence, August 31, 1955. AI Magazine, 27(4), 12. https://doi.org/10.1609/aimag.v27i4.1904

The MIT Press Reader. (2021, January 14). *The Czech play that gave us the word 'robot'*. The MIT Press Reader. Retrieved September 20, 2022, from https://thereader.mitpress.mit.edu/origin-word-robot-rur/

Nielsen, J. (2020). *10 usability heuristics for user interface design*. Nielsen Norman Group. Retrieved September 20, 2022, from https://www.nngroup.com/articles/ten-usability-heuristics/

Grudin, J. (2009, Winter). AI and HCI: two fields divided by a common focus. *AI Magazine*, *30*(4), 48+ https://link.gale.com/apps/doc/A216267417/AONE?u=oslo&sid=bookmark-AONE&xid=d85837d1

Grewal, P. D. (2014). A critical conceptual analysis of definitions of artificial intelligence as applicable to computer engineering. *IOSR Journal of Computer Engineering*, *16*(2), 09–13. https://doi.org/10.9790/0661-16210913

Hundt, A., Agnew, W., Zeng, V., Kacianka, S., & Gombolay, M. (2022). Robots enact malignant stereotypes. *2022 ACM Conference on Fairness, Accountability, and Transparency*. https://doi.org/10.1145/3531146.3533138

Roomba® robot vacuums. Roomba® Robot Vacuum Cleaners | iRobot®. (n.d.). Retrieved September 20, 2022, from https://www.irobot.com/en_US/roomba.html

Yang, Q., Steinfeld, A., Rosé, C., & Zimmerman, J. (2020). 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). https://dl.acm.org/doi/abs/10.1145/3313831.3376301