1. Module One

1.1 AI and interaction with AI

1.1.1 History of AI

In 1956 the term Artificial Intelligence was initially used by an American Mathematician and logician John Mc Carthy in a workshop at Dartmouth Conference thus he is also known as father of AI as he was the first to introduce and give an idea about AI.[1]

1.1.2 Definitions of AI

Carthy defined AI as "the science and engineering of making intelligent machines." Every aspects of learning or any other feature of intelligence can be principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves.[2]

Minsky defined AI as "the science of making machines do things that would require intelligence if done by men"[3]. "A machine with the general intelligence of an average human being. At that point the machine will begin to educate itself with fantastic speed."(Jonathan Grudin, 2009 / Time Magazine, November 20, 1970)

"Artificial intelligence (AI) is the simulation of human intelligence processes by machines, especially computer systems. Specific applications of AI include expert system, natural language processing, speech recognition and machine vision."[4]

---

[1] https://www.javatpoint.com/history-of-artificial-intelligence

[2] https://www.dartmouth.edu/~vox/0607/0724/ai50.html

[3] https://www.britannica.com/biography/Marvin-Lee-Minsky

[4] https://searchenterpriseai.techtarget.com/definition/AI-Artificial-Intelligence

1.1.3 My definition of AI

The term AI is combination of two words artificial and intelligence, where artificial means man made and intelligence meaning the ability to acquire and apply knowledge and skills. So simple we could say AI is a manmade machine which is capable of learning and obtaining skills and knowledge. AI is a board topic, but according to my point of view on the basis of the given definitions above, AI is a machine that learns, generate and be able to perform those task that requires human intelligence automatically. Learning can be learning human behaviors or universal phenomena that would for example help in weather forecast. Automatically in the sense that it implies what it had learn and behave, auto driving is an good example of AI performing task.

AI is an super version of human brain, what our brain is capable of doing, is implemented in AI. As a normal person, we might not have explored or have lacked to fully use all functionality of brain for example memory storage, recalling, recognition, logical reasoning, decision making and many more. Thus AI illuminates such incapability of humans for ease of human.

1.1.4 Company working with AI

Tesla is a company that has been producing electric vehicles especially cars, focusing on auto driving. The company aims to lunch fully autonomous cars in future. They have been using "thinking" algorithms, sensors, navigations, that allows auto pilot. AI is therefore used for safety and convenience with focus on minimizing or nullifying road accidents.

1.2 Robots and AI systems
1.2.1 History of Robot

It seems to be very difficult to find out who exactly coined the term Robot and it has been used in different plays and comic, The name of Russian-born American science-fiction writer Isaac Asimov appears the most, as he coined the term *Robotics* in his book which was collection of his short stories ***The complete Robot*** written in between 1939 till 1977 where he presented three laws of robot: It must not injure human beings, follow human orders unless it would conflict first one and protect its own existence keeping in mind first and second law. Though the word robot was just found written in papers, the first ever commercial, digital and programmable robot was built

by George Devol in 1954. The robot was given name 'Unimate' was supposed to move hot pieces of metal.

1.2.2 Definitions of Robot

Robot is "a machine that can perform a complicated series of tasks automatically".(Oxford Dictionaries)[5]Thus robot are automatic machine that are supposed to one or more task without human involvement.

According to Merriam Webster 'Robot is 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)". [6]

By the definition robot focuses on the task of robot that is it supposed to do human activities without human intervention but with greater efficiency and effectiveness.

1.2.3 My definition of Robot

Robot are the machine that does repetitive and complex job with greater effectiveness and efficiency automatically. Thus robots are task oriented, auto mechanism that are either programmed to one task or more simultaneously or one at a time for industrial purposes that contribute to large production.

1.2.4 Relation Between AI and Robot

Robot is not different from AI rather it is a sub-topic that can be considered under the umbrella term AI. Robot is an artifact that does specific function that it is programmed for. Robot can have AI as a part of its programming, but we cannot say robot a AI, it's an example or platform for application if AI.

---

[5] https://www.oxfordlearnersdictionaries.com/definition/english/robot

[6] https://www.merriam-webster.com/dictionary/robot

1.2.5   Physical robot

Waiter Robot, these kind of robot are found in different restaurants of the world, where it serves food from kitchen counter to your table. It mostly famous in Japan and China. It has navigations if all the tables so it moves around following that navigation. It has few button intact in it that commands it to move. It stops in case if there is some hindrance on its way.

1.3   Universal Design and AI systems
1.3.1   Definition of 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. The universal design is simply a design that is applicable in whole universe that cope with the physical, geographical and intellectual difference, in other word suitable for all.

The 7 main principles of Universal design are :
Equitable use: People of all capability could use the design.

Flexibility: The design should accommodate with wide range of individual preferences.

Simple and intuitive use: Must be easily understandable by all

Perceptible information: The information provided must be properly understood with the help of text, images, videos, voice notes, sound, movement or combinations of these.

Tolerance for error: Must be able of minimizing or avoiding error, in case of unintentional error must have measures to recover.

Low physical effort: The design must be efficient, comfortable for use and with less feeling of fatigue.

Size and space for approach of use: The design must have suitable size and space for reach, manipulation and use.

1.3.2   Potential of AI

We could possibly develop something which gives it reasoning and explanation. As human perceives new things and corelate it. AI could make some similar simulations where it identifies

new things. It could be face reading , image fetching, recognition and recalling. Or even give its own point of view on the given text or summarize on its own.

1.3.3 Understand and understanding

In WCAG 2.1 the principle understandable refers to the information and the operation of the users interface must be understandable by the users. Any text in the interface must be readable, button in the interface must be easy to interpret what it does. The users must be able to navigation at which phase they are , must be able to either go back or be aware of upcoming phases. Whereas in HCI AI-interface guideline, the term understanding is used referring to the system that must understand its users behaviors. The system must be able to learn from the past action of the users and use them as important data to facilitate the users in future. Thus this answer the question "Do machine understand?" The AI integrated system are built to understand users and their needs.

2. Module Two

2.1 AI-infused systems

2.1.1 Characteristics of AI-infused systems

The characteristics of an AI-infused system might differs according to different application and might varies on the basis of user requirement. But in general a system with AI must be accurate, adoptable, transparent. It must adopt to the users' behavior, for example when it comes to recommending a movie in Netflix, the system must learn what are the preference of the users. It should have a technique of self-learning as technology is dynamic, software keep evolving whereas users action keep wavering. A system must have a quality to grasp these and upgrade by itself. Thus a system must focus on high recall rather than high precision to achieve higher level of acceptance and accuracy of the system (Kocielnik et al. (2019).

The another essential feature an AI infused system must have is accuracy. The output that's being given by the system must be precise and correct. It also includes the processing time and reliability. AI provides feedback to any queries or input but the feedback must be from a authentic source, useful and meaningful. For example a security system takes voice and face recognition as a controlling method, it must have highly concrete recognition pattern that's gives errorfree results.

AI system is capable of doing several thing without even notifying user of those. But instead AI must be able to explain its limitation what it can do and what it cannot. Thus transparency in the system is another characteristic, it must be able to explain any action it takes and how it is functioning or how it came to a conclusion. The term Explainable AI came into exitance after very long time AI was termed. Thus a system must generate explanations, support its outcome with relevant reasonings.

One the other hand, there are several positive and negative rummers about AI since a very long time. Since it is just a machine and several has also experienced AI not performing as intended. Sometime AI failed to provide faultless feedback in some cases. It might respond differently to the same text input over time. "AI-infused systems may demonstrate unpredictable behaviors that can be disruptive, confusing, offensive, and even dangerous." (Amershi et al., 2019). So far AI is still developing, we can presume such drawbacks would be illuminated.

2.1.2 Example of AI-infused system

Facebook, is one of the application that is used by almost more than 5% of world population. It has been using AI the most. Even though we might uses Facebook as social network platform. It is remotely making use of AI to predict what we are posting, writing. It has developed very efficient picture recognition system, that recognize your picture and suggest if you might want to tag yourself to the picture where you are present, its recognizes the faces on the picture you have uploaded and tag people from your friend list automatically. It also learns form user behaviors and suggest several content, videos or advertisement. These days you do not need to like or subscribe a page or channel , based on the click, time you stayed watching a certain video is used to recommend similar content. Lately, we can see if you use some emojis or letter it can learn the sentiment behind it and emphasize it, for example when you use heart emoji, you can see heart shaped balloon floating in your screen, when you comment the word "congratulation" it changes its color to red automatically without you enabling these features. Thus, Facebook possess the features of accuracy and adaptation.

2.2 Human-AI interaction design

2.2.1 Take-aways from Amershi et al. (2019) and Amershi et al. (2019)

In Amershi et al. (2019), he has undergone through a 150 recommendations from multiple sources comprising them to 20 and evaluate them by the uses of detailed user study, heuristic evaluation and expert evaluation. And introduced 18 guidelines as contribution to HCI, with the motive to helps design and evaluate AI-infused system so that people can understand, trust and can engage with effectively.

Similarly Kocielnik et al. (2019), has focused on improving user satisfaction and user acceptance of AI infused system via expectation adjustment techniques to meet the expectation of the users instead of deception as used in marketing or in-depth involved understanding shaping user mental models as used in intelligible AI works. The technique is developed to make end users understand the general mechanisms in which people learn new information with an ease despite engaging end-users in complex understanding of underlying algorithms. This article also highlights that high precision rather than high recall of a system performing at the same level of accuracy can lead to much lower perceptions of accuracy and decreased acceptance.

2.2.2 Facebook and AI Design Guidelines

In my opinion, Facebook adhere Guideline3 "Time services based on context" and Guideline13 "Learn from user behavior". Facebook has a several features that is time services based, the notification pop up when there is some picture you posted a year ago or more. It evens makes a shot video for your friendship anniversary with pictures , messages, quotes, common likes etc. Facebook learn from user behavior , as it is an expert at personalizing experience on the basis of user's passed action. It evens notice the search we did in google or other web browsers, as a result advertisement of related search appears on Facebook screen.

Whereas deviates Guideline 11 "Make clear why the system did what it did", though Facebook has several features and implementation, it does not explains why it did certain activities, it suggest us the advertises but how it knew that that user looking after, we cannot find logical reasons behind such actions. Similarly, Facebook seems to violate Guideline 12 "Remember recent interactions"

as it cannot tell us which post we were looking or reading at before sometime, which pages or profile we visited.

2.3   Chatbots / conversational user interfaces

Chatbot is a automated communication platform where natural language processing system is highly used. A system communicates with human as a human replication. Several business organization are using chatbot to answer the frequent query from the customers, some of them just give information that are preset where as some are highly capable of understanding all your queries, process your queries and provide relevant feedback or event perform some task.

2.3.1   Challenges in the design of chatbots

The challenges in design of chatbot can be failing to bridge the gap between user expectation and system operation in other word understanding user intent(Luger, E., & Sellen, 2016). It is very difficult to make a chatbot understand about different type of users as they might uses different language, words to express single thing. As a human we can notice the emotion of other human during conversation, if he is sad, happy or angry but how a chatbot will understand such human emotions. The another challenge can be accuracy, the user might not get the expected result every time. User acceptance is another challenge if the user is willing to use chatbot or not. Chatbot cannot make sure if the user have user have understood the use of chatbot or if the user is satisfied. Thus a chatbot must understand user needs to provide appropriate feedback than just giving list of the recommendation that are available (Følstad & Brandtzæg, 2017). Chatbot has to handle errors and try to elimination them in future. In some chatbot, it processes a input in a same way even if when we use the same input multiple times, thus it should not repeat same mistake time and again.

2.3.2   Guidelines G1 and G2 in Amershi et al. (2019) to overcome challenges

G1- Make clear what the system can do. G1 explains that design has to help the user understand what the AI system is capable of doing. This guideline is helpful for user to the make them accept the system. If a user is unknown of what a system can do then, user might neglect the system, thus Chatbot must make their user understand about their capabilities and encourage them to communicate via text or verbally.

G2-Make clear how well the system can do what it can do. G2 describes that a system must be capable of explaining how well a system can perform its intendent tasks. In G1 makes sure that user are made familiar with its functionalities but argues on make users also understand how a system can give a response of the query with full accuracy. Thus the combination of G1 andG2 can in a great extent helps to overcome the challenges in the design of chatbot.

3. Module Three - Human AI collaboration

3.1   Robots , its Level of Autonomy and Explainability

There are still a misconception about Robots, people prejudices Robots based on some movies, books where robot are often used to kill or destroy. Through the advancement and development in robotic field, these misconceptions are gradually being eliminated. Robot are used in several sectors as an agent or part of team. In Philips at al. (2016) papers he has given a taxonomy and examples of human-robots collaboration. He has represented the traditional way of using animals to perform several task as an analog to show the possibility of robots in the future where robots are not just a machine rather team mates which he calls human.-robot teams. In his paper he has given examples of robots that were built to replicate tasks of animals that results physical benefits, emotional benefits and cognitive benefits. He has also discussed that trust is an essential dimension for building effective interactions between future human-robot teams, as trust ensures that functional relationships between humans and robotic teammates can be formed and sustained (Philips at al. 2016).

Similarly there is huge confusion about what level of autonomy that must be given to a robot, since it is just man made machine , there is question whether it should control human or be controlled by human. In Shneiderman (2020) paper , he has presented the two dimension two-dimensional framework of Human- Centered Artificial Intelligence that focuses on high levels of human control and high levels of automation that will result trust, reliable and safe computer application. There has also been the argument on the situations where there is excessive human control and excessive automation. Figure 1 shows the two dimensional framework very clearly, where the box at lower left sides represents less human control and less autonomy the examples of these are clock, mouse trap. The box at lower right denotes high automation but low level human control, defensive

weapons systems, pacemakers are some examples. Similarly the box on higher left side refers to higher level of human control but lower level of automation for example bicycle and piano The higher right side indicates high level of human control with high level of automation such as elevator , camera. The yellow triangle represent the idea goal where the designs are trusted, reliable and safe.
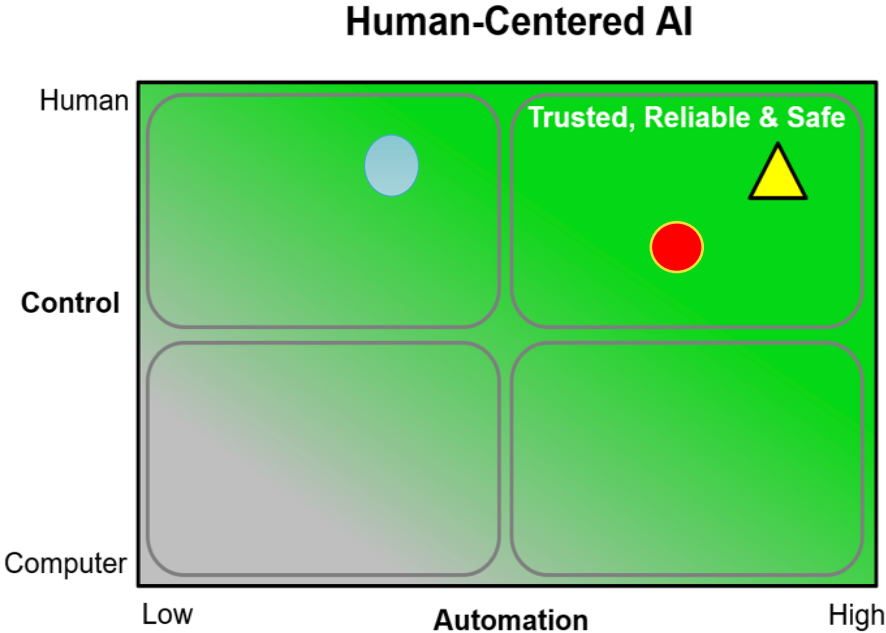

Figure1:Two-dimensional framework

Though there has been several technological development in the field of robot and AI and robotics is considered as one the branch of AI, there is a the question that how these robots or any AI infused system are acquiring the data. There is higher demand of the explainability of such system. In the paper Hagras at al. (2018) refers explainable AI as action AI which is easily understood and analyzed by humans and to be explainable the AI needs to be transparent, have a cause, not be biased, be fair, and safe.

3.1.1 T-Hawk robot

Among all the examples given by Philips at al. (2016) in his paper, I like the concept of T-Hawk robot that is used by military for search and rescue operations which replicate sheepdogs, sled dogs and specialist animals. The purpose of this robot is to search extremely risky environments that poses threat to humans such as detecting bombs, calculating destruction done by natural disasters.

This robot is has high level of human control with autonomous flight and navigation functionality (Wikipedia I, n,d). Among the Sheridan & Verplank's ten levels of automation , I think T-Hawk falls *under "Level 5: Computer executes alternative if human approves"*. And according to the Shneiderman (2020) two-dimensional framework it might fall somewhere around top left box. As shown in Figure 1 , the blue circle point out it's places in the framework. As operators controls the robot throughout the operation, there is excessive human control.

It has been once recoded that T-Hawk crash landed due to loss of control, thus I think it will be advantageous to increase the level of autonomy of this robot to eliminate such damages. T-Hawk has a control system that is operated by the operators inside a van or safe areas the whole flight is monitored, but if it is provided with higher autotomy, there will not be the need for such control. It could be equipped with more sensors so that auto piloting might be possible without human control.

The robot lacks expainability and need to develop methods of explainable AI in future. By the appearance of this robot is difficult understand what is its purpose. Only few trained operators or developer can operate this devices. On the other hand the information provided by T-Hawk in the form of picture or videos are free from bias as it captures what it see via its camera. This also support another focus of explainable AI that is safety since it performs dangerous task that might a be threat to human life. As they can reduce cost, perform actions that have a higher risk, increase in productivity will help to create more efficient services (Hagras, 2018, p.29).

3.1.2 NAO robot

Another example I find interesting is NAO robot which was built to gives emotional support and to take care of children's behaviors. NAO is a humanoid robot that resembles a human more like a child with the height of 58cm which was evolved in 2006 and gradually upgrading, NAO[6] is a latest version launched in 2018 (SoftBank Robotics, 2020). NAO robot was conducted for the purpose of using it to induce care-taking behaviors in children (Phillips et al 2016. p.107). Throughout the development phase it has been used for education and research. It is a programable robot that has be used in several scenarios thus has be commercialized and used as an assistant by companies and healthcare centers to welcome, inform and entertain visitors (SoftBank Robotics, 2020).

NAO performs several task without human intervention. It has 25 degrees of freedom which enables to move and adapt to the environment (SoftBank Robotics, 2020). It is programmable and can be used for different scenarios, it is capable of communicating with humans and can navigate itself from one place to another. Since it required human to be programmed and it can perform certain task autonomously, in my opinion if falls under *Level 7: The computer executes automatically, then necessarily informs the human"* from Sheridan & Verplank's ten levels of automation. Whereas I think, NAO falls on upper right quadrant in the two-dimensional framework (Shneiderman, 2020). In figure 1, the red circle points the place of NAO, where there is high level of autonomy with average level of human control.

In my point of view if NAO is given more higher level of automation than it has at the present time, it might be problematic. As it has been used in health sectors and other sectors to provide information, there might be chances of robot deciding itself and not providing valid information. It has also been recruited to take care of children, high level of autonomy without human supervisor may lead to unpredicted situations like children end up learning bad habits.

The structure of the NAO somehow resemble a child , thus in general everyone can anticipate it can move , communicate and sense. Indeed NAO can move freely, can interact with human, it can

recognize and communicate in 20 different languages, has seven touch sensor that helps to perceive environment and locate himself in space and can recognize shapes, objects and even people (SoftBank Robotics,2020). These function shows that it is transparent, and it is programmed for certain tasks so cannot be biased. It has been commercialize, so anyone can buy them and use it this proofs that it is safe to use. Thus, I believe NAO can be categorized as a robot having higher explainable AI features.

References:

Moor, J. (2006). The Dartmouth College artificial intelligence conference: The next fifty years. *Ai Magazine*, *27*(4), 87-87.

Marvin Minsky, by Michael Aron Dennis in Encyclopædia Britannica, Aug 5, 2020

The amazing ways Tesla is using Artificial Intelligence and Big Data, Bernard Marr (https://www.bernardmarr.com/default.asp?contentID=1251)

Andrej Karpathy - AI for Full-Self Driving at Tesla(https://techau.com.au/andrej-karpathy-head-of-ai-at-tesla-talks-full-self-driving-challenges/)

Who coined the term 'Robotics'?, Sandford Report (https://cs.stanford.edu/people/eroberts/courses/soco/projects/1998-99/robotics/history.html)

Definition of Robot , Merriam Webster (https://www.merriam-webster.com/dictionary/robot)

 Robot waiters, euronews (https://www.euronews.com/2020/06/10/restaurant-in-the-netherlands-to-use-robots-instead-of-waiters)

Universal design and its principles (http://universaldesign.ie/What-is-Universal-Design/The-7-Principles/)

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)

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)

Liao, Q. V., Gruen, D., & Miller, S. (2020). 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. (https://dl.acm.org/doi/abs/10.1145/3313831.3376590)

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)

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)

Luger, E., & Sellen, A. (2016). 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. (https://www.microsoft.com/en-us/research/wp-content/uploads/2016/08/p5286-luger.pdf)

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

Phillips, E., Ososky, S., Swigert, B. and Jentsch, F. 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, (2016) pp. 1553 – 1557
DOI: https://www.researchgate.net/publication/306003139_Human-Animal_Teams_as_an_Analog_for_Future_Human-Robot_Teams_Influencing_Design_and_Fostering_Trust

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

Wikipedia I. (n.d). T- Hwak (robot). Gathered on the 9th of November 2020 from https://en.wikipedia.org/wiki/Honeywell_RQ-16_T-Hawk

SoftBank Robotics -NAO Gathered on the 9th of November 2020 from https://www.softbankrobotics.com/emea/en/nao