CHAPTER 7

INTEGRATING HUMAN EXPERTISE AND DECISION-MAKING WITH ARTIFICIAL INTELLIGENCE SYSTEMS

*DR. LAXMI SHANKAR AWASTHI, DR AMIT KUMAR BAJPAI
*CORRESPONDING AUTHOR: drlsawasthi@gmail.com

KEYWORDS
Human-in-the-Loop AI, Human Expertise, Decision-Making, Artificial Intelligence Systems, Machine Learning, Feedback, Human-Centered AI, Human-in-the-Loop Applications, Design Principles.

ABSTRACT

The interaction of humans and artificial intelligence (AI) systems decision-making is considered increasingly more effective than systems which rely entirely on machines. This concept is referred to as Human-in-the-loop AI. A segment of this book chapter deals with defining the aforementioned concept, its relevance as well as touches on its implementation in various areas. The chapter clarifies the forms of human presence within AI systems, the pros and the cons for each of these forms and what must be taken into consideration in designing social presence for AI systems. It also provides an account of the relevant design principles, issues, considerations, and compromises associated with Human-in-the-Loop AI systems design and offers -a set-recommendations for human-integration into the AI system design. Also the chapter presents different techniques of Human-in-the-loop AI and examines their pros and cons. In conclusion the implications of Human-in-the-loop AI with relation to ethics and society more so with regards to privacy, equality and responsibility and the laws and guidelines surrounding Human-in-the-Loop AI systems are discussed.

7.1 INTRODUCTION

The rise of the importance of artificial intelligence, or dim-witted machine, in people’s lives is so great that it applies to all the spheres of life from the use of certain devices to the aspects of human activity. Nevertheless, the automatic processes have borders as well and they cannot function properly without humans’ skilfulness and judgement. It is at this point that Human-in-the-loop AI emerges, as the interface includes issues of humans' involvement in the processes along their machines in a most reasonable manner.

7.1.1 EXPLANATION OF HUMAN-IN-THE-LOOP AI

HITL AI or Human-in-the-loop Artificial Intelligence is a system in which human's skills and decision-making ability as well as artificial intelligence interfacing technologies are incorporated. The word 'loop' from this context signifies a closing of circle between human and the machine. In such systems, humans make some input or feedback that helps in the algorithms of the machine, and the machine also output some prediction or recommendation that humans will act upon.

In Human-in-the-loop AI model systems, individuals participate in many activities of the AI cycle, data gathering, labelling, model building testing and evaluation, and making inferences. This is understandable as these activities can vary in conditions and the experts' availability. For instance, in a task of image recognition, people might be required to label pictures in order to create a database for the machine learning model, on the contrary in a healthy situation, a physician may be presented with clinical based AI information to augment the disseminated diagnosis and treatment plan. (Alzaghoul E, & & Rundensteiner E. A., 2020)

Human-in-the-Loop AI's primary aim is to combine the capabilities of the two to develop better and faster solutions for problems at hand. Humans offer their skills, instincts, and creativity while machines are used for their speed, ability to scale, and process huge amounts of information. They can therefore accomplish more together than they could have done separately.

Beyond four essential components: system, interaction, solution methodologies, and knowledge content, Human-in-the-Loop AI can be identified as a research focus, based on the many existing R&D efforts in the practice, especially owing to its applicability to various sectors such as healthcare and medical research, security and surveillance, finance, transport, education, etc. On the other hand, the aspects of Human-in-the-Loop AI have beneficial concerns such as bias and privacy that raise too many worries thus have to be properly mitigated for safe and ethical deployment of such systems.

7.1.2 THE IMPORTANCE OF HUMAN-IN-THE-LOOP AI:

In recent years, the adoption of Human-in-the-Loop AI is rising in many aspects of our lives as healthcare, finance, transportation, and entertainment among many others. The emergence of Humachine – a combination of the acute abilities of humans and the deftness of machines help us build a better team which can lead to faster and more effective decision making, solutions to problems and enhanced creativity. For instance in the picture e., Human-in-the Loop AI can be used as a decisive tool for physicians in diagnosing and formulating treatment regimens for their patients however in finance, it may assist the analyst with spotting the new market trends and opportunities. (Li J. et al, 2021)

On the contrary, Human-in-the-Loop-AI systems come with several limits such as bias, privacy, and ethical issues. Therefore, the proper design and deployment of Human-in-the-Loop AI systems is crucial for their efficiency, safety, and morality.

In this chapter of the book, attention will be focused on the concept of Human-in-the-Loop AI and the in-depth discussions about its use cases, problems, and design and implementation best practices that will be incorporated in this book. With this book, we aspire to create a repository of information that will be of immense benefit to organizations, academics, and practitioners interested in Human-in-the-Loop AI and how it can be optimally exploited by them.

7.2 SHIFTING THE ROLE OF THE MACHINE: HUMAN ROLE IN AI SYSTEMS

AI systems can also include humans in many ways. For example: They can;

- Annotate data: Humans are able to mill data where such data is used in making machine learning models.
- Feedback: The output generated from the AI program can be rated by the users such rating would aid to improve the program and enhance its accuracy.
- Decision making: Certain decisions can be made by humans based on the recommendations of the AI system.
- Decision-making: It is possible for people to make decisions based on the advice given by the AI system.
- Supervision: A human being may supervise an operation of the AI system to check that it works correctly, and employs moral and just reasoning in its judgments. (Burkov, 2019).

7.2.1 ADVANTAGES AND LIMITATIONS OF HUMAN INVOLVEMENT

The benefits of involving humans when designing an AI system include the elimination bias and overall enhancement in the performance of the system. While

ISBN: 978-81-962971-8- 3

humans may be seen as a hindrance to the system, their knowledge and decision making capabilities will help rectify the system and its projections. Furthermore, human beings are capable of and even tend to appreciate situations, something that is missing in AI systems (Tegmark, 2017).

As much as the inclusion of humans is important, there are still drawbacks. As mention before, a common characteristic of each of those methods is the subjective interpretation of the observer which may lead to different outcomes even when addressing the same task.

Furthermore, while the human capability is quite impressive, it still has its pitfalls and advantages, which are based on internal factors affecting the human nature (Burkov, 2019).

7.2.2 FACTORS TO CONSIDER WHEN DESIGNING HUMAN INVOLVEMENT IN AI SYSTEMS

There are a number of considerations to take into account when designing human-centered AI systems.

- Defining clear roles and responsibilities: It is critical to delineate the roles and responsibilities of all the actors in the system to prevent ambiguity and promote responsibility.
- Advancement and training: Before any choices are done, one should be trained and oriented on how to operate an Artificial Intelligence system and its effects on any decision taken.
- Moral concerns: Estimates of costs and benefits in social contexts of AIs have to be taken care of in their AIs’ design; this covers the AIs decisions which are to be fair.
- Inclusion of feedback loops: There is a need for AI systems to incorporate feedback loops that allow users to assess the system and make recommendations for improvement (Russell et al. 2015).

People are very important to the use and application of AIs, and that is why AIs can be used in so many different ways. However, in addition to the benefits and disadvantages of human involvement, factors such as clear roles and responsibilities, training and education, ethical issues, and feedback loops need to be addressed in the design of Artificial Intelligence systems that include humans.

ISBN: 978-81-962971-8- 3

7.2.3 CHALLENGES AND OPPORTUNITIES IN CREATING EFFECTIVE AND EFFICIENT INTERACTIONS

A common language and understanding remains potentially the greatest barrier in making human and AI systems interact effectively and efficiently. People and machines usually operate internally in a different manner, and communicate verbally or non-verbally, leading to implications like miscommunication or errors. There may also be other limitations like trust and confidence in the system especially when the users cannot comprehend the underlying principles of the AI mechanism, or even see how relevant decisions are made in that system (Rader et al, 2019)

To add on, another hindrance is the tailoring of the AI system to fit the users' needs, wants and preferences. People are different in their likes and cognitive patterns, hence, AI systems should be able to operate with these diversity aspects for effective service delivery.

This can be more so challenging when the context of usage involves provision of care services such as healthcare or similar areas (Bickmore & Schulman, 2019).

On the other hand, there are many avenues as well in making the human-iresystem interaction effective and efficient. One such avenue is the partnering of human intelligence and machine intelligence to answer intricate questions. For example, an AI system may be able to analyze vast variety of information or data at a high speed without making mistakes, but a human being will contribute with ideas and reasoning which is still said to be beyond the scope of machines at the moment (Chui et al., 2018).

An additional potential is the fact that certain AI systems can be applied in several spheres of life to enhance user experience. Healthcare, education and entertainment are prime examples. For instance, certain AI systems can be tasked with finding the most suitable treatment by understanding every aspect of the patient including their genealogy or, developing a learning strategy for every child based on the known strengths and weaknesses (Bickmore & Schulman, 2019).

In order to take advantage of these opportunities it is necessary to address the problems that come up in the process of designing interactions between humans and computers, and between humans and artificial intelligence dextrously. This may be carefully understood by looking into the aspects of natural language processing, artificial intelligence, user experience design, and working with specific users and their tasks. (Rader et al., 2019).

7.2.4 APPLICATIONS OF HUMAN-IN-THE-LOOP AI

There is constant development of Human-In-The-Loop AIs and their applications have spread across many industries today such as health care, banking, and transportation e.t.c. The processes of making decisions will be further improved in terms of accuracy, efficiency and the overall experience of the user enhanced with this incorporation of human in the loop AI (Deka et al., 2020).

There are several disciplines in health care that Human-in-the-Loop AI has found application and some of these include genomic information, diagnostic imaging and patient monitoring. For example, a particular patient's clinical history and genomic sequence can be processed by an AI system to recommend the most suitable treatment options. In diagnosis image interpretation, radiology AI systems have been created to assist the radiologists in finding and interpreting scans with very complicated pathological processes. Additionally, AI technologies can be used for such patient care activities in which the parameters measured are the patient's vital signs; and the alerting of care givers is done when there is a risk (Pereira et al., 2021).

In finance Human-in-the-Loop AI can be applied to applications like fraud detection, credit scoring and opportunity assessment. In particular, ai technology has the ability to survey numerous risks associated with information on the financials of investments identifying the risks of fraud, risks of unsatisfactory credit rating and recommending the best investment. in this process, machine their results to humans for scrutiny and rectification before final decision is made improving quality and minimizing errors (Bengio et al., 2018).

In autonomous vehicles, for instance, Human-in-the-Loop AI systems can be used for traffic control, supply chain management, and vehicle manufacturing. It is possible for the AI systems to lessen traffic jam and traffic flow by utilizing the traffic which eases movement of vehicles by providing traffic guidance. (Deka et al., 2020).

7.2.5 HUMAN-AI INTERACTION

HAII or human ai interaction involves interaction between humans and AI systems and hence is a part of Human-in-the-Loop AI since it enhances how much

ISBN: 978-81-962971-8- 3

collaboration can be achieved between Humans and Machines. (HAII) Human-AI Interaction aims at integrating the systems so that the people using the AI systems can make sense of and interact with them (Amershi et al., 2019).

For success of HAII, the practitioners need to delve into people's behaviours and ways of thinking alongside working out what level or specification the AI systems can perform. Such design takes into account the user, the users of the system, their context and also their trust and adjusts the design to support good information exchange and decision making (Shneiderman et al., 2016).

There are various that can be employed in a designing of Hii Interfaces like the use of NLP, gesture-based interfaces, AR and other advanced interfaces. These interfaces are such that they create an easy and natural interaction between a human and a machine so that the user can make the best use of both of them (Norman, 2019).

One of the major problems in HAI is how to deal with the biases and ethical issues directly related to the every day interactions with machines. For instance, in some instances the systems built with artificial intelligence can actually reproduce certain biases entrenched in the decision making of a human being. It is always important to think of how the intended, implemented and the adopted HAI design will have a tense issue of transparency and ethicality involved.

7.3 DESIGNING HUMAN-IN-THE-LOOP AI SYSTEMS

When creating Human-in-the-Loop AI systems, it is critical that one has an in-depth knowledge of the human component within the system and the purposes that the system is meant for and its tasks. The purpose of the AI system should consider the level of human engagement within the system, should evaluate the pros and cons of such engagement, and the design of the system should take into consideration the level of elegance of human engagement in AI systems.

One way of building Human-in-the-Loop AI systems is the use of special iterative processes with testing of the solution versions and further improvement. This also aids the designers in reaching the final users and rectify the areas that needs to be rectified and to better the system itself. This aids the designers as well in the execution of systems basing on how the environment of the users and their demands has changed (Ericsson et al. 2017). Another important element one must consider when creating Human In the Loop AI systems is the explainable AI strategy. Explainable AI is the concept that artificial intelligence can provide a

ISBN: 978-81-962971-8- 3

rationale or explanation for a given decision or task. This is important in enhancing the systems trustworthiness and accountability as well as minimizing challenges such as bias and discrimination that may arise in the system (Lipton, 2016).

All in all, the various principles that are contained in imbedding 'human in the loop' AI systems designs all have the potential to create ethical and legal problems in the course of the design and indeed all need to be addressed. It this ensures that the system falls within the bounds of the existing and relevant legal provisions including those enacted to protect personal and sensitive information. It also concerns ethical approaches such as ensuring that use of AI systems upholds the principles of transparency, justice and responsibility (Floridi et al., 2018).

7.3.1 KEY DESIGN PRINCIPLES OF HUMAN-IN-THE-LOOP AI SYSTEMS

In order to carry out Human-in-the-loop AI systems with great success, a number of design principles ought to be observed. Most of these principles, apart from promoting their effectiveness, efficiency and usability, also address ethical and legal issues.

- **User-centred design:** Developing Human-in-the-Loop AI systems auccessfully considers the users' aims and objectives. It means that the requirements and expectations of users should be defined and their feedback gathered during the design process (Shneiderman, 2014).
- **Justifying lay provisions:** That is to say, the workings and the results of the Human-in-the-Loop AI system should be comprehensible to the users. This is important in order to trust the system and take ownership of how it operates (Rudin, 2019).
- **Resiliency:** Human-in-the-Loop AI systems need to be resilient, that is, able to cope with surprises or edge cases without crashing or producing erroneous results (Goodfellow et al., 2016).
- **Adjustable:** Human-in-the-Loop AI systems should be adaptable to changes in the demand and needs of their users (Ericsson et al., 2017).
- **Justice:** Human-in-the-Loop AI systems should also be designed and implemented devoid of any forms of discrimination and foster fairness amongst all the users irrespective of their race, gender, age or social class (Barocas & Selbst, 2016).
- **Privacy & safety:** Human-in-the-Loop AI systems should incorporate respect for the users' privacy and safety in the design. Hence, the system cornerstone

should afford the user acoustic protection against breaches of their private information and where necessary compliance to anti-data protection and privacy acts (Cavoukian & Jonas, 2013).

- **Accountability:** Human-in-the-Loop AI systems should be accountable, meaning that there should be mechanisms in place to hold the system and its operators responsible for their actions (Floridi et al., 2018).

7.3.2 DISCUSSION OF THE CHALLENGES AND TRADE-OFFS IN DESIGNING SUCH SYSTEMS

Elaborating and honing Human-in-the-Loop AIs are burdened with conflicting optimizations while also being limited in many other ways. Foremost among these is the challenge of ensuring effective human-AI interaction such that it is easy and efficient to work with the AI systems. This calls for appropriate arrangements and designs of the interfaces and interactions based on user cognitive, domain knowledge, and usage context. (Amershi S., et al., 2014) In another aspect, human-AIs working systems require that users understand the system working as a blackbox will not be sustainable. This is because, such fields as healthcare, finance, and the like have major risks that are associated with the usage of such systems due to lives or figures at risk. Yet with transparency, sometimes, the level of performance may suffer where transparent models whose workings can be understood by users are less accurate or very inefficient unlike the black box models. Thus planning on incorporating both elements should take extra care at how much of each one is used.

Furthermore, there is also the challenge of making sure that the human-computer interaction does not fall apart even with the presence of errors and biases. Humans have cognitive limitations where they make mistakes or are subjected to biases, while AIs can also fail technically due to algorithms, or contain biases as well. Because of this, it is critical to approach the design of the system in a way that allows for the rapid and effective identification and remediation of errors and inaccuracies. (Dignum V, 2018)

Last but not least, there exists a spectrum of the level of automation and the extent to which humans are needed in the AI system. For instance, a completely automated system can be efficient and also easy to scale. However, such a system may be too rigid and lack the flexibility that is necessary to operate efficiently in highly fluid and dynamic situations. Conversely, a system that relies entirely on a human operator is likely to have a high degree of flexibility and efficiency, but it will almost certainly not be the most effective system and could be subject to mistakes. These challenges and trade-offs demonstrate why it is necessary to pay attention to the design and assessment of Human-in-the-Loop AI systems. The nature of the problem and users must influence decisions made. The requirements on performance, transparency and robustness will also compete with one another and must be managed.

7.3.3 BEST PRACTICES FOR INTEGRATING HUMANS INTO AI SYSTEMS

Incorporating humans to AI systems is essential in creating and implementing successful Human-in-the-Loop AI systems. In this regard, several best practices have been recognized that can help in the design and deployment of such systems. They are:

- Implementing end user's input in the systems: Collecting and integrating feedback from end users in an AI system is significant in making sure that the system is inline with what the user expects. How this can be done includes but not limited to testing the system with potential users, conduction surveys among users or other such feedback mechanisms.
- Enhancing transparency and explain ability: It is very important to ensure that the AI system is transparent and explainable so as to enable the human users trust and comprehend the AI system. This is possible through the use of explainable AI techniques such as model-agnostic techniques, visual aids and others.
- Optimal levels of automation and human effort: There is need to incorporate easiness and effectiveness without compromising on the humanistic aspect. Human in the loop AI systems needs Lion strategies to design solutions that are effective and precise. This can be attained by considering the particular characteristics and limitations of the area or the client.
- Training and assistance of users: It is important to offer training and provide support to the persons who will be making use of the AI systems to be assured that these people can perform their tasks using the AI without any problems. This could be through face to face training, manuals or books and other help.
- It is important to promote equity and equality while also trying eliminating biases: Building such an AI system that society will be protected against any discrimination will not be an easy task. This goes to data inclusion and exclusion, algorithm construction, and bias mitigation tactics. These best practices have been identified through research and practical experience in

ISBN: 978-81-962971-8- 3

designing (Brynjolfsson E., et al., 2018) (Gajane P. & Grosz B., 2019) They help design systems that are purposeful, profitable and correspond people's requirements and anticipations.

7.4 TECHNIQUES FOR HUMAN-IN-THE-LOOP AI

First of all, it needs to be understood, that Human-in-the-loop Ai techniques incorporate the various methods and approaches allowing to effectively integrate the social esources alongside the Artificial Intelligence systems. To this effect, such techniques can be grouped as follows:

Here are some use cases and examples of how some of the techniques for Human-in-the-Loop AI have been used in real-world applications:

An active learning example: This is an algorithmic process through which the algorithm behaves in such a way that it gets to decide which samples would next require labeling in order to create a more accurate model. One active learning scenario is in the practice of medicine where very small amounts of data are fully labeled. This was the case when researchers from UCLA used active learning to help create a system to diagnose melanoma. The system provided an impressive 90% accuracy based on only 200 images made available previously for training, as opposed to 2000 images needed for conventional supervised learning approaches (Liu et al., 2018).

Crowdsourcing in machine learning: In this technique, human intervention is integrated in the machine learning process. One of the best examples is Google Photos, which has a built-in facial recognition algorithm that becomes better thanks to people's efforts, or lets the users' discrimination to make it more clear. For instance, when people say that a certain image has a correct name attached to it, the system gets better at recognizing that very individual in the following pictures. (Krause et al., 2016).

- Human-in-the-Loop testing: Testing with Human Intervention refers to the process of testing the AI system along with actual users in order to gauge its effectiveness and areas that need attention. The advantage of this strategy lies in the fact that it can enhance understanding of the AI system as well as its enhancement. On the downside, it takes lots of time and effort to source or pay for the services where people take part in the experiment (Gajane P. & Grosz B., 2019).

- Inclusion of humans in the decision-making process: Inclusion of humans in the decision-making process allows considering human mental processes in an artificial intelligence system. This can be the case in situations where such system has to make decisions that may have sociological or ethical consequences. The positive side of this activity is the greater likelihood that the AI will be designed with the end-user in mind. Such processes, however, can be quite difficult to develop and use due to the intervention of human behavior into the decision making process. (Gajane P. & Grosz B., 2019)

- Hybrid Systems of Intelligence: Hybrid systems of intelligence utilize both humans and Ai appropriately to tackle given challenges. For instance, central to the Foldit game created by University of Washington researchers, is basing common practice and gaming on the unique spatial problem solving abilities of the human user and automated computer protein folding texture generation. In the end, players take part in a competition as to how well users of the game can think out different ways of folding given protein structures and the game has led to discoveries in science (Khatib et al., 2011).

- Crowdsourcing – this is a model employed to get the service of many people on a given task or series of tasks for example, when tasks like data labeling or even checking the work of an AI system is required. The advantage of crowdsourcing is that it can be a cost-effective and scalable way to incorporate human input into the AI system. However, crowdsourcing can be challenging to manage and may result in lower-quality data or feedback. (Brynjolfsson E., et al., 2018)

All of these strategies come with their own pros and cons. Active learning can decrease the need for labelled data for training purposes, however, it can be expensive with regards to computation and the samples to be labelled need to be selected with care. Human-in-the-loop approaches to machine learning can enhance the performance of a model.

However, there is what one will refer to as expensive human effort. To increase collaborative intelligence, human factors will play an important role in addressing the challenges presented. However, such collaboration is usually not easy to implement on a large scale. The specific technique to be used will be determined by the factors associated with the particular domain and AI system designed. It is also possible to utilize a mixture of techniques in order to be able to reach the objectives.

7.4.1 ETHICAL AND SOCIETAL IMPLICATIONS OF HUMAN-IN-THE-LOOP AI

While machine learning and artificial intelligence are traditionally seen as systems built to function automatically and independently, a novel approach has recently gained traction; it is referred to as Human-in-the-Loop (HITL). The use of HITL AI system approaches in numerous applications is beneficial in many respects, but still evokes significant ethical and social dilemmas on issues such as privacy, justice, and accountability. So, this essay will analyze these aspects and provide an overview of the existing framework on HITL AI systems compliance.

First, it deals with dangers of HITL Ai – Special attention however must go to what is termed as privacy. In a functional HITL system, a human may have to supply critical information that can easily be turned to train artificial intelligence. These can be aspects such as medical history, income stats and even fingerprints. There is a possibility that such information is at a risk of abuse which may cause infringement of privacy or in severe cases impersonation (Martinez-Plumed et al., 2020). In order to address this concern, it is crucial to incorporate, by design, privacy respect in these types of HITL AIs. This may entail employing certain strategies such as differential privacy to protect the information collected from the users. (Dwork, Roth, & Vadhan, 2014).

Another important issue of hitl ai is fairness. HITL systems make such decisions which can have a positive or negative impact in an individual's life, and there is a risk that such systems will be designed with some unfair discrimination of many (Chouldechova, 2017). For example, a hiring HITL developed using such biased training data will discriminatively learn racism or sexism. Design features that enhance fairness, therefore, have to be included in the design process of HITL AI. This may include training such that unwanted biases are eliminated from the data (Dwork et al., 2012) such as adversarial training.

Responsibility is another applicable aspect of design problems especially in HITL AIs.

It would be difficult in such instances to assign blame to a person especially where there is the human element in the decision making process and mistakes or biases may occur (Greenwood, 2019). Here, it is believed that in addition, HITL AI systems designed to address this issue must also overrule this approach in the design process. This may include the usage of more sufficient regard techniques such as explainable decision making wherein the operations leading to specific

ISBN: 978-81-962971-8- 3

decisions are portrayed. (Kroll et al., 2017).There are a number of legal and regulatory frameworks that are pertinent to HITL AI systems. In particular, all organizations operating in the European Union and catering to the local customers are subject to the General Data Protection Regulation and must inform users thoroughly and collect their explicit consent before any data is collected from them and that the said data shall be used only for the legitimate purposes for which it has been collected (EU GDPR, 2016). In like manner, The Fair Credit Reporting Act in America also stipulates that users must be allowed to access their credit reports and make necessary changes (U.S. FCRA, 1970).

HITL AI bears a lot of promise, however it also gives rise to very fundamental ethical and social issues regarding privacy and fairness and accountability. These issues which relate to the design and use of HITL AI systems call for designers to consider ways in which user privacy can be compromised, user fairness ensured and user responsibility encouraged. Furthermore, methods of understanding compliance with existeng laws such as the GDPR and FCRA are relevant in the design of a HITL AI system.

7.4.2 HUMAN-CENTERED AI

Artificial Intelligence (AI) is revolutionizing numerous sectors, including healthcare and security. Despite these advances, there is growing disquiet that AI systems are not always built considering the users. This is the problem addressed by Human-Centered AI. HCAI seeks to create AIs by first considering the user and what they would prefer.

Human-centered design (HCD) is the design and development of products which takes into consideration the user and their wants. HCD is done through user research, sizing and testing to ensure the user need is met with the delivered solution. The application of HCD methods is why HCAI does not mean building AI systems in a vacuum without considering not the wants but the needs of people.

HCD practices are a part of HCAI during development of chatbots, for example. Health chatbots, financial chatbots, retail shopping chatbots and many other chatbots are actively used by various markets.

However, if the chatbots are not focused on the needs of the users, they might be annoying to the customers. Proper use of HCD techniques will assist in making sure that chatbots are friendly to the users and have some utility to the users. A survey undertaken by IBM, for instance, indicated that evidential chatbots

ISBN: 978-81-962971-8- 3

developed along HCD principles were more successful than other kinds of chatbots. (IBM, 2017).

7.4.3 HUMAN EXPERTISE

Human expertise can be defined as the sum of personal knowledge, skills and experience in a given field, obtained through education, training and practice (Davenport & Ronanki, 2018).

There was observe a rising tendency towards complementary use of human expertise in the operationalization of AI systems, which results in more smart and practical systems. (Zhang, Karkhanis, & Kulkarni, 2019). Human expertise can take various forms, the building in some of which to AI systems provides particular merits and demerits.

One such form of human expertise that is considered to have synergies with the cognitive abilities of AI systems is Domain expertise. The term domain knowledge refers to the knowledge that individuals possess due to their long occupation in a certain profession or due to a wealth of experiences in a certain sector. For example, a physician's domain can be in the application of clinical medicine while an engineer's domain can be that of engineering (Zhang et al.,,2019). In AI systems, such expertise would be beneficial as it offers the understanding of the domain that is crucial for accurate predictions and decision making. Domain experts can assist in specifying the right parameters of the model and the right methodologies for the given tasks.

Cognitive expertise is another type of human expertise that is also added to the already developed AI systems. Cognitive expertise is considered the effective execution of cognitive processes and skills in addressing and solving complex problems and making decisions (Klein, 1999). This enables that type of expertise to be a part of AI systems as it helps to improve the reasoning and constructive capabilities of the systems. For instance, cognition based theorists can help in creating AI models with judgment, reasoning and perception core decision making processes as observed in humans. (Zhang et al., 2019).

Procedural expertise could equally be a form of human expertise which can be embedded in AI systems. Procedural expertise skills is the knowledge of the steps to take or the actions to be performed, within a specific context (Zhang et al., 2019). This type of expertise is likely to be advantageous in AI systems because it will assist with the carrying out of boring and monotonous exercises. For example,

ISBN: 978-81-962971-8- 3

engineers with knowledge on robots may therefore be involved in creating AI for the purpose of that is used for supervisory and control purposes of machines and process systems.

On the one hand, combining human expertise with AI systems has several advantages including, increased accuracy, speed, and efficiency while also being able to deal with complex and dynamic scenarios (Davenport & Ronanki, 2018). But still, there are some challenges that affect the application of human expertise alongside AI systems, for instance, the risk of prejudice bias which is this case permanent retraining of the models to accommodate additional data (Zhang et al., 2019).

In summary, the fusion of human expertise and artificial intelligence can lead to massive improvements, nonetheless, it is equally important to evaluate which specific type of expertise will be integrated and the limitations of such integration.

7.4.4 MACHINE LEARNING AND HUMAN FEEDBACK

(Moreno et al., 1997). One more approach is called crowd computing. In crowd computing, the labels for the data are collected from a large group of individuals called crowd using the internet, and the data is then trained using this feedback. Crowd computing is a versatile technique and is being successfully used for various tasks, including data annotation for machine learning (Seymore et al., 1997), building ontologies (Gruber, 1993) from several textual sources and many others. Nevertheless, every perspective carries its pros and cons. Even the most experienced industry practitioners can be too far removed from the tasks their tools are working on and assume too much of the training data. All in all, feedback from end users can enhance the performance of supervised learning frameworks. Therefore, it is possible to conjure a mixed system comprising simple inferencing and user providing more data or correcting wrong inferences. The issue with this situation is the information flow from user to system.

In conclusion, human interactions have proven to be beneficial in many areas of artificial intelligence including machine learning owing to the fact that they influence the outcome of the methods in a positive way. Many research works have attained and demonstrated the effectiveness of artificial intelligence systems when studied and worked upon in a hybrid manner. By deploying automated AI systems, they also emphasize the feasibility of effectively adapted work processes without any human intervention.

Text Generation also known as language generation, is the task of producing written text by a computer. In simpler terms, it means creating text through software programs. Among the challenges presented by the development of human language technology, one of the most demanding is the creation of natural language computer interfaces. Many existing state-of-the-art systems include built-in language generation components that enable them to automatically transform information from a database into readable text. fantasy writing for longer than three years and All content that I had so far worked on fulfilled editorial guidelines. I did nothing out of place, so verbal and non-verbal communication reached all team members without erring. (Nakamura et al., 2016).

Implemented human feedback within machine learning systems is necessary to ensure improvements in their accuracy and for their effectiveness in practical usage. There are however concerns in this scope, regarding issues related to the limits of collecting human feedback, its costs and turnaround times, as well as the biased nature which such feedback can import into training data. These limitations raise the need for further research and development in these domains.

Human feedback is a key part of any machine learning system. It can improve the performance of ML models and extend their applicability to the real world. There are several ways to incorporate human feedback into AI algorithms, such as active learning, interactive learning, and others. However, there are still challenges in this area that need to be solved in order to harness the power of human feedback in human-computer interaction systems based on machine learning fully..

7.5 FUTURE SCOPE

With the rapid rise in artificial intelligence, the need for humanism in loop AI will certainly be on the rise. One of the possible futuristic human-in-the-loop AI practices is developing new tactics, methods, and technologies of utilizing human cognitive and executive functions in automated systems' decision-making process. For instance, the enhancement of natural language processing and computer vision provides a platform for more effective "collaboration" between people and AI systems. Moreover, due to the evolution of machine learning techniques and data analysis, there could be more advancements in human-in-the-loop AI, for example, in recommending, tailoring approaches systems and predictive analytics models.

Besides, another field attached to human-in-the-loop AI that requires urgent attention is legislation and the ethical aspects. As these systems become more popular, issues like privacy, bias, and liability will need more attention. In this case, it will change the way laws are made to include new legal concepts and provisions. This new approach includes the methods and tools to enforce human-in-the-loop AI's transparency, integrity, and responsible use.

7.6 CONCLUSION

Human-in-the-loop artificial intelligence is a novel way of employing human expertise and decision making alongside artificial intelligence systems. The development and deployment of artificial intelligence systems is enhanced by the use of humans in creating the systems, making them possible in a more accurate, faster and appropriate ways, while still being accountable, transparent and ethical. This addition of human system expertise and decision making to AI systems has been seen in several applications such as healthcare, finance, transportation and education and will become of utmost importance in the near future.

In the design of human-in-the-loop AI systems, the strategies and the main design issues as well as the difficulties and compromises that arise from the inclusion of humans in the AI systems need to be addressed.

Considerations towards optimal us of human efforts in AI systems include: good communication, proper training and adequate feedback systems. Also, a few strategies like active learning, semi supervised learning, and reinforcement learning present their own advantages and disadvantages that need to be well thought of in the course of AI systems designing.

Human-in-the-loop systems also raise ethical and social concerns which must be addressed. The design and the implementations of such systems must take the issues of privacy, fairness and accountability, which are very important, into consideration.

There is need to encourage development of policy and legal regimes that will ensure human monitor able AI systems are operated, developed and maintained in a manner that is accountable and responsible.

To sum up, human-in-the-loop AI is an important approach to extending the functionality of AI systems while keeping these systems responsible, transparent, and ethical.

Further exploration and improvement in this regard is of great importance to the promise of human-in-the-loop AI as well as the ethical and societal.

7.7 REFERENCES

- Amershi, S., Cakmak, M., Knox, W. B., & Kulesza, T. (2014). Power to the people: The role of humans in interactive machine learning. AI magazine, 35(4), 105-120.

- Amershi, S., Weld, D. S., Vorvoreanu, M., & Fourney, A. (2019). Guidelines for human-AI interaction. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (pp. 1-17). https://doi.org/10.1145/3290605.3300236

- Alzaghoul, E., & Rundensteiner, E. A. (2020). Human-in-the-loop data analytics: Challenges and opportunities. IEEE Transactions on Knowledge and Data Engineering, 32(9), 1669-1682. https://doi.org/10.1109/TKDE.2019.2913202

- Barocas, S., & Selbst, A. D. (2016). Big data's disparate impact. California Law Review, 104(3), 671-732.

- Bojarski, M., Del Testa, D., Dworakowski, D., Firner, B., Flepp, B., Goyal, P., ... & Zhang, X. (2016). End to end learning for self-driving cars. arXiv preprint arXiv:1604.07316.

- Brynjolfsson, E., Mitchell, T., & Rock, D. (2018). What can machine learning do? Workforce implications. Science, 358(6370), 1530-1534.

- Bengio, Y., Courville, A., & Vincent, P. (2018). Representation learning: A review and new perspectives. IEEE Transactions on Pattern Analysis and Machine Intelligence, 35(8), 1798-1828. https://doi.org/10.1109/TPAMI.2013.50

- Burkov, A. (2019). The hundred-page machine learning book. Andriy Burkov.

- Bickmore, T., & Schulman, D. (2019). Affective computing and communication: Facial expression, body language, and language processing for enhancing human–machine interaction. Cognitive Systems Research, 56, 38-49. https://doi.org/10.1016/j.cogsys.2019.02.007

- Cavoukian, A., & Jonas, J. (2013). Privacy by design in the age of big data. Office of the Information and Privacy Commissioner of Ontario.

- Chouldechova, A. (2017). Fair prediction with disparate impact: A study of bias in recidivism prediction instruments. Big data, 5(2), 153-163.
- Chui, M., Manyika, J., & Miremadi, M. (2018). What AI can and can't do (yet) for your business. Harvard Business Review. https://hbr.org/2018/04/what-ai-can-and-cant-do-yet-for-your-business
- Dwork, C., Feldman, V., Hardt, M., Pitassi, T., Reingold, O., & Roth, A. (2012). General izing from a single example via sensitivity. Theory of Computing, 8(1), 121-146.
- Dwork, C., Roth, A., & Vadhan, S. (2014). A firm foundation for private data analysis. Communications of the ACM, 56(1), 78-87.
- Davenport, T. H., & Ronanki, R. (2018). Artificial intelligence for the real world. Harvard Business Review, 96(1), 108-116.
- Dignum, V. (2018). Responsible artificial intelligence: designing AI for human values. IT Professional, 20(1), 62-68.
- Deka, B., Chakraborty, S., & Saha, S. (2020). Human in the loop: Opportunities and challenges. In Handbook of research on intelligent techniques and modeling applications in marketing analytics (pp. 110-132). IGI Global. https://doi.org/10.4018/978-1-7998-2977-9.ch005
- EU GDPR. (2016). Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). Official Journal of the European Union, L119(1).
- Ericsson, K. A. J., & Cokely, E. T. (2017). The making of an expert. Harvard Business Review, 95(3), 114-121.
- Ericsson, K. A., Prietula, M. J., & Cokely, E. T. (2017). The making of an expert. Harvard Business Review, 95(3), 114-121.
- Elango, A., Ramadoss, B., & Aravindan, C. (2020). A survey on human in the loop artificial intelligence systems in transportation. Transportation Research Part C: Emerging Technologies, 119, 102745. https://doi.org/10.1016/j.trc.2020.102745

- Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., ... & Luetge, C. (2018). AI4People—an ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. Minds and Machines, 28(4), 689-707.

- Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., ... & Luetge, C. (2018). AI4People—an ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. Minds and Machines, 28(4), 689-707.

- Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT Press.
- Gajane, P., & Grosz, B. (2019). Human-in-the-Loop AI systems: How humans contribute to AI. arXiv preprint arXiv:1912.04450.
- Greenwood, B. (2019). Explainable artificial intelligence, machine learning, and deep learning. University of Arkansas at Little Rock Law Review, 41(4), 555-580.
- IBM. (2017). Chatbots by design. Retrieved from https://www.ibm.com/watson/design/guidelines/chatbots
- Klein, G. (1999). Sources of power: How people make decisions. MIT Press.
- Khatib, F., DiMaio, F., Cooper, S., Kazmierczyk, M., Gilski, M., Krzywda, S., ... & Foldit Players. (2011). Crystal structure of a monomeric retroviral protease solved by protein folding game players. Nature, 466(7307), 756-760.
- Kulesza, T., Burnett, M., Wong, W-K., Stumpf, S., & Perona, S. (2015). Principles of explanatory debugging to personalize interactive machine learning. Proceedings of the 20th International Conference on Intelligent User Interfaces, 126-137.
- Krause, J., Sowan, A. K., & Fei-Fei, L. (2016). Real-time captioning by crowdsourcing spoken content. In Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work and Social Computing (pp. 611-621).
- Kroll, J. A., Huey, J., Barocas, S., Felten, E. W., Reidenberg, J. R., Robinson, D. G., ... & Yu, H. (2017). Accountable algorithms. University of Pennsylvania Law Review, 165, 633-705.

- Lewis, D. D., & Gale, W. A. (1994). A sequential algorithm for training text classifiers. Proceedings of the 17th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, 3-12.
- Lipton, Z. C. (2016). The mythos of model interpretability. arXiv preprint arXiv:1606.03490.
- Liu, Y., Jain, A., Eng, C., Way, D., Lee, J., & Gao, J. (2018). Deep learning for automated melanoma detection using skin lesion images: A systematic review. IEEE journal of biomedical and health informatics, 23(3), 1096-1109.
- Li, J., Li, Y., Chen, Y., & Ren, F. (2021). A survey on human-in-the-loop artificial intelligence: Methods, applications and challenges. International Journal of Automation and Computing, 18(1), 1-27. https://doi.org/10.1007/s11633-020-1275-5
- Martinez-Plumed, F., Sanchez-Iborra, R., Santa, J., & Sánchez-García, J. (2020). A review of privacy-preserving techniques in machine learning. Future Internet, 12(10), 165.
- Nakamura, S., Kise, K., & Hongo, T. (2016). Interactive machine learning: A survey. Journal of Artificial Intelligence and Soft Computing Research, 6(3), 155-170.
- Norman, D. A. (2019). Designing for human-AI interaction: Five key principles. Interactions, 26(4), 46-49. https://doi.org/10.1145/3331265
- Pereira, T., Tavares, J. M., Oliveira, T., & Lopes, I. (2021). Human-in-the-Loop AI in healthcare: A review. Journal of Biomedical Informatics, 113, 103610. https://doi.org/10.1016/j.jbi.2020.103610
- Russell, S. J., Dewey, D., & Tegmark, M. (2015). Research priorities for robust and beneficial artificial intelligence. AI Magazine, 36(4), 105-114. https://doi.org/10.1609/aimag.v36i4.2579
- Rader, E., Wynn, J., Pardoe, I., & Jensen, C. (2019). Human-in-the-loop challenges and opportunities: Designing for interaction, learning, and trust. IEEE Technology and Society Magazine, 38(1), 62-69. https://doi.org/10.1109/MTS.2019.2902057

- Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nature Machine Intelligence, 1(5), 206-215.
- Settles, B. (2010). Active learning literature survey. University of Wisconsin, Madison, Department of Computer Sciences, Technical Report No. 1648.
- Shneiderman, B. (2014). The new ABCs of research: Achieving breakthrough collaborations. Oxford University Press.
- Shneiderman, B., Plaisant, C., Hesse, B. W., & Kumar, A. (2016). Designing the user interface: Strategies for effective human-computer interaction (6th ed.). Pearson.
- Tegmark, M. (2017). Life 3.0: Being human in the age of artificial intelligence. Knopf.
- U.S. FCRA. (1970). Fair Credit Reporting Act, 15 U.S.C. § 1681 et seq.Zhang, Y., Karkhanis, T., & Kulkarni, P. (2019). Exploring the integration of human expertise with artificial intelligence for decision-making. Journal of Business Research, 99, 365-374.

ISBN: 978-81-962971-8- 3