Scientific Research Publishing

Open Journal of Social Sciences, 2025, 13(3), 64-80
https://www.scirp.org/journal/jss
ISSN Online: 2327-5960
ISSN Print: 2327-5952

Trusting Artificial Intelligence: A Qualitative Exploration of Public Perception and Acceptance of the Risks and Benefits

Omonzokpia George Okoidigun[1], Eyere Emagbetere[2]

[1]School of Information technology, University of the Cumberland, Williamsburg, Kentucky, USA
[2]Department of Mechanical Engineering, Federal University of Petroleum Resources, Effurun, Nigeria
Email: emagbetere.eyere@fupre.edu.ng

How to cite this paper: Okoidigun, O. G., & Emagbetere, E. (2025). Trusting Artificial Intelligence: A Qualitative Exploration of Public Perception and Acceptance of the Risks and Benefits. *Open Journal of Social Sciences, 13,* 64-80.
https://doi.org/10.4236/jss.2025.133006

Received: October 26, 2024
Accepted: March 8, 2025
Published: March 11, 2025

Abstract

This paper aims to research the perceptions of the populace on the dangers and advantages of Artificial Intelligence (AI), and the effects of those perceptions on the trust and acceptance of AI. A population of ninety participants from San Francisco, California were interviewed and their responses were analysed qualitatively, resulting in an intriguing mix of both positive and negative perceptions. While new opportunities are seen to be helpful especially in making life more convenient and helping solve world's problems on one hand, AI is seen to have some negative consequences. About 60% are concerned about the likelihood of AI to replace people in their jobs, 47% are concerned with its potential to be biased, and 52% are concerned with its likelihood to invade people's privacy. Concerning the specific qualities, the trust was identified as an important measure. 58% of the participants questioned the openness of AI's actions and their objectives. In addition, 62% participants showed a high interest in better governance and disclosure in AI development and use. This work provides insights into the fact that there is a need to increase public trust in AI through establishing higher levels of transparency, ethical standards accompanied by regulation. All these insights necessitate extending the discussions on AI's governance and emphasizing that we must achieve innovation in this sphere, while taking into consideration certain key values.

Keywords

Artificial Intelligence (AI), Public Perception, AI Trust, AI Risks and Benefits, AI Ethics, AI Governance

1. Introduction

Artificial Intelligence (AI), which used to be a marginal branch of computer sci-

ence, has become a universally incorporated technology in today's existing society. In several essential fields of our lives, including medical services, financial and banking services, education, and entertaining services, AI systems are becoming an inseparable part of our daily lives (Haleem, Javaid, & Khan, 2019; Prianto, Viony Sumantri, & Sasmita, 2020). Such integration, therefore, leads to massive improvement in productivity, creativity, and solution finding mechanisms. For instance, modern solutions that are based on artificial intelligence algorithms allow to forecast diseases, manage the course of learning, and facilitate car driving (Amini, Bagheri, & Delen, 2022; Thadeshwar, Shah, Jain, Chaudhari, & Badgujar, 2020). Though such advances are encouraging, more worrisome steady trends of the adverse impact of AI on the society are being witnessed.

As mentioned above, the risks adversity linked with AI is huge and involves several elements. These are risk of job loss by replacement with AI, adverse prejudices/biases in the algorithms used by AI, infringement on individual's privacy, possibility of the use of AI with negative effects on the individual or the society (Acemoglu, 2021; Rakowski, Polak, & Kowalikova, 2021). In addition, the sophistication of AI systems is only set to grow and so far, regulation has failed to keep up with the rate of advancement and this is in terms of who is accountable and for how long. These matters relate to a more general social discussion that tends to be characterized by regularly shifting between concerns about the impact of Artificial Intelligence, and about its potential effects.

There are important reasons why the public's attitude toward AI, its opportunities, and threats should be understood. First, societal attitude is an essential key determinant to the uptake and embrace of AI systems. Research has indicated that when people have trust in technology, they will use such a technology while mistrust will make people resist or reject the use of a particular technology (Kahma & Matschoss, 2017; Chou & Gaysynsky, 2023). This is especially significant in sectors such as the healthcare industry where the reliability of algorithms to diagnose, recommend treatments or medication is consequential. Second, which is the perception that the public has about AI, assists policymakers and technologists, who are in a position of nurturing the ethical AI future, to do so accordingly. The public concern or negative attitude towards AI can also incite the requirement of intense regulation and exploring ways on how to get back the trust of the people.

Although there is a recent surge in the amount of work focusing on AI, one of the areas that have not been explored is the public's attitudes toward both the benefits and the threats that AI poses. A vast majority of current works concerns itself with the technological aspect of AI or the ethical concerns from the technical point of view only (Li, Ruijs, & Lu, 2023; Tippins, Oswald, & McPhail, 2021). Furthermore, there is very little scientific literature that involves a more complex analysis of people's attitudes towards AI and how those attitudes are ambivalent. This study aims to address that by offering a qualitative analysis of how the public perceives AI with the view of establishing how people manage to traverse the landscape of belief and disbelief, trust and mistrust and thus acceptance.

The purpose of this study is three-fold. First, it seeks to establish the variables that control trust in AI among the public. Second, it examines how the public and different stakeholders understand possibilities of AI risky and positive in different fields. Last, it explores the patterns of going beyond the identified perceptions to looking at AI adoption and policies at the societal level. In light of these considerations, the research questions guiding this study are as follows:

1) Which aspects play the leading roles within trust in AI by the public?

2) In what ways do individuals conceive the risks and benefits of AI, and how do these perceptions affect the level of his or her acceptance for such developments?

3) What are the wider implications of the given part of public perception for the further development of AI and regulation of this field?

If we do not understand the answers to these questions, then we cannot build a society that wants and supports AI technologies to be created and deployed to its potential. It means that, as the AI becomes more and more prominent, it is essential not to disregard people's attitude toward it as an important aspect to consider in the development of the best strategy of using the benefits of AI while eliminating its drawbacks.

2. Literature Review

2.1. Overview of AI and Public Perception

Sources have associated Artificial Intelligence with being the main key to the progression of technology and the social world. The awareness and understanding of AI by the public, nonetheless, include variety of factors involving a positive and negative view about the technology. There are a few reported works that analyse how people perceive AI; it was established that although people are optimistic about AI in general, they remain concerned about its effect on employment, privacy and ethics (Gupta & Mishra, 2022). Another factor includes the way that AI and its capabilities and uses are depicted in the media, and popular culture. Movies, television shows, books, and news have premised Artificial Intelligence in the polarized and inclusive categories of humanity's advancement and saviour or as humanity's doom as it portrays catastrophic outcomes (Cui & Wu, 2021; Dieter & Gessler, 2021; Shank & DeSanti, 2018). This duality is evidence of the state of hope that is associated with the application of artificial intelligence but dread that is associated with AI.

2.2. Risks and Benefits of AI

The matters of risk with the help of AI applications are among the key hot topics of discussion. Indeed, one of the most often mentioned threats is so called automation threat which can lead to job loss. Several research has indicated that industries that are more susceptible to automation include manufacturing, transportation, and retailing which can imply massive unemployment (Ernst, Merola, & Samaan, 2019; Chaudhary & Tyagi, 2019). However, wisdom systems them-

selves may be unwise; precisely, they can be ridden with bigotry, which implies that fairness in various sectors such as employment, policing, and credit can be elusive (Chanda & Banerjee, 2024; Cossette-Lefebvre & Maclure, 2023). These risks are worsened by the fact that most AI programs are known to have black boxes thus it is hard for the public to comprehend on how decisions are made eroding their trust even further (von Eschenbach, 2021; Bélisle-Pipon, Monteferrante, Roy, & Couture, 2023). Inadequate privacy remains the other main concern because people may feel that someone is spying on them due to the cameras mounted by the AI program. The more advanced the AI systems are integrated in people's everyday lives the amount of data collected and analysed rises dramatically. This has raised concerns about surveillance and a possibility of the abuse of person information as pointed by Almeida, Shmarko, & Lomas (2022).

On the other hand, the implication of AI is highly beneficial as has been widely stated. AI can be effective in changing industries because of its capabilities to enhance various efficiency, precision, and decision-making outcomes. In healthcare for instance, artificial intelligence is being employed to forecast patient progression, prescribe treatment depending on the individual's needs and even in invasive surgery. In environmental science, AI is making an impact in areas such as climate change tracking, resource management and the defining of sustainable approaches (Kaack et al., 2022; Tuli et al., 2022). All these benefits depict how AI can be used to solve some of the existing problems facing the world today.

2.3. Trust in Technology and AI

The adoption of AI, and its acceptance, depends on the level of trust in the technology, and its effectiveness. Various researchers have participated in a discourse on the idea of trust in technology where scholars pointed out that transparency, accountability, and users' control contribute significantly to the trust model of a given technology (Shin & Park, 2019). In the case of AI, trust issues arise as a result of opacity most of them possess, where the user cannot completely explain the rationale behind the decision making (von Eschenbach, 2021). Such lack of clarity means scepticism and resistance to use of AI technologies, especially when the result of a decision made by an AI is going to be large, for example in autonomous cars or cases involving the criminal justice system (Nakashima, Mantovani, & Machado Junior, 2024; Hobson, Yesberg, Bradford, & Jackson, 2023). Trust is also underlined as one of the AI concerns in ethical considerations for AI as seen in other studies. The ethical frameworks for AI allude to the three fundamental principles of fairness, accountability, and transparency (FAT) that should be implemented in AI for them to be considered safe in the society (Hoffmann, Roberts, Wolf JD, & Wood, 2019). These frameworks request participation for the emergence of artificial intelligence technologies to avoid presenting new types of bias or adding to existing bias.

2.4. Gaps in Existing Research

Despite a well-developed stream of research on the threats and opportunities of

AI and the role of trust in them, research on the community's nuanced conceptions of AI remains lacking. Instead, most works are devoted to certain facets of AI, which might cover its technological potential or socio-ethical consequences, all viewed in terms of an authority and provided by an authority's representatives. There is comparatively little academic discussion of how different members of the public make sense of the relative threat and opportunity of AI in their lives. Specifically, there is scant literature about the quantitative research on the specific topics and perspectives regarding AI which people may have. The result of analysing these two views is an important factor to consider when creating AI technologies that would be efficient and acceptable by the public. That is why this study aims to fill these gaps by carrying out a qualitative analysis of different aspects of AI, as well as people's trust or mistrust towards them, and the ways they go about it. In this vein, it intends to broaden the knowledge of factors that predispose the public to AI and offers recommendations that may assist in the development of AI technologies.

3. Methodology

3.1. Research Design

This paper uses a qualitative research methodology to identify key perceptions that people have about AI, and how such perceptions shape the attitudes that people have towards AI and their willingness to accept it. Qualitative method is suitable for this research since it seeks to explore the black and white of people's perception on issues such as divorce, separation and single parenthood. The primary mode of data collection used in the study is the semi-structured interviews; further analysis is done thematically as it will help to decode themes emerging in data.

3.2. Participant Selection

Using purposive sampling, participants for this study were identified because of their potential to provide vital information to the study that is relevant to the research questions (Campbell et al., 2020). The study population targeted the general public, and all possible efforts were made in order to make the participants diverse in terms of age, sex, education level, occupation, and experience with the AI technologies. Such demographic variety is important for understanding as many attitudes toward AI as possible, starting with those who work in industries using the technology sometimes experimentally to those who quite rarely interact with it in their daily practices.

In this study, 90 participants were selected from different organizations, Internet groups, and social networking sites. They were picked from San Francisco, California; a city that is associated with technological advance and a cosmopolitan population. This location was selected to gather various opinions of people about AI since the city is a hub of technology and it comprises people of different backgrounds. The inclusion criteria allowed only subjects who were at least 18 years old and who had at least basic understanding of the concept of AI, even if they

had no technical background. This criterion was chosen as it allows the participants to really interact with the topic not excluding those who may use AI on a casual basis as well as those who work with it every day.

3.3. Data Collection

Data were collected using semi-structured interviews which made this research more directional while at the same time being very free to capture participants' attitude towards AI. This paper's interview guide was designed from the research questions and early literature regarding public attitude towards artificial intelligence (De Sousa, De Melo, Bermejo, Farias, & Gomes, 2019). To achieve this, it comprised questions that would allow the gathering of more elaborate information regarding the participants' perceptions of risks, opportunities, trust and acceptance of AI. Example questions included:

- "What comes to mind when you think about Artificial Intelligence?"
- "Can you describe any experiences you've had with AI technologies?"
- "How do you feel about the increasing role of AI in various aspects of society?"
- "What concerns, if any, do you have about AI?"
- "What benefits do you see AI providing, and how do they affect your view of AI?"

Face-to-face interviews or interviews with use of video conference was employed following the participant's preference or what was convenient to the participant. They took about 45 to 60 mins each and all the interviews were recorded with participants' permission for the purpose of transcription and analysis. The interviews took about 6 months, so enough time was given to recruit the participants and make arrangements for the interviews.

3.4. Data Analysis

To undertake the analysis of the data, thematic analysis process was employed and this was based on the framework developed by Braun & Clarke (2006). Since the aim of the study was to examine and compare the respondents' experiences, perspective and perceptions, thematic analysis was applied as it allows for data analysis and reporting of patterns within the data. The analysis proceeded in the following stages:

1) Familiarization with the Data: The first procedure was taken to transcribe the interviews in details and then to recounting the tapes several time to ensure adequate orientation on the content.

2) Generating Initial Codes: The next stage included analysis of the collected data The simplest or basic analysis done in this study was coding the data. This process involved underlining of portions in the text that seemed important or interesting concerning the research questions. Manual coding was adopted whereby data was analysed using qualitative data analysis software known as NVivo that helps in storing, retrieving and sorting codes.

3) Searching for Themes: Coding was done in a systematic manner and the

codes were sorted and categorized in to themes that encompassed visible patterns in the data collected. Themes were therefore identified inductively thus giving the researcher the natural opportunity to code the themes rather than having to force the data into pre-determined boxes.

4) Reviewing Themes: These are the emerged themes that were finally evaluated and cleaned up to meet the research questions and data collected. This process called validation involved comparing the identified themes with the entire data set for consistency.

5) Defining and Naming Themes: Concerning the above-mentioned processes, it is in this stage that the issues under analysis are named in order to records their essential features, i.e., the themes. Each theme was also discussed with the aid of participant quotes where relevant in elaborating on each of the established themes.

6) Writing Up: The last part included the process of combining the themes in dealing with the research questions, and putting the research into context of prior research studies.

3.5. Ethical Considerations

The findings of this research conformed to ethical practices for conducting research among people. Earlier this any participant in the study received written information about the objectives of the study, their role in this process, as well as the necessary methods and measures to protect the participant's confidentiality and anonymity. Each of the participants signed a consent form that indicated he or she authorized the research, agreed to cooperate fully and understood his or her rights; right to withdraw from the study at any time with no punishment.

Personal identifiers of all participants were omitted upon data collection, and pseudonyms were employed in the place of the participants' real names in the transcriptions and summaries of the studies. The data were kept secured and retrievable only by the authorized team members of the study and all the collected data will be deleted after a period that is admissible according to the institutional data disposal policies. Also, the study obtained approval from the Ethical Review Board (ERB) of the host institution, to make sure the conduct of the research was in harmony with the laid down ethical practices.

4. Results

4.1. Overview of Themes

The study of the interview responses obtained from ninety respondents led to the emergence of several themes associated with people's attitudes to AI. These themes shed light on the working of the individuals in the context of trust, risk, and opportunities of AI. Some of the major themes that are being discussed include Major Theme 1: Ambivalent Attitudes towards AI, Major Theme 2: Trust and Distrust, Major Theme 3: Positive and Negative Impacts, Major Theme 4: Desire for Regulation and Transparency. The detailed presentation of these themes is made below together with the specific percentages that captured the respondents with

such perceptions.

4.2. Ambivalence toward AI

Altogether, 65% of the respondents reported that they are uncertain toward AI, which means that they recognize that AI both can be helpful and dangerous to them. This was, in many cases, a thoughtful understanding of AI's potential and great performance combined with concern over its potential consequences. Concretely, as shown in Figure 1, 42% of the respondents supported the AI for improving the daily life by facilitating the tasks, however, the positive approach towards AI is accompanied by concerns about AI-related unemployment and risk of people losing the control over systems. Indeed, one of the participants said, "I gave full marks for AI for how it can make things such as shopping or customer service quicker and way more personalized but at the same time I gave full marks for concern that we have been outsourcing our decision-making powers to machines". This theme relates to the mixed feelings that people have in relation to technological developments, and the positive and the negative in equal measure.

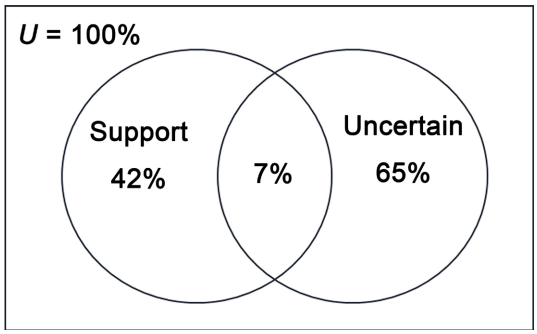

Figure 1. Venn diagram showing the support and uncertainty for AI among respondents.

4.3. Trust and Distrust

One of the interesting findings was that trust in AI was an antecedent of the AI technologies acceptance by the respondents. As represented in Figure 2, the most concern was reported in the employment of AI with about 58% of the respondent's voicing pessimism in areas such as the manner AI works or the motives of the firms that create the AI applications. Such scepticism tended to be accompanied by ignorance of how AI actually functions or potential reasons behind the use of AI.

Another respondent articulated this concern as follows: "I still doubt the ability of something that is hard for me to understand, AI sounds like the black box—decisions are being made but nobody knows how or why and, to be honest, who is going to stop the companies from making a profit from us."

Yet 36% of the respondents said they would be more trusting with AI, if they knew more about the inner workings of AI and if there are better checks and balances in place to prevent exploitation.

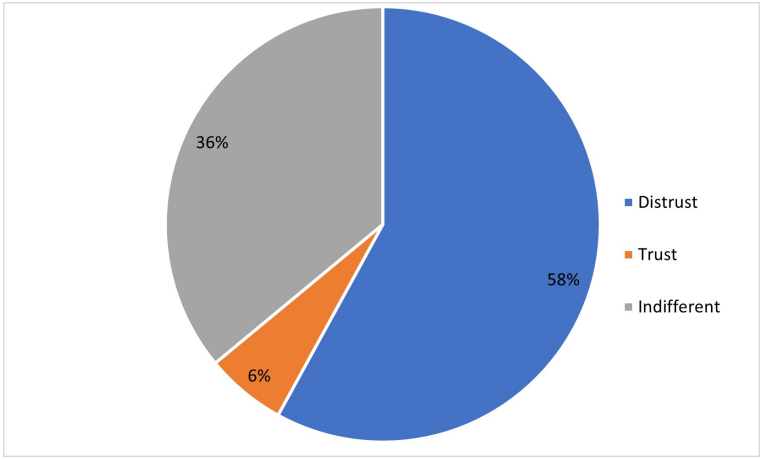

Figure 2. Attitude towards AI: trust and distrust.

4.4. Perceived Risks

Potential threats linked with AI emerged as another broad concern resolving around the possible dangers posed by artificial intelligence. According to our respondents as can be seen in Figure 3, the most common risk linked to AI was job displacement: 60% of the participants were concerned that AI could cause the dismissal of human workers in different fields. Unsurprisingly, this concern was most strongly expressed for respondents engaged in manual or routine forms of employment and those states most likely to see automation of their job roles.

Indeed, one of the participants working in the retail said, "I look at self-checkout machines and think one day there won't be cashier positions anymore and people like me will be out of work."

Further risk regarding the AI advancement indicated by 47% of respondents was the risk where bias was introduced and upheld reinforcing inequities. They feared that AI systems once developed may even reinforce prejudice in such areas as employment, access to credit and policing.

"AI is only as good as the data it is trained on", said one respondent. "If that data is biased, then the AI will be biased too. That is a lot of unfair decisions being made". It is worth to note, that privacy was named as a key challenge too with 52% of participants stating that they feel uncomfortable that their personal data is processed and analysed by AI. These feelings of concern were particularly brought out mainly in relation to surveillance and misuse of data noting other broader issues of data privacy.

However, other risks rarely mentioned by 13% of respondents includes liability, cyber security, misinformation and transparency.

4.5. Perceived Benefits

55% of respondents said there is a lot of gain as with the targeted application of AI; it improved people's convenience in today's busy and fast-paced society. For instance, 38% of the respondents identified AI's usefulness in providing customized services, product recommendations or identifying the rightful route in

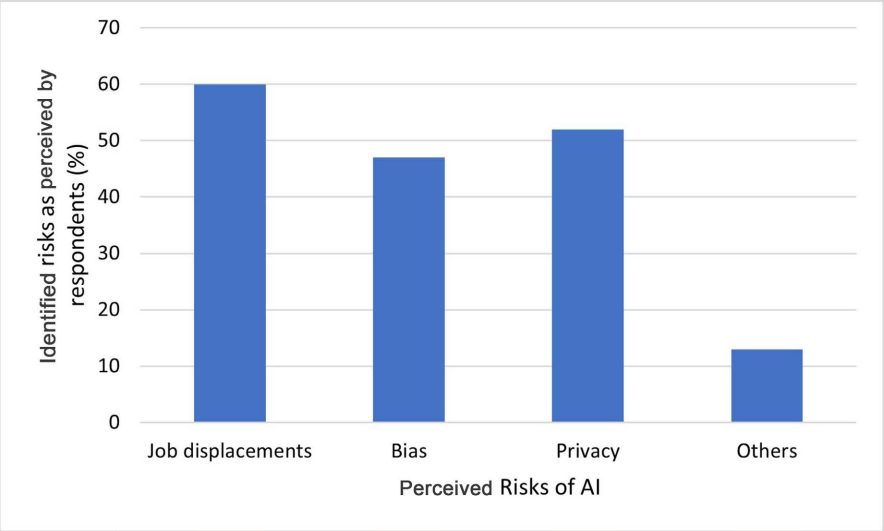

Figure 3. Perceived risks by percentage.

travelling among others. One of the respondents commented that the ability to be given directions and avoid traffic and the ability to be given music that is believed to be suitable for the user is like having a personal assistant. Apart from ease, an equal number of the respondents perceived AI to have positive impacts in fields such as healthcare and environmental segments. There is a "can-do" mentality about AI addressing difficult challenges at work involving healthcare where disease diagnosis is expected to be performed more effectively or how energy should be managed in a way that will reduce the adverse effects on the world's climate. It remains, though, that there is certain potential: "AI could really help with things like climate change by making our use of resources… more focused," a respondent added. "It may sound quite trivial and simple perhaps it can ease life so much and at the same time save the earth". This theme demonstrates that the people acknowledge that AI has a positive influence on the society where the technology is applied, especially to the issues of the world.

4.6. Desire for Regulation and Transparency

Perhaps the most revealing insight into the feeling of the population towards the use of AI can be concluded from a combination of the following excerpted questions and answers given during interviews; Interviewer; Is there anything that you would like to see change when it comes to the use of AI? About 62% of respondents said that there was a need to have a strict policy that can govern the use of AI to avoid misuse of the technology. The participants also agreed that the AI technologies in use today should be regulated to prevent such effects such as bias and job loss, or invasion of the privacy of the users. One respondent said, "There should be more regulation, which will keep prejudice of AI away from people, there should be legislation that will compel organisations into being responsible for how they use AI."

Transparency was also another factor with majority of the participant at 54%

saying that companies and developers should reveal more information about how the AI work and make decisions. Some of them gave an indication that the level of transparency must be improved so that they can easily accept the AI Technologies. "I believe if people knew more about how those systems function, they would be more trusting." Said one of the respondents. "There is still much of it that seems to occur out of sight, and that makes it also hard to believe…" This theme raises awareness about the role of effective laws and corporate management in building people's trust in artificial intelligence.

5. Summary of Findings

The views of this study show that perceptions about AI are not as straightforward as some previous research suggested. There is certainly awareness of AI's advantages, especially in terms of optimisation and in response to the world's problems, but the negative outlooks regarding the impacts of job loss, social unfairness, invasion of privacy and the lack of explication of how those impacts come about are strongly felt. Trust becomes the prominent feature; a third of the participants were not satisfied with how AI was being developed and implemented and called for more regulation into this. Thus, the presented results indicate the need for further action to fight the negative attitudes toward AI by enhancing the transparency of the actions, providing the ethical requirements, and establishing the strong protective legislation. Thus, this study also helps to fill the gap in the discussions regarding AI by identifying how ordinary people understand and weigh pros and cons of AI, and the findings made can be useful for AI's further evolution and utilization.

6. Discussion

6.1. Interpretation of Key Findings

This study has supplied the scholarly community and the general public with useful information about the diverse and multilayered ways through which the general public feels about AI. The study also notes an enormous acceptance of AI in the same way people embrace innovation or new technologies, but at the same time, people have many concerns as to AI's possible negative effects. Such an attitude is rather ambiguous and can be seen as an expression of other social concerns regarding the progress of technological innovations in people's lives.

This work also revealed that trust and scepticism were two of the most important trends that people felt about artificial intelligence. Participants understood the potential of the speeds and convenience AI can offer, but they also showed a rather primal fear with regards to AI and the organizations that use them. This scepticism is in line with earlier studies showing that trust in AI is quickly ruined by opacity of algorithms, and apparent lack of responsibility for AI interventions (Robinson, 2020; Purves & Davis, 2022).

Coveted risks about AI, with emphasis on job replacement, bias in AI solutions and privacy infringement formed the basis of participants' perceptions. Such risks

are in line with the literature that claims that these risks are major fundamental challenges to the broad adoption of AI (Acemoglu, 2021; Chaudhary & Tyagi, 2019). The notion of losing a job, be it real or imagined, is a signal of an acute anxiety, which is shared by societies all over the world regarding the future of work mitigated by the advancement of artificial intelligence. Other concerns raised by participants regarding AI include the reinforcement of bias and violation of individual rights are other aspects of the ethical implications of developing and implementing of AI.

However, the study also found signs of awareness of the considerable potential that AI has, including the spheres like health care, environmental conservation, and general ease in daily life. This work, therefore, propounds the fact that though the public is sceptical of AI, there is latent positivity toward the technology and the solutions that it holds to some of the society's biggest problems. Such optimism is due to the fact that like any potential technology, if applied responsibly AI is accepted as a technology which can and will, do more good than harm to the society at large (Toll, Lindgren, Melin, & Madsen, 2020; Damioli, Van Roy, & Vertesy, 2021). In Interview, one common wish was expressed to have more regulation and clarity regarding AI. This showed that participants felt a prescriptive need for ethical and regulatory measures in regards to the development and usage of AI systems that are non-partisan, responsive and transparent. This finding supports the increasing voices within the scholarly and policymaking circles for proper principles of AI that can eliminate the side effects of the technology in addition to enhancing its advantages (Makridakis, 2017; Liang et al., 2022).

6.2. Some Implications for AI Development and Deployment

The findings of the present study hold important implications regarding the advancement and application of AI systems. First of all, the given results can be discussed focusing on the values related to the relevance of trust and its circumstances such as openness and accountability. Businessman and software engineers should reduce the complexity of a process of the use of artificial intelligence. This could entail producing simpler and more engaging introductions to how AI systems operate, or clarifying what data was used to "teach" these systems and that they might not be free of prejudice.

In addition, the research shows that there is a need for the governments to come up with thorough legal policies which capture ethical issues that relate with AI. Policy makers should talk about the creation of the rules for the minimization of the risks which are connected with job replacement, employment discrimination, and violations of the privacy but also for the generating of the conditions for the reasonable and fair AI usage. Such regulations should be developed in consultation with the public in order to respect the nation's values since they are indeed regulations of the nation. The need for ethical AI is perhaps supported by the fact that there is a market for AI products and services that are FAT. Some of these principles can include how the data is collected, how the data is used and ensuring

that consumers are aware of this information. For organisations that embrace and incorporate these principles, their AI technologies or systems may likely enjoy the approval and acceptance of the public. This could include the adoption of the ethical AI frameworks that are under the FAT principles as well as participating in engagements that would ensure that the stakeholders are also implementing ethical AI systems that are socially acceptable.

6.3. Contribution to Existing Literature

This research expands upon the current literature insofar as it offers a well-developed qualitative analysis of how the public experiences embodied by "ordinary" people understand the risks and benefits of AI. Although a lot of studies have employed quantitative and qualitative method to analyse the efficiencies and the ethical issues of AI from the standpoint of experts, this research also helps to get an understanding of the general populace's fears and expectations in relation to AI.

The findings also contribute to the existing knowledge of how people are "of two minds" about AI. Whereas the previous research concluded that there is a balanced mixture of positive and negative views of people towards AI, this work gives a better understanding of how these perceptions are developed and how optimism and scepticism are embedded in the opinions about the future of AI. In this sense, by looking at the concrete fears in the form of scepticism that is rooted in AI concern, this study provides a better understanding of what has to be done to ensure the public trust in the AI. Furthermore, the study prescribes the urgent call for more regulation and enlightening practices, which supports the discourse on Artificial Intelligence's governance rather strenuously stresses the need for ethical norms and adequate legal frameworks that would be safe and suitable for the public. This contribution is timely given that the application of AI technologies is gaining momentum in the many sectors across the society, thus points are very relevant when asking how to address and manage misuse and malfunctions of these systems.

6.4. Limitations and Future Research

However, it is also important to understand that this study as valuable as it is for understanding the public perception of AI has some limitations. The main drawback is the rather limited and geographically narrow range of participants' views, which may not be typical for the rest of the world. Possible future research could work around this by using larger and different target samples, and even carry out cross-cultural studies in an attempt to establish whether or not perceptions of AI differ cross culturally. One of the limitations is the use of self-reported data only, which may be affected by such bias as the desire to present oneself in a positive light or participants' insufficient knowledge of AI. As the current study targeted participants with a wide range of experience with AI, future works could include enhancing the educational part or using a combination of survey questionnaires, observations, and experiments. Last of all, the further studies might be dedicated

to the analysis of the impact of educational campaigns on attitudes toward AI. Based on the findings of this study, enhancing the awareness of AI can contribute to the generation of trust and pointing to the further research about the effective means of increasing the trust and understanding of the population with regard to AI. It could be used to design awareness creating campaigns or, perhaps educational initiatives with a view of promoting a healthier perception of AI systems.

7. Conclusion

The aim of this research was to find out how the public perceive AI, with special emphasis on the opportunities and threats of AI and the way these perceptions shape the trust and acceptance of AI solutions. While carrying out the research, 39 interviews were conducted with a diverse cross-section of the population, which allowed identifying several important trends, which helped to reveal diverse positive and negative attitudes towards AI.

These insights show that while there is a lot of understanding regarding the opportunities that are accruable from AI such as efficiency gains, customization, enhanced innovation in areas like healthcare and environmentalism, there is also alarm. Out of the different issues, job loss, unfair dismissal, privacy violation and the fact that most AI are opaque posited a lot of work against the general acceptance by the public. Such apprehensions are coupled with the fact that a significant number of participants perceived AI technologies as "black boxes", that is, systems of which they are unable or find it difficult to model or inspect carefully. Indeed, one of the study's biggest accomplishments is to ascertain the high demand of the public in terms of more regulation and disclosure of Artificial Intelligence. Whilst respondents wanted to ensure that AI was used beneficially to promote equity and fairness, they were equally concerned that there must be standard ethical procedures that would reduce the effects of the dangers of AI. This finding corroborates the emerging belief within the literature, as well as in the global policy and technological communities, that strong governance is the only way to build trust in AI.

The findings of this study with particular reference to public attitudes to AI point to the benefit of conducting a more extensive dialogue regarding the future of AI. The developers, companies, and policymakers should be more open to declare what they are doing and develop strategies to overcome all those ethical issues of AI. In this way, they will be able to contribute to the creation of a more accurate relationship between the members of the society and the AI technologies, in order to pave the way to the responsible innovations for the benefit of everyone. All in all, this research benefits the field of AI by offering a complex view of how laypeople evaluate the opportunities and threats of AI. It points to deficit in the current ways of making AI and the regulation of AI more open and democratic so that the public have a say in how the technology is developed and managed. This means that as AI progresses and it is integrated into society, it will be essential to ensure that the public has confidence in the technology. The research should be

continued in the future focusing on the ideas that have been investigated in the course of this work to increase awareness of the public and to reach the further effective perspective of the attitudes towards AI.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

Acemoglu, D. (2021). *Harms of AI.* National Bureau of Economic Research.

Almeida, D., Shmarko, K., & Lomas, E. (2022). The Ethics of Facial Recognition Technologies, Surveillance, and Accountability in an Age of Artificial Intelligence: A Comparative Analysis of US, EU, and UK Regulatory Frameworks. *AI and Ethics, 2,* 377-387. https://doi.org/10.1007/s43681-021-00077-w

Amini, M., Bagheri, A., & Delen, D. (2022). Discovering Injury Severity Risk Factors in Automobile Crashes: A Hybrid Explainable AI Framework for Decision Support. *Reliability Engineering & System Safety, 226,* Article 108720. https://doi.org/10.1016/j.ress.2022.108720

Bélisle-Pipon, J., Monteferrante, E., Roy, M., & Couture, V. (2023). Artificial Intelligence Ethics Has a Black Box Problem. *AI & Society, 38,* 1507-1522. https://doi.org/10.1007/s00146-021-01380-0

Braun, V., & Clarke, V. (2006). Using Thematic Analysis in Psychology. *Qualitative Research in Psychology, 3,* 77-101. https://doi.org/10.1191/1478088706qp063oa

Campbell, S., Greenwood, M., Prior, S., Shearer, T., Walkem, K., Young, S. et al. (2020). Purposive Sampling: Complex or Simple? Research Case Examples. *Journal of Research in Nursing, 25,* 652-661. https://doi.org/10.1177/1744987120927206

Chanda, S. S., & Banerjee, D. N. (2024). Omission and Commission Errors Underlying AI Failures. *AI & Society, 39,* 937-960. https://doi.org/10.1007/s00146-022-01585-x

Chaudhary, V., & Tyagi, S. (2019). A Study of the Impact of AI on the Job Market More Opportunities & More Threats. *African Diaspora Journal of Mathematics, 22,* 166-169.

Chou, W. S., & Gaysynsky, A. (2023). A Relationship-Centered Approach to Addressing Mistrust. *Journal of Communication in Healthcare, 16,* 320-323. https://doi.org/10.1080/17538068.2023.2258683

Cossette-Lefebvre, H., & Maclure, J. (2023). AI's Fairness Problem: Understanding Wrongful Discrimination in the Context of Automated Decision-Making. *AI and Ethics, 3,* 1255-1269. https://doi.org/10.1007/s43681-022-00233-w

Cui, D., & Wu, F. (2021). The Influence of Media Use on Public Perceptions of Artificial Intelligence in China: Evidence from an Online Survey. *Information Development, 37,* 45-57. https://doi.org/10.1177/0266666919893411

Damioli, G., Van Roy, V., & Vertesy, D. (2021). The Impact of Artificial Intelligence on Labor Productivity. *Eurasian Business Review, 11,* 1-25. https://doi.org/10.1007/s40821-020-00172-8

De Sousa, W. G., De Melo, E. R., Bermejo, P. H., Farias, R. A., & Gomes, A. O. (2019). How and Where Is Artificial Intelligence in the Public Sector Going? A Literature Review and Research Agenda. *Government Information Quarterly, 36,* Article 101392. https://doi.org/10.1016/j.giq.2019.07.004

Dieter, D. G., & Gessler, E. C. (2021). A Preferred Reality: Film Portrayals of Robots and AI in Popular Science Fiction. *Journal of Science & Popular Culture, 4,* 59-76.

https://doi.org/10.1386/jspc_00025_1

Ernst, E., Merola, R., & Samaan, D. (2019). Economics of Artificial Intelligence: Implications for the Future of Work. *IZA Journal of Labor Policy, 9,* 1-35. https://doi.org/10.2478/izajolp-2019-0004

Gupta, A., & Mishra, M. (2022). Ethical Concerns While Using Artificial Intelligence in Recruitment of Employees. Business Ethics and Leadership. *Business Ethics and Leadership, 6,* 6-11.

Haleem, A., Javaid, M., & Khan, I. H. (2019). Current Status and Applications of Artificial Intelligence (AI) in Medical Field: An Overview. *Current Medicine Research and Practice, 9,* 231-237. https://doi.org/10.1016/j.cmrp.2019.11.005

Hobson, Z., Yesberg, J. A., Bradford, B., & Jackson, J. (2023). Artificial Fairness? Trust in Algorithmic Police Decision-Making. *Journal of Experimental Criminology, 19,* 165-189. https://doi.org/10.1007/s11292-021-09484-9

Hoffmann, A. L., Roberts, S. T., Wolf, C. T., & Wood, S. (2019). Beyond Fairness, Accountability, and Transparency in the Ethics of Algorithms: Contributions and Perspectives from Lis. *Proceedings of the Association for Information Science and Technology, 55,* 694-696. https://doi.org/10.1002/pra2.2018.14505501084

Kaack, L. H., Donti, P. L., Strubell, E., Kamiya, G., Creutzig, F., & Rolnick, D. (2022). Aligning Artificial Intelligence with Climate Change Mitigation. *Nature Climate Change, 12,* 518-527. https://doi.org/10.1038/s41558-022-01377-7

Kahma, N., & Matschoss, K. (2017). The Rejection of Innovations? Rethinking Technology Diffusion and the Non-Use of Smart Energy Services in Finland. *Energy Research & Social Science, 34,* 27-36. https://doi.org/10.1016/j.erss.2017.05.024

Li, F., Ruijs, N., & Lu, Y. (2023). Ethics & AI: A Systematic Review on Ethical Concerns and Related Strategies for Designing with AI in Healthcare. *AI, 4,* 28-53. https://doi.org/10.3390/ai4010003

Liang, W., Tadesse, G. A., Ho, D., Fei-Fei, L., Zaharia, M., Zhang, C. et al. (2022). Advances, Challenges and Opportunities in Creating Data for Trustworthy AI. *Nature Machine Intelligence, 4,* 669-677. https://doi.org/10.1038/s42256-022-00516-1

Makridakis, S. (2017). The Forthcoming Artificial Intelligence (AI) Revolution: Its Impact on Society and Firms. *Futures, 90,* 46-60. https://doi.org/10.1016/j.futures.2017.03.006

Nakashima, H. H., Mantovani, D., & Machado Junior, C. (2024). Users' Trust in Black-Box Machine Learning Algorithms. *Revista de Gestão, 31,* 237-250. https://doi.org/10.1108/rege-06-2022-0100

Prianto, Y., Sumantri, V. K., & Sasmita, P. Y. (2020). Pros and Cons of AI Robot as a Legal Subject. In *Proceedings of the Tarumanagara International Conference on the Applications of Social Sciences and Humanities (TICASH 2019)* (pp. 380-387). Atlantis Press. https://doi.org/10.2991/assehr.k.200515.067

Purves, D., & Davis, J. (2022). Public Trust, Institutional Legitimacy, and the Use of Algorithms in Criminal Justice. *Public Affairs Quarterly, 36,* 136-162. https://doi.org/10.5406/21520542.36.2.03

Rakowski, R., Polak, P., & Kowalikova, P. (2021). Ethical Aspects of the Impact of AI: The Status of Humans in the Era of Artificial Intelligence. *Society, 58,* 196-203. https://doi.org/10.1007/s12115-021-00586-8

Robinson, S. C. (2020). Trust, Transparency, and Openness: How Inclusion of Cultural Values Shapes Nordic National Public Policy Strategies for Artificial Intelligence (AI). *Technology in Society, 63,* Article 101421. https://doi.org/10.1016/j.techsoc.2020.101421

Shank, D. B., & DeSanti, A. (2018). Attributions of Morality and Mind to Artificial Intelligence after Real-World Moral Violations. *Computers in Human Behavior, 86,* 401-411. https://doi.org/10.1016/j.chb.2018.05.014

Shin, D., & Park, Y. J. (2019). Role of Fairness, Accountability, and Transparency in Algorithmic Affordance. *Computers in Human Behavior, 98,* 277-284. https://doi.org/10.1016/j.chb.2019.04.019

Thadeshwar, H., Shah, V., Jain, M., Chaudhari, R., & Badgujar, V. (2020). Artificial Intelligence Based Self-Driving Car. In *2020 4th International Conference on Computer, Communication and Signal Processing (ICCCSP)* (pp. 1-5). IEEE. https://doi.org/10.1109/icccsp49186.2020.9315223

Tippins, N., Oswald, F., & McPhail, S. M. (2021). Scientific, Legal, and Ethical Concerns about Ai-Based Personnel Selection Tools: A Call to Action. *Personnel Assessment and Decisions, 7,* Article 1. https://doi.org/10.25035/pad.2021.02.001

Toll, D., Lindgren, I., Melin, U., & Madsen, C. Ø. (2020). Values, Benefits, Considerations and Risks of AI in Government. *Journal of E Democracy and Open Government, 12,* 40-60. https://doi.org/10.29379/jedem.v12i1.593

Tuli, S., Gill, S. S., Xu, M., Garraghan, P., Bahsoon, R., Dustdar, S. et al. (2022). HUNTER: AI Based Holistic Resource Management for Sustainable Cloud Computing. *Journal of Systems and Software, 184,* Article 111124. https://doi.org/10.1016/j.jss.2021.111124

von Eschenbach, W. J. (2021). Transparency and the Black Box Problem: Why We Do Not Trust AI. *Philosophy & Technology, 34,* 1607-1622. https://doi.org/10.1007/s13347-021-00477-0