



Perceptions of Potentials and Risks of Artificial Intelligence Across Diverse Levels of AI-literacy

Zeba Khadhijah¹ and Bhupesh Kumar Mishra^{2,*}

¹*School of Business, Computing and Social Sciences, University of Gloucestershire, UK*

²*Centre of Excellence for Data Science, Artificial Intelligence and Modelling (DAIM), University of Hull, UK*

Abstract: Perceptions of the potential benefits and risks of artificial intelligence (AI) vary widely among people with different levels of knowledge about the technology. By better understanding these variations, we can identify and address challenges in the adoption and responsible use of AI for all. This study employs a qualitative approach to explore these differences through criterion sampling. A total of 15 participants who self-identified as belonging to one of three AI-literacy categories (Limited, Intermediate, and Expert) were interviewed. The audio-recorded and transcribed interviews were thematically analyzed, identifying differences in understanding among different literacy groups regarding AI's impact on employment, creativity, bias, trust, data privacy, and the need for regulation and awareness. The study reveals that varying levels of AI-literacy shape perceptions of bias, privacy, creativity, job security, and trust. A key finding was that participants with limited AI-literacy did not link AI to data and therefore did not associate concerns like data privacy with AI, and also did not believe that AI can be biased, as they viewed bias as a human quality rather than something a machine could possess. These findings underline the need for initiatives that strengthen AI-literacy across different AI user groups to enable informed and equitable engagement with AI technologies, while also highlighting the need for education, transparency, and policy measures that build trust and accountability. In addition, how diverse users understand and interact with AI has important implications for developers, educators, and policymakers seeking to advance explainability and inclusive literacy.

Keywords: artificial intelligence, AI-literacy, perceptions of AI, bias in AI, trust in AI, need for AI awareness and regulation

1. Introduction

Artificial intelligence (AI) advancements have brought about transformative benefits for humans across various domains. AI is often described as 'the ability of machines to perform tasks that normally require human intelligence' [1]. However, this definition can be confusing because there are many tasks that AI can perform in ways human intelligence cannot, such as processing millions of data points in seconds, and many tasks humans can do in ways that AI simply cannot, such as critical thinking and reasoning based on real-world experience, demonstrating empathy, etc. Perhaps this lack of clarity about what AI is allows people to perceive AI in very different and sometimes inaccurate ways. Fictional representations have painted an image of AI taking over the planet and posing an existential threat to humans [2].

On the other hand, news narratives or marketing efforts by companies creating AI may overly emphasize the immense potential of AI, which may cause people to ignore the limitations of AI technology and demonstrate 'automation bias', that is, blindly trusting outputs from machines to always be accurate [3]. For example, Wang and colleagues' [4] experimental study found that people with low computer literacy provided lower fairness evaluations of algorithms than people with high computer literacy and that there was a bias

to rate an algorithm as fairer if they received a more favorable outcome, especially among participants who have attained a lower level of education. Misconceptions and biases like these become a barrier for people to embrace a technology that can make their lives better in revolutionary ways. Perspectives on AI advancement vary widely, particularly between AI developers and everyday users, who often differ in both their understanding of the technology and their expectations of its capabilities. For instance, previous literature notes that many explainable AI (XAI) techniques seek to explain the working of models to those with a background in data and technology and do not cater to laymen [5]. This can lead to inequalities by not adequately equipping some groups of users to leverage AI to access the full scope of its benefits, thus slowing down socio-technological progress. Therefore, to create more user-friendly services, there is a strong need for more research to help developers of AI understand how all their end users understand and interact with AI. Secondly, in a world where AI is rapidly advancing, governments need to think about how to help people become more AI literate so that they can adapt to their evolving workplace, as well as promote the safe and responsible use of AI.

To design effective interventions to improve the AI-literacy of the diverse public, we need to know where they currently stand in their understanding and use of AI. Few existing studies explore people's perceptions of AI and AI-anxiety [4, 6]. and the general perception of AI [7, 8]. The present study aims to examine how people with different levels of familiarity with AI technology perceive AI and the various ethical issues surrounding it, through in-depth semi-structured

*Corresponding author: Bhupesh Kumar Mishra, Centre of Excellence for Data Science, Artificial Intelligence and Modelling (DAIM), University of Hull, UK. Email: bhupesh.mishra@hull.ac.uk

interviews. To the best of our knowledge, no existing studies have undertaken comparisons across different AI-literacy categories, making this study a novel contribution to the field. The analytical findings revealed that participants across different AI-literacy levels expressed both optimism and concern, with lower literacy groups showing misconceptions about bias, data use, and privacy, while higher literacy groups emphasized transparency and explainability. The study also highlights that AI is seen as useful but raises issues around creativity, job security, fairness, and trust, underscoring the need for education, clear communication, and strong policies. Overall, the study demonstrated that improving AI-literacy and accountability is essential to ensure equitable and responsible adoption.

2. Literature Review

2.1. Narratives and their influence on public perception of AI

Portrayals of AI in popular culture can influence the way we perceive AI. AI in fictional narratives is often portrayed in a humanoid form [2], and this anthropomorphism can cause us to visualize AI as more than just a tool. Fictional representations have caused people to have either exaggeratedly utopian views of what AI can achieve (e.g., eliminate poverty) or amplified dystopian views (e.g., robots take over the world) [9]. In a survey study conducted by Cave et al. [10] with a sample that was nationally representative of the UK population, it was found that 55% of participants feared that AI would make humans obsolete, and 44% feared that AI may enable computers to become more powerful than humans. However, in recent years, more practical and immediate concerns about AI have made their way from academic debates into mainstream media. Ouchchy et al. [11] examined the portrayal of various ethical issues relating to AI in newspaper articles and found that more practical concerns (e.g., safety, data privacy) have been discussed in the media in recent years, often catalyzed by widely discussed events such as the Tesla autopilot accident. With the introduction of ChatGPT at the end of 2022, a new record was set for the fastest user adoption of a new application [12] and discussions around generative AI and large language models (LLMs) became more popular. People’s understanding of the power of AI expanded beyond what they had seen in movies and books, and new concerns such as the accuracy of generated content and the lack of transparency by companies building these models, were brought to the forefront [10].

2.2. Theoretical underpinnings: technology acceptance model

While AI tools hold transformative potential, whether they are effectively adopted or integrated ultimately depends on how people perceive and engage with them. The most popular theoretical framework exploring what affects our technology adoption and use is the technology acceptance model (TAM). Originally proposed in 1989, the model outlined how perceived usefulness, perceived ease of use determine if a technology will be adopted and used [11]. The model has been continuously expanded over the years to include the effects of variables like social influences, computer self-efficacy and computer anxiety [12]. Drawing on this model, the present study seeks to explore similar factors of social perceptions, trust in AI and AI-anxiety to understand the unique challenges towards responsible AI adoption. In a recent study by Schiavo and colleagues, variables such as AI-anxiety and AI-literacy have been explored through the lens of the TAM, demonstrating the influence of these variables in perceptions of AI and, therefore, its acceptance and adoption [13].

2.3. AI-anxiety

Historically, new technology is often initially met with anxieties over its potential. For instance, ‘automation anxiety’ became prominent during the Industrial Revolution when machinery started replacing manual labor, and ‘computer anxiety’ gained popularity in the 1980s [13]. Johnson and Verdicchio [14] described AI-anxiety as the fear of out-of-control AI but rejected its validity and indicated that AI-anxiety is a result of misconceptions, such as a future fully controlled by AI that leaves out human beings. This confuses what autonomy means for humans (e.g., free will) with what autonomy means for a computer (e.g., creation of data based on parameters inside the machine). Li and Huang’s [13] study applied integrated fear acquisition theory and conducted a confirmatory factor analysis of questionnaires to identify the dimensions described below in Table 1, thus expanding our understanding of AI-anxiety beyond just the fear of a futuristic world controlled by AI. Similarly, Wang and Wang [15] created a tool to measure AI-anxiety, and their factor analysis revealed four dimensions — (i) the anxiety associated with learning and keeping up with various uses of AI, (ii) the fear that AI will replace humans and cause people to lose jobs, (iii) fears that AI will lead to loss of human control and robot autonomy, (iv) and being intimidated by AI.

Table 1
Li and Huang’s dimensions of AI-anxiety

Dimension of AI-anxiety	Description
Privacy violation anxiety	Anxieties associated with data leakage
Bias behavior anxiety	Concern about AI that discriminates unfairly
Job replacement anxiety	Fears that AI will cause humans to lose their jobs
Learning anxiety	Anxieties due to feeling that AI is too complex to learn and keep up with
Existential risk anxiety	Worry that AI may harm humans to achieve a goal or cause societal harm
Against ethics anxiety	Disturbed that people may form emotional bonds with AI, that AI can be deceptive, etc.,
Artificial consciousness anxiety	Anxiety that AI can achieve the same level of consciousness as humans and thus threaten the status of humans
Lack of transparency anxiety	Concerns that it is difficult to understand why AI has made certain decisions or where it has gone wrong

2.4. AI-literacy: why it matters

AI-literacy has been defined as an individual’s ability to critically evaluate, use, and effectively communicate with AI [16]. Long and Margeko provide a comprehensive AI-literacy framework identifying five areas of competencies: knowing what AI is, how it works, what it can do, how to use it ethically and understanding human–AI interactions and perceptions [16]. Similarly, Ng et al. [17] conducted a comprehensive review of AI-literacy literature and identified four aspects: knowing and understanding AI, applying AI knowledge, critically evaluating AI and understanding AI ethics. Building on these frameworks, the present study examines perceptions of AI related to bias, transparency, job replacement, and creativity across individuals with varying levels of AI-literacy.

People’s education and professional background tend to impact their AI-literacy skills. For example, Sartori and Bocca’s study [18] found that people with a computer science or IT background were much more likely to correctly identify applications that use AI, which indicates that individuals who study or work in a computer/information technology field are more likely to have better AI-literacy skills. From an AI developer perspective, it is important to understand how people of different backgrounds interact with and understand AI so that these applications cater to all users [19, 20]. Poor AI-literacy perpetuates misconceptions about AI, and this is a major cause of AI-anxiety [13, 14] which may hinder people’s acceptance of useful technology. People’s perceptions of AI, which are affected by their level of AI-literacy [20], could affect how AI tools are developed in the future and how they are regulated [11]. Therefore, there is a need to improve AI-literacy among the public to ensure the safe and responsible adoption of AI, and this begins with understanding how perceptions differ across literacy levels.

3. Study Design and Setup Methodology

This study adopts an exploratory qualitative design to examine how individuals with different levels of AI-literacy perceive AI, particularly in relation to ethical concerns, perceived risks and benefits and regulatory considerations. This study has been reviewed and approved by the ethics lead at the University of Gloucestershire.

3.1. Sampling technique

Criterion sampling method was used in this study -individuals known to the researcher and the research supervisor’s network were invited to participate if they fit the following inclusion criteria:

- 1) Must be over the age of 18.
- 2) Must consent to the interview being audio recorded.
- 3) Must reside in the United Kingdom.
- 4) Should belong to one of the following categories: a. Having limited knowledge and/or use of AI; b. Having an intermediate level of AI knowledge/Regularly using some form of AI; or c. Expert in the field of AI and data.

The researcher approached individuals who they deemed suitable for each of the above categories (5 participants per category), but the participants also self-identified their AI-literacy group in the short form completed before participating in the interview (which matched our estimation of their literacy levels in all 15 participants). As a small-scale exploratory study to yield initial insights into differences in perceptions of AI, non-probability criterion sampling is a pragmatic and relevant choice as it targets individuals whose literacy level falls within the inclusion requirements. Qualitative studies are known for their smaller sample sizes in comparison to quantitative studies. Ritchie et al. [21] explain that in qualitative research a phenomenon only needs to appear once to form an analytical point; the frequency of incidence is not necessarily a concern, as this study is not trying to establish statistical significance. Also, to do justice to the immense amount of information gained from a single data collection point, it is ideal and practical to keep sample sizes small.

It is important to note that the present study makes no claims of generalizability of its findings, as the 15 participants interviewed in the study cannot reflect the entire diversity of views that the whole population may have. The aim of qualitative research is never to generalize in a strict sense, but it has been argued that once the presence of an experience has been identified through qualitative research, we do know that it is available within a culture or society, even if we do not know exactly who and how many people experience it — and in this sense, there is some conceptual expansion and generalizability [22]. The goal of the present study is only to serve as a starting point to explore the presence of differences in AI perceptions among people with different backgrounds. The findings of the study will have important implications in identifying objectives and points of exploration for larger studies.

3.2. Participants’ characteristics

A total of 15 participants were included in the study [23]. All participants’ names have been changed to protect their anonymity. Tables 2, 3, and 4 summarize the characteristics of the participants and also their diversity, who together represent a broad spectrum of individuals interacting with AI at different levels.

Table 2
Participants characteristics — limited AI knowledge (Group 1)

Pseudonym	Gender	Age	Ethnicity	Background
Betty	Female	51	White/British	Administrative assistant
Dave	Male	20	White/British	TV production student
Samuel	Male	62	White/British	Church Minister
Aarti	Female	26	South Asian/British South Asian	Teaching assistant
Leela	Female	23	South Asian/British South Asian	Recent psychology graduate

Table 3
Participants characteristics — intermediate AI knowledge (Group 2)

Pseudonym	Gender	Age	Ethnicity	Background
Ahmed	Male	29	Arab/Middle Eastern	Data science student
Annie	Female	38	White/British	Business engagement head
Cathy	Female	23	White/British	Artificial intelligence student
Paul	Male	27	Black/African	Data science student
Ramya	Female	25	South Asian/British South Asian	Junior data scientist

Table 4
Participants characteristics — advanced AI knowledge (Group 3)

Pseudonym	Gender	Age	Ethnicity	Background
Margaret	Female	38	White/British	Founder and CEO of Data Science Consultancy
Hrithik	Male	42	South Asian/British South Asian	Engineer and PhD Data Science candidate
Sneha	Female	36	South Asian/British South Asian	Digital data officer
Shankar	Male	40	South Asian/British South Asian	Senior lecturer of data analytics
Dhruv	Male	44	South Asian/British South Asian	Professor of AI

3.3. Data collection

3.3.1. Tools and procedure

Based on the research questions and previous literature, relevant areas to explore with participants were identified, and interview question guides were formulated accordingly. Care was taken to ensure that all questions were open-ended, and participants were encouraged to elaborate on their answers. The interview guide was reviewed by an expert and deemed to be comprehensive and suitable for the present study. Informed consent and demographic information forms were distributed to all the participants, who were briefed about the study and possible risks, and they had the opportunity to ask any questions. Individuals who then voluntarily agreed to participate signed the consent form and filled out a short form that collects demographic information such as gender, age, ethnicity, etc., as well as other information relevant to the study, such as their professional background and the AI-based applications with which they are familiar. Following this, a time slot was agreed upon for the interview, either face-to-face or online. Interviews were audio recorded (with participants’ consent) and transcribed using Otter AI transcription software (basic version). The transcripts were manually proofread, and any incorrectly transcribed words or phrases were corrected. These transcripts were analyzed as described below.

3.3.2. Reflexivity

A qualitative researcher needs to demonstrate reflexivity [24], that is, to constantly examine their own biases, assumptions, values, and perspectives and be aware of how they may affect the research process [25]. In the present study, the primary researcher is aware that characteristics such as her age, gender, ethnicity or being a data science student, can unknowingly affect the way participants respond to the interview questions – for instance, if a participant was of a similar age to the researcher, it may cause them to be more comfortable and respond more candidly. Some participants were known acquaintances of the researcher, which may have had an impact on the rapport during the interview and the way they responded to the questions. The researcher is passionate about building unbiased and ethical AI and strongly advocates for diversity in the data workforce and for building explainable AI systems. Therefore, it is possible that the researcher’s personal bias can assign more importance to views that align with the researcher’s perception during the qualitative analysis process. The researcher has made every effort to bracket off this bias by avoiding any assumptions and performing qualitative coding with an open mind to ensure that final themes are reflective only of participants’ opinions and not her own. It is worth noting that many experts believe that the researcher’s personal bias should not be viewed as a demerit of qualitative work, but the subjective interpretation of a researcher who is aware of their biases is the essence of the qualitative research process that makes sense of the findings [26, 27].

3.4. Deductive thematic analysis

Thematic analysis (TA) is a method for identifying, analyzing, and interpreting patterns of meaning, known as ‘themes’, within qualitative data. We use the six-phase approach described by Bruan and Clarke [27]: (i) Researcher reads the data and familiarizes themselves with it, (ii) assigns labels to segments of text based on content called ‘codes’, (iii) collates codes into themes based on meaning, (iv) reviews themes and retains those relevant to research questions, (v) Ongoing analysis to define and refine the themes, (vi) producing the final report of findings/themes. This study uses deductive thematic analysis, that is, previous literature provided a framework that informed the qualitative coding process and interpretation of data to form the final themes. The qualitative coding and analysis process was done manually without the use of any software.

4. Results and Analysis

The transcribed data have been qualitatively analyzed under 7 main themes.

4.1. Impressions of AI and its potential

This theme was derived from participants’ responses to the open-ended question “What first comes to your mind when you hear the term AI or Artificial Intelligence?”

4.1.1. AI as a useful tool

Analysis of transcripts revealed that the majority of participants across the three groups of AI-literacy saw AI as a powerful tool that can make human life easier and more convenient. While all participants acknowledged the usefulness of AI tools, some participants noted that the user experience can be frustrating at times, for example, when the AI tool does not understand their accents (in the case of voice assistant devices) or when the content it generates is inaccurate and lacks references (e.g., answers provided by Chat GPT in some cases). Overall, participants acknowledged that there are both positive and negative aspects to this technology, but most showed optimism about AI’s potential. Figure 1 depicts the word cloud that appeared from the transcribed dataset, highlighting participants’ first impression of AI as a powerful tool associated with advancement and convenience.

4.1.2. AI in people’s minds

Ramya, 25, a female junior data scientist, says, “The beauty of it is that it’s become so user-friendly, and what we now call AI is not as scary anymore. It’s not walking robots with their intelligence. It’s just day-to-day things like creating a PowerPoint presentation template or a simple chatbot.” ChatGPT appears to be more commonly associated with AI than other applications, such as voice assistants or recommendation systems, especially among the participants in the low AI-literacy category, even though the latter

over low-risk everyday jobs may have a harsh economic impact on the common people.

Group 3 participant Margaret emphasized that advancements in AI do not mean that entire job roles would cease to exist, but just that certain tasks are removed from people, giving them more time to spend on other productive tasks. Participants also suggested that human expertise and connection are important in many job roles, and human oversight will always be needed in case technology fails. Shankar discusses the importance of education and policy to minimize the economic impact on people who may lose jobs due to rapid AI advancements: “There is no question that AI will impact workforces. The question is how we will collectively handle the impact. Countries and governments may start looking into investing in things like universal credit for people to gain additional training to better acclimatize to the digital and AI world.” Group 3 and Group 2 participants tended to emphasize the need for people to upskill and improve their AI-literacy to cope with the changing world, more than Group 1 participants did.

4.4. Bias in AI

Group 1 participants were more likely to believe that AI is unbiased, as they view it as a machine that is programmed to be neutral. “I suppose AI shouldn’t be able to treat people unfairly; it should be programmed to look past any kind of biases humans may have. So, where humans can judge people based on a whole range of characteristics, in theory, a computer should have the ability not to do that, I think.” — Betty. However, Group 2 and Group 3 participants were more aware of the issues of bias in AI due to poor data quality used for training AI systems. Margaret stresses the importance of data quality and expresses concern about people using AI without understanding that AI is only as good as the data you feed into it and that the world is filled with flawed data. While it may not be possible to remove bias in the natural world or control for unknown biases, participants suggest various strategies to mitigate bias: (i) Using carefully created balanced and updated training data sets, (ii) Thorough testing and repeated refining of AI tools before rolling out to the market, (iii) doing random checks with the output, (iv) using these tools sensibly with human instinct and alongside other tools to make better informed and fairer decisions, (v) and encouraging more women and individuals from other minority groups to take up careers in data and AI.

4.5. Data privacy

When it comes to data privacy, most of the participants were concerned about their data privacy and security, but the level of concern varied. Group 1 participants did not connect data and, therefore, data privacy, with AI developments. Betty and Aarti both said that while they have always been concerned with data privacy, they do not see how that relates to AI. Leela feels uncomfortable with the fact that her phone is always ‘listening’ and feels that it is too invasive. On the other hand, Dave feels that our data is already out there, and if one has not done anything wrong and has nothing to hide, they should not worry too much.

Group 2 participant Ahmed expresses concerns about data leakage and cyber-attacks, and how they can be harmful to the public. Participants in this group hold that tech companies should be stricter about data security, only collect and access data that they need, and follow the GDPR (General Data Protection Regulation) and other relevant guidelines strictly. Paul says that if he knows that a certain company or brand is not very big on data security, he will switch to a competitor that values data privacy more. While most Group 3 participants believe that data security is a concern, they also acknowledge that there may be a trade-off between convenience and privacy concerns, as can be seen in the following quotes:

“My phone recognizes my face and unlocks it — this is driven by AI. It has data about my facial features, and I’m allowing access. I haven’t really looked into how or where or in what form this data is stored. But I still make use of it because of the convenience it brings. Now, if I had to unlock my phone and, in the process, get a one-time password (OTP) to another phone or write a long, complicated password each time I’m unlocking it, it’s a huge inconvenience. So, there is this balance of convenience and privacy. Also, I am somehow assured to some extent that the data is going to stay on my phone in an encrypted format and that it may store certain recognizable features of my face and not entire images.” — Shankar.

“I use certain social media apps, even though I have heard of instances where our data in these apps has been misused. My friends and family live abroad, and these platforms make it very easy to connect with them. When they introduced more security features, I turned those settings on and that did give me some peace of mind.” — Sneha. An exception was Margaret (group 3), who has a background in marketing. She feels that public perception of data breaches is exaggerated and that companies only use data to market things to you. She says that even in case of breaches, there are usually simple ways to seek remedies (e.g., calling your bank and blocking your credit card) and it is not as scary as popular media makes it out to be.

4.6. Trust in AI

Regarding trust in AI, Group 1 participants say that having a vague understanding of how AI works can help them trust it more. Aarti and Leela say that while they do not fully understand how AI works, the output it provides seems logical, which is enough for them to trust and use it. Samuel discusses the need for us to acknowledge that AI is not flawless, and question it, rather than blindly trust it. When it comes to trusting AI to make decisions, most participants in this group preferred the ultimate decision to lie with humans. Leela says that humans can think in deeper ways than a robot can, and thus decisions in important areas should be made by humans. Group 2 participants acknowledged that it may not be possible to always understand the technical workings of AI systems which can be sensed in Annie’s comment: “Unless you are a technical expert, I don’t think you’ll ever fully know how an AI makes its decisions, unless you understand how neural networks are built or things like that. What I’d want is evidence to show that the decisions it’s making are correct. So, if there was some validation saying that this AI system has been used for the last two to three years and correctly predicted these things and has had positive impacts in these ways — that would give me confidence in it.”

Participants in Group 2 also said that they would be more confident in using an AI-based tool if subject experts were involved in the development of AI systems (e.g., doctors involved in developing AI tools for healthcare) and if there was evidence of scrutiny and vigorous testing in the development process. They also preferred human intervention for final decisions to be made in areas like healthcare, justice or finance. Group 3 participants advocate for the need to understand how AI works. Margaret says that she would want to see evidence of data quality control checks and statistical indicators of performance to trust an AI system. Shankar and Dhruv advocate for building more explainable AI (XAI) to understand and build trust in black box models. While Margaret believes that AI could have better accuracy in making healthcare decisions, Sneha and Shankar talk about how human intervention is needed to make sure that the AI systems are not enhancing biases existing in society. Dhruv points out that healthcare systems are overburdened and having human oversight to make each decision may not be possible. He believes that human oversight should be at the stage where AI is being trained to ensure that it functions without bias.

4.7. Need for awareness and regulation

When it comes to awareness and regulation, participants across all groups of AI-literacy unanimously agreed that it is essential to develop regulations and policies specific to the use of AI. Group 1 participants express concern about certain people or companies misusing AI for personal gain in ways that harm more vulnerable people. For example, Leela discusses how deepfakes have been used in problematic ways to destroy the reputation of many successful women. Participants identify regulation as the only solution to minimize the harm. Group 2 participants also agree that regulations are essential for the responsible development and use of AI and identified the importance of laws such as GDPR and the EU AI Act, but many also advocated for a globally or at least regionally agreed-upon, domain-specific framework, for the safe use of AI technology, developed in consultation with various stakeholders and experts.

Group 3 participants felt that regulation should focus on increasing awareness about AI, improving transparency and reducing misuse. They emphasized the importance of people understanding the basics of AI, where it works well and where it does not, and the fact that it is not perfect but only as good as the data it was trained on. Sneha suggests that raising awareness among the public should be a shared responsibility of the government, AI developers, teachers, and domain experts (e.g., doctors for healthcare AI tools). Shankar acknowledges that it can be difficult to operationalize guidelines sometimes, but developing the guidelines is an important first step. Dhruv emphasizes the need for regulations not to be so restrictive that they stifle innovation, especially for smaller organizations that do not have the resources to work within heavy restrictions.

5. Discussions

Many previous studies have established that a higher digital literacy is associated with more positive attitudes towards the use of computer systems [28–31]. Being well-literate in AI would not only entail being able to use AI-based applications but also being able to critically think about these systems and the various ethical issues they present [17, 32]. In this study, participants belonging to three different AI-literacy categories (limited AI knowledge, intermediate, and expert) shared their views on various aspects of artificial intelligence. Existing literature has indicated that those with limited AI-literacy are likely to subscribe to dystopian views, such as a world fully controlled by AI that leaves out humans [14, 16].

5.1. Perceptions of AI

While Dave, a member of the limited AI-literacy group, did discuss his fears of AI eventually controlling everything, most of the interviews, regardless of AI-literacy levels, demonstrated more immediate concerns. While some of the interview questions may have caused participants to focus on realistic concerns, the initial questions were left open-ended for participants to freely talk about any apprehensions about AI that came to mind. Recent developments in the field, such as ChatGPT and other platforms, have likely allowed people's perceptions to move beyond fictional narratives. With AI being much more tangible now, very dystopian views have likely become less common. However, there were some inaccuracies in how some Group 1 (limited AI-literacy) participants perceived AI. For instance, they were not aware that AI is trained by using lots of data and thus do not connect the issue of data privacy to AI and did not believe that AI can treat people unfairly, as computer programs should be objective and not biased the way humans do. These misconceptions can be unsafe and risky in a world that is increasingly adopting AI, as people may blindly trust AI systems without asking questions about important issues such as data quality or bias.

5.2. AI and creativity

AI has the potential to reshape human creativity. Cremer et al. [32] suggest three possible futures: one where there is a bloom of AI-assisted innovation, one where machines monopolize creativity or one where human-made creative works command a premium due to their authenticity. While some participants in this study expressed concern that human-generated art may lose its value with the rise of AI-generated art, there is research evidence indicating the opposite. Horton et al. [33] found that people devalue art labelled as AI-generated and tend to consider human works more creative. Zhou and Lee [34] found that the visual novelty of AI-generated artworks declines over time and tends to have a similar style. They note that artists who can produce novel works regardless of their use or non-use of AI tend to be evaluated favorably. Elfa and Dawood [35] suggest that AI can examine existing works and patterns very quickly and generate patterns, which can save a lot of cognitive effort for artists, and the work generated by AI can inspire artists to build new things. The participants in our study suggest that adapting to new ways of thinking that utilize AI to build creative works and policies that protect human artists and their artworks, might be the best way forward.

5.3. AI and job replacement

When it comes to fear of losing jobs, a previous study [36] highlighted that young people, non-regular employees and those working in clerical and manufacturing jobs perceived the highest subjective risk of being replaced in their jobs by AI, while people with higher education, especially in science and engineering, and those in jobs that require professional qualifications or licenses perceived the least subjective risk of losing their jobs. While many participants believed that administrative tasks could be performed by AI easily, the interview with an actual administrative assistant revealed that they did not fear AI taking over their jobs any time soon because a lot of their role involves coordination with multiple stakeholders, and the systems are not advanced enough to perform those tasks yet. Most of us have had at least one experience with customer service chatbots, often repeating the same information and not providing us with an actual solution. This would indicate that humans are still required to do these jobs until the time comes when AI can be more efficient at these tasks. Regardless, it is undeniable that AI is increasingly being integrated into the workplace, and we must invest in upskilling and improving our AI-literacy to become more comfortable with the use of AI. The study also pointed out that governments must strategize ways to manage the rapid progress of AI developments without economically impacting the common people. Some suggestions made by other academics [37] as well as experts in the present study include the introduction of Basic Income or Universal Credit schemes to cushion the economic impact for people in at-risk occupations, accompanied by the introduction of training programs to equip individuals with the required skills to integrate into the changing workspace.

5.4. AI and bias

Out of the five participants in this study belonging to the limited AI-literacy group, two participants said that they do not believe that AI can treat people unfairly, and one participant believed that AI should be fairer than humans since it cannot have the kind of bias we do, and only two participants believed that AI could be biased due to flaws in its design or that AI could embody the biases that its creators have. One of the reasons why people are likely to believe that AI would be unbiased is due to the lack of awareness that AI algorithms need to be trained with data, and this data can often be biased, which is aligned with multiple instances of unfairness in AI systems that have been previously

documented. For example, many facial classification systems are more accurate for light-skinned and male faces than dark-skinned female faces because the training data was predominantly made up of white male images [38, 39]. Similarly, an algorithm used to determine who should qualify for extra healthcare facilities used the cost of healthcare expenses as a predictor, which was biased against minority communities who had much higher healthcare needs but accessed healthcare services less frequently due to lower incomes [40]. In these regard, this study supports the argument that we must start educating everyone about the issue of fairness so that they are equipped with the knowledge and skills to use AI responsibly.

5.5. AI and data privacy

One of the core components of AI applications is the need for a lot of data to build an application, and there needs to be security measures and privacy preservation techniques applied at various levels of the data architecture [41]. Participants in the lower AI-literacy category did not seem to connect data to AI or demonstrate an understanding that data privacy issues are also linked with AI. Previous studies have suggested the need for an ethical trade-off to build efficient AI systems, where data privacy and access are balanced [42]. Furthermore, participants in our study also acknowledged that there is a trade-off between allowing access to personal data and the convenience we get because of it (e.g., personalized recommendations on Netflix), but emphasized the importance of companies taking steps to comply with at least the minimal industry standards for data privacy and security. We would also suggest that companies take the initiative to explain what data is being collected and how it will be used to improve the systems in a user-friendly manner, without technical jargon.

5.6. Trusting AI and accountability

When it comes to trusting AI, there are two different takes. Many of the participants in this study who belong to the higher literacy categories work in the AI field and therefore consider it very important for them to know how a system works or make decisions to use it or recommend it to their clients. On the other hand, some participants feel that it is not possible to understand the technical details of every AI system, and what is more important is evidence that the system is indeed efficient and accurate. However, Group 3 participants are optimistic about being able to better understand the workings of deep learning, highlighting more recent research into explainable AI techniques (XAI) that work to break black box models and understand the workings of these models. Nonetheless, they have also shown concern about existing XAI methods (e.g., Local Interpretable Model-Agnostic Explanations (LIME), and Layer-wise Relevance Propagation (LRP) as they tend to have more technical explanations and do not cater to the layman [5]. This highlights the need for policies that encourage makers of explainable AI to have simplified explanations that can be understood by all and are freely available to any user. This is aligned with the study of Ehsan et al. [20] where participants with an AI and non-AI background were exposed to three different types of explainable AI: A rationale-generating (RG) robot that explains the reasoning behind its decisions in natural language, an Action Declaring (AD) robot that states its decisions without providing reasoning, and a Numerical-Reasoning (NR) robot that “thinks out loud” by simply outputting the numerical Q-values for the current state with no language component.

It was found that even the group that had a non-AI background trusted the numbers generated by the NR robot more, without understanding what they meant, just because it seemed smart and calculated. While numerical metrics such as accuracy can be good indicators for the reliability of a model, they can also be unreliable in cases where training data quality is substandard, causing models to

overfit. Therefore, amateur users need to develop skills to understand what these metrics mean and critically evaluate the reliability of research that claims high accuracy. With regards to trusting AI to make decisions in healthcare, finance, and criminal justice, most participants felt that they would need a combination of both AI and human experts. Having human oversight would act as a safety net in case of poorly developed AI systems that show bias, while powerful AI systems can process data rapidly and simplify people’s jobs in these fields. Therefore, a combination would be the most responsible way to use AI at present.

5.7. Implications

This study has many strong implications for AI developers, policymakers and educators. Firstly, our findings provide valuable insights into the key areas where people have misconceptions or a lack of awareness about how AI systems work. For instance, many users are unaware of the critical role that data plays in shaping AI outputs or believe that AI systems are inherently unbiased. This highlights the need for targeted awareness campaigns aimed at correcting these misconceptions. Educators must prioritize teaching common people the skills necessary to critically evaluate AI tools, ensuring they can make informed decisions in a data-driven world. Policymakers should consider supporting initiatives that promote AI-literacy and data transparency, creating an informed public that can engage with AI systems more confidently and responsibly. Secondly, for AI developers aiming to maximize their user base, building user trust is essential. Our findings show that people are more likely to trust AI systems when they are provided with evidence of the system’s accuracy, even if they do not fully understand the underlying technical mechanisms. This highlights the importance of transparency. By openly publishing performance metrics and outcomes in layman’s terms, AI companies can foster greater trust among users. Moreover, this transparency can bridge the gap between expert and non-expert users, ensuring broader accessibility and adoption. Thirdly, the study’s findings illustrate the critical importance of privacy and cybersecurity in shaping user behavior. Users are highly sensitive to how their data is handled, and a perceived lack of robust security measures can lead to a significant loss of trust. If AI developing companies fail to invest adequately in cybersecurity, users are not only less likely to engage with their products but are also more inclined to switch to competitors that demonstrate stronger privacy protections. This highlights that maintaining high standards of cybersecurity is not just a legal requirement but a key competitive advantage in the AI marketplace. For policymakers, these findings emphasize the need for stronger regulations focused on data protection, ensuring that organizations are held accountable for safeguarding user information. By implementing policies that mandate transparency in data practices and enforce strict cybersecurity standards, governments can help build a more secure and trustworthy AI ecosystem.

6. Conclusions

This study explored a range of opinions of people of diverse AI-literacy backgrounds. Most participants saw AI as a tool that makes life very convenient and believe that AI will be increasingly integrated into various aspects of our lives in the near future. Some common concerns expressed by participants included bias due to poor data quality, people’s lack of awareness about the accuracy of AI predictions, misuse of AI tools by bad actors, and managing the economic impact on individuals whose jobs are at risk due to AI automation. The study found that participants with lower AI-literacy levels were unaware of the link between data and AI. The lack of understanding in how AI is built also led them to think that AI cannot be biased, as they think of bias as a quality only humans exhibit. This study makes a unique contribution to existing literature as it is one of the first studies to qualitatively examine

ethical considerations with AI in a diverse AI-literacy population. It also has many practical implications for AI developers and educators.

6.1. Limitations and directions for future research

A key limitation of the present study is the use of a non-probability sampling method, which can sometimes lead to a less representative sample. For example, all participants classified as having limited AI-literacy in this study still possessed at least basic digital literacy (i.e., comfortable with using computers for basic functions like web browsing, email and others), which may skew the results toward those who are somewhat familiar with technology. This creates a potential bias, as the findings would have likely indicated a much larger gap between expert and amateur users in understanding and adopting the use of AI if we had included participants who also had lower levels of digital literacy or less exposure to technology. To address these limitations in future research, it is crucial to include more diverse samples with individuals of varying levels of digital literacy. Cultural factors can also influence attitudes about AI. For instance, Pew Research Centre's 2020 survey [43] noted Eastern countries like Japan showing more positive attitudes towards AI than Western countries like the US. Therefore, future research with international and cross-cultural samples could offer valuable insights into how perceptions of AI ethics, trust, and transparency vary across different societies, which may also inform region-specific policy development. While the present study included a variety of industry and academic experts, there were no policymakers in the sample. Including participants directly involved in shaping AI governance in future research would provide valuable insight into the practical challenges of establishing regulations. Their perspectives would complement those of the general public, offering a fuller understanding of how ethical principles are translated into policy and how trust in AI can be fostered. In terms of methodology, if larger qualitative datasets are collected in future studies, machine learning-based sentiment analysis models could be explored as a means of analyzing these responses. While such models can efficiently detect patterns across large datasets, their limitations, particularly in handling linguistic nuance and capturing the full depth of human emotion, must be acknowledged and carefully addressed.

Ethical Statement

This study did not require formal ethical approval because the School of Business, Computing, and Social Science at the University of Gloucestershire does not issue IRB-style approval numbers for Master's-level human-subjects research. The project was reviewed and internally approved by the school-level ethics team in accordance with the University of Gloucestershire's ethical policy, but no formal ethical approval reference number is provided for Master's dissertations.

Conflicts of Interest

The authors declare that they have no conflicts of interest to this work.

Data Availability Statement

Data available on request from the corresponding author upon reasonable request.

Author Contribution Statement

Zeba Khadhijah: Methodology, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Writing – review & editing, Visualisation, Project administration.
Bhupesh Kumar Mishra: Conceptualisation, Methodology, Software,

Validation, Resources, Writing – review & editing, Visualisation, Supervision.

References

- [1] Sheikh, H., Prins, C., & Schrijvers, E. (2023). Artificial intelligence: definition and background. In H. Sheikh, C. Prins, and E. Schrijvers (Eds.), *Mission AI: The new system technology* (pp. 15–41). Switzerland: Springer. https://doi.org/10.1007/978-3-031-21448-6_2
- [2] Chubb, J., Reed, D., & Cowling, P. (2024). Expert views about missing AI narratives: Is there an AI story crisis?. *AI & Society*, 39(3), 1107–1126. <https://doi.org/10.1007/s00146-022-01548-2>
- [3] Wang, R., Harper, F. M., & Zhu, H. (2020). Factors influencing perceived fairness in algorithmic decision-making: Algorithm outcomes, development procedures, and individual differences. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–14. <https://doi.org/10.1145/3313831.3376813>
- [4] Parycek, P., Schmid, V., & Novak, A. S. (2024). Artificial intelligence (AI) and automation in administrative procedures: Potentials, limitations, and framework conditions. *Journal of the Knowledge Economy*, 15(2), 8390–8415. <https://doi.org/10.1007/s13132-023-01433-3>
- [5] Ras, G., Van Gerven, M., & Haselager, P. (2018). Explanation methods in deep learning: Users, values, concerns and challenges. In *Explainable and interpretable models in computer vision and machine learning*. Switzerland: Springer International Publishing.
- [6] Yeh, S. C., Wu, A. W., Yu, H. C., Wu, H. C., Kuo, Y. P., & Chen, P. X. (2021). Public perception of artificial intelligence and its connections to the sustainable development goals. *Sustainability*, 13(16), 9165. <https://doi.org/10.3390/su13169165>
- [7] Lai, M. C., Brian, M., & Mamzer, M. F. (2020). Perceptions of artificial intelligence in healthcare: Findings from a qualitative survey study among actors in France. *Journal of Translational Medicine*, 18(1), 14. <https://doi.org/10.1186/S12967-019-02204-Y>
- [8] Nash, D. M., Thorpe, C., Brown, J. B., Kueper, J. K., Rayner, J., Lizotte, D. J., & Zwarenstein, M. (2023). Perceptions of artificial intelligence use in primary care: A qualitative study with providers and staff of Ontario Community Health Centres. *The Journal of the American Board of Family Medicine*, 36(2), 221–228. <https://doi.org/10.3122/jabfm.2022.220177R2>
- [9] Acampa, S. (2025). Imagining AI: Exploring narratives and sociotechnical imaginaries between utopias and dystopias. *SN Social Sciences*, 5(7), 90. <https://doi.org/10.1007/s43545-025-01130-9>
- [10] Cave, S., Coughlan, K., & Dihal, K. (2019). “Scary robots” examining public responses to AI. In *2019 AAAI/ACM Conference on AI, Ethics, and Society*, 331–337. <https://doi.org/10.1145/3306618.3314232>
- [11] Ouchchy, L., Coin, A., & Dubljević, V. (2020). AI in the headlines: The portrayal of the ethical issues of artificial intelligence in the media. *AI & Society*, 35(4), 927–936. <https://doi.org/10.1007/s00146-020-00965-5>
- [12] Milmo, D. (2023). ChatGPT reaches 100 million users two months after launch. Retrieved from: <https://www.theguardian.com/technology/2023/feb/02/chatgpt-100-million-users-open-ai-fastest-growing-app>
- [13] Li, J., & Huang, J. S. (2020). Dimensions of artificial intelligence anxiety based on the integrated fear acquisition theory. *Technology in Society*, 63, 101410. <https://doi.org/10.1016/j.techsoc.2020.101410>

- [14] Johnson, D. G., & Verdicchio, M. (2017). AI anxiety. *Journal of the Association for Information Science and Technology*, 68(9), 2267–2270. <https://doi.org/10.1002/asi.23867>
- [15] Wang, Y. Y., & Wang, Y. S. (2022). Development and validation of an artificial intelligence anxiety scale: An initial application in predicting motivated learning behavior. *Interactive Learning Environments*, 30(4), 619–634. <https://doi.org/10.1080/10494820.2019.1674887>
- [16] Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. In *2020 CHI Conference on Human Factors in Computing Systems*, 1–16. <https://doi.org/10.1145/3313831.3376727>
- [17] Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence*, 2, 100041. <https://doi.org/10.1016/j.caeai.2021.100041>
- [18] Sartori, L., & Bocca, G. (2023). Minding the gap (s): public perceptions of AI and socio-technical imaginaries. *AI & Society*, 38(2), 443–458. <https://doi.org/10.1007/s00146-022-01422-1>
- [19] Nourani, M., Roy, C., Block, J. E., Honeycutt, D. R., Rahman, T., Ragan, E. D., & Gogate, V. (2022). On the importance of user backgrounds and impressions: Lessons learned from interactive AI applications. *ACM Transactions on Interactive Intelligent Systems*, 12(4), 1–29. <https://doi.org/10.1145/3531066>
- [20] Ehsan, U., Passi, S., Liao, Q. V., Chan, L., Lee, I. H., Muller, M., & Riedl, M. O. (2024). The who in XAI: How AI background shapes perceptions of AI explanations. In *2024 CHI Conference on Human Factors in Computing Systems*, 1–32. <https://doi.org/10.1145/3613904.3642474>
- [21] Ritchie, J., Ormston, R., McNaughton Nicholls, C., & Lewis, J. (2013). *Qualitative research practice: A guide for social science students and researchers*. USA: Sage Publications.
- [22] Mulisa, F. (2022). When does a researcher choose a quantitative, qualitative, or mixed research approach?. *Interchange*, 53(1), 113–131. <https://doi.org/10.1007/s10780-021-09447-z>
- [23] Mishra, B. K. (2025). Unpublished internal document [Unpublished raw data]. University of Gloucestershire.
- [24] Jamieson, M. K., Govaert, G. H., & Pownall, M. (2023). Reflexivity in quantitative research: A rationale and beginner’s guide. *Social and Personality Psychology Compass*, 17(4), e12735. <https://doi.org/10.1111/spc3.12735>
- [25] Dovi, C. (2011). *Introducing qualitative research in psychology*. UK: McGraw-Hill Education.
- [26] Palaganas, E. C., Sanchez, M. C., Molintas, M. P., & Caricativo, R. D. (2017). Reflexivity in qualitative research: A journey of learning. *The Qualitative Report*, 22(2), 426–438. <https://doi.org/10.46743/2160-3715/2017.2552>
- [27] Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- [28] Saklaki, A., & Gardikiotis, A. (2024). Exploring Greek students’ attitudes toward artificial intelligence: Relationships with AI ethics, media, and digital literacy. *Societies*, 14(12), 248. <https://doi.org/10.3390/soc14120248>
- [29] Abima, B., Engotoit, B., Kituyi, G. M., Kyeyune, R., & Koyola, M. (2021). Relevant local content, social influence, digital literacy, and attitude toward the use of digital technologies by women in Uganda. *Gender, Technology and Development*, 25(1), 87–111. <https://doi.org/10.1080/09718524.2020.1830337>
- [30] Nikou, S., De Reuver, M., & Mahboob Kanafi, M. (2022). Workplace literacy skills—How information and digital literacy affect adoption of digital technology. *Journal of Documentation*, 78(7), 371–391. <https://doi.org/10.1108/JD-12-2021-0241>
- [31] Omrani, N., Riviuccio, G., Fiore, U., Schiavone, F., & Agreda, S. G. (2022). To trust or not to trust? An assessment of trust in AI-based systems: Concerns, ethics and contexts. *Technological Forecasting and Social Change*, 181, 121763. <https://doi.org/10.1016/j.techfore.2022.121763>
- [32] Amankwah-Amoah, J., Abdalla, S., Mogaji, E., Elbanna, A., & Dwivedi, Y. K. (2024). The impending disruption of creative industries by generative AI: Opportunities, challenges, and research agenda. *International Journal of Information Management*, 79, 102759. <https://doi.org/10.1016/j.ijinfomgt.2024.102759>
- [33] Horton Jr, C. B., White, M. W., & Iyengar, S. S. (2023). Bias against AI art can enhance perceptions of human creativity. *Scientific Reports*, 13(1), 19001. <https://doi.org/10.1038/s41598-023-45202-3>
- [34] Zhou, E., & Lee, D. (2024). Generative artificial intelligence, human creativity, and art. *PNAS Nexus*, 3(3), 52. <https://doi.org/10.1093/pnasnexus/pgae052>
- [35] Ali Elfa, M. A., & Dawood, M. E. T. (2023). Using artificial intelligence for enhancing human creativity. *Journal of Art, Design and Music*, 2(2), 3. <https://doi.org/10.55554/2785-9649.1017>
- [36] Prokopowicz, K. (2022). Technological unemployment: Who is afraid of losing their job to a robot?. *Journal of Business Management*, 20. <https://doi.org/10.32025/JBM22001>
- [37] McGaughey, E. (2022). Will robots automate your job away? Full employment, basic income and economic democracy. *Industrial Law Journal*, 51(3), 511–559. <https://doi.org/10.1093/indlaw/dwab010>
- [38] Yucer, S., Tektas, F., Al Moubayed, N., & Breckon, T. (2024). Racial bias within face recognition: A survey. *ACM Computing Surveys*, 57(4), 1–39. <https://doi.org/10.1145/3705295>
- [39] Kotwal, K., & Marcel, S. (2025). Review of demographic fairness in face recognition. *IEEE Transactions on Biometrics, Behavior, and Identity Science*. <https://doi.org/10.1109/TBIOM.2025.3601217>
- [40] Vartan, S. (2019). Racial bias found in a major health care risk algorithm. Retrieved from: <https://www.scientificamerican.com/article/racial-bias-found-in-a-major-health-care-risk-algorithm/>
- [41] Dai, D., & Boroomand, S. (2022). A review of artificial intelligence to enhance the security of big data systems: State-of-art, methodologies, applications, and challenges. *Archives of Computational Methods in Engineering*, 29(2), 1291–1309. <https://doi.org/10.1007/s11831-021-09628-0>
- [42] Bak, M., Madai, V. I., Fritzsche, M. C., Mayrhofer, M. T., & McLennan, S. (2022). You can’t have AI both ways: Balancing health data privacy and access fairly. *Frontiers in Genetics*, 13, 929453. <https://doi.org/10.3389/fgene.2022.929453>
- [43] Kapur, S., (2021). Reducing racial bias in AI models for clinical use requires a top-down intervention. *Nature Machine Intelligence*, 3(6), pp.460–460. <https://doi.org/10.1038/s42256-021-00362-7>

How to Cite: Khadhijah, Z., & Mishra, B. K. (2026). Perceptions of Potentials and Risks of Artificial Intelligence Across Diverse Levels of AI-literacy. *Artificial Intelligence and Applications*. <https://doi.org/10.47852/bonviewAIA62025489>