Through the Lens of Humanity: Unveiling Perspectives on AI Bias

A qualitative analysis of how (diverse) individuals perceive and interpret bias within artificial intelligence (AI) systems.

Student Name: Hrishita Pramanik
Student Number: 558475

Supervisor: Dr Aviv Barnoy

Media & Business
Erasmus School of History, Culture and Communication
Erasmus University Rotterdam

Master's Thesis
June 2024
Word count: 20225

Through the Lens of Humanity: Unveiling Perspectives on AI Bias

A qualitative analysis of how (diverse) individuals perceive and interpret bias within artificial intelligence (AI) systems.

Abstract

This thesis paper examines what diverse individuals perceive as bias in the AI technologies and how they understand and respond to the embedded biases of man in AI algorithms. The focus of this paper is on perceptions of AI bias, rather than the than the actual biases present in the algorithms themselves or the actions that may result from these perceptions. A qualitative design is used to present a detailed description and in-depth understanding of how biases are perceived and experienced among different ages and nationalities. It will combine semi-structured interviews with participants from varied backgrounds to provide a nuanced view of how biases are perceived and experienced across different ages and nationalities. Significant are the findings about what will be aware and concerned at different levels of augmentation and mitigation by AI about social and ethical challenges. Indeed, this has critical importance in developing transparent, fair, and accountable AI and will be relevant in the more general discussion related to gaining public trust with the deployment of such technologies in an ethical manner. This paper presents challenges that AI has brought on public trust and ethical consideration; hence, it will enhance the visibility of research within the academic audiences and practitioners in fields concerned with AI ethics, technology policy, and the sociotechnical consequences of artificial intelligence.

KEYWORDS: Artificial Intelligence, AI Bias, Perceptions of Technology, Ethical AI, Sociotechnical Systems, Qualitative Research, Thematic Analysis, Technology Policy, User Experience, Algorithmic Accountability

Preface & Acknowledgements

I remember dreading writing another thesis after I just wrote my bachelor's thesis, I thought it would be impossible to write another thesis right after, at a higher academic level and a longer read. I thought it was going to be impossible till I ended up doing it again. On that note, I want to give my first acknowledgement to my parents, for continuously doing everything in their power to make me a better and well-rounded person. For helping me see the world and all the opportunities the world has to offer to me. For always putting their kids first. For almost every good thing in my life is because of them. Without them, I would not be here, and because my parents believed in me, I believe it gave me the strength to continue to make them proud, Mama and Papa Pramanik, I hope I made you proud again.

To all my participants who have given time out of their day to help me with my research, I am deeply thankful for your cooperation and kindness, without each of you, this research could not have been conducted and your insights were invaluable. I would also like to thank my supervisor Aviv Barnoy for helping during this challenging process to make the best version of this thesis. For both, I owe my deepest sincerity.

I also want to thank Carlo Meijers, for giving me sound advice and always being there for me emotionally, as this was a transformative and challenging process, having you by my side made it easier.

I almost forgot to add this part, but I thank myself for making it despite all the differences I have faced in my personal life. Moving all the way from Kolkata, India to where I am right now, I have lived in places like Hong Kong, China and Hamburg, Germany and now Rotterdam, The Netherlands and moving constantly was difficult for child me. I had to restart my life a few times, find new friends and find myself every time I moved. Having my little sister Nikita with me did make it easier. She inspires me to be strong. Though I learnt to adapt and learn new languages, I always felt that home was never a place for me, but rather the people present that made me feel like I was at home, and to those people, thank you for always sticking around for me, thank you for being a real friend.

To my grandparents, thank you for believing in me and may you always be proud of me.

Lastly, *bless all forms of intelligence.*

Table of Contents

1. Introduction

The introduction of new technologies has often been met with a mix of awe, skepticism, and critical inquiry. From the invention of the radio and television to the broad use of the internet, every technological advancement in history has been met with conversations about its potential for social change and implications (Wolff, 2021, p. 1). These conversations have frequently interchanged between utopian visions of democratization and empowerment, and dystopian fears of control and loss of autonomy. Artificial intelligence (AI), which embodies the pinnacle of technological advancement and simultaneously reflects our most profound moral and ethical concerns, is currently at the forefront of such discourses (Tai, 2020, p. 339).

The term 'AI bias' describes the systematic and potentially unfair results that might occur when AI systems are developed by ingrained or learned prejudices (Suresh & Guttag, 2021, p. 3). This bias typically originates from various sources, including the design and algorithms that may unintentionally favor particular outcomes, or the data used to train AI systems, which may contain historical human biases or imbalances in representation (Suresh & Guttag, 2021, p. 3). The term "AI bias" can refer to many problems, including disparities in facial recognition technology and biased decision-making in AI-based hiring tools (Suresh & Guttag, 2021, p. 3).

The societal relevance of studying citizens' perceptions of AI bias cannot be overstated as AI technologies are becoming more integrated and embedded into everyday life. From mundane daily interactions with digital assistants to important decision-making in the fields of healthcare, finance, and law enforcement, AI's impact goes beyond its technological utility and has important societal and ethical implications (Kraus et al., 2021, p. 1).

Although people's perceptions of these biases are the focus of this research, it's crucial to distinguish these perceptions from the real behavioral acts or policy decisions that are influenced by AI. Perceptions can inform stakeholders' approaches to regulating and developing AI, thereby shaping the framework for ethical AI deployment (Kraus et al., 2021, p. 1). Nevertheless, this impact on practice and policy is a different line of analysis, where policymakers' and technologists' perceptions of bias must be interpreted separately from their actual actions or choices. This paper emphasizes the necessity to distinguish between the societal impact of perceived prejudices and the specific steps made to address them.

1.1 Research Design

This research addresses AI bias within a socio-technical framework that distinguishes between human perceptions and the actions that result from these perceptions. Specifically, the chosen methodology for this paper is designed to evaluate not only how participants perceive AI biases, where each perception can be driven by individual experiences, cultural backgrounds, and societal interactions, but also how such perceptions have led to subsequent behaviors or decisions concerning AI technologies.

The primary research question is: “**How do different individuals perceive bias in AI?**”. This research is also complemented by four sub-questions which will help further answer the main research question; “How do concerns about data integrity and algorithmic transparency influence perceptions of fairness and trust in AI systems”, “What are the common patterns of AI utilization among participants in their personal and professional activities?”, “What ethical concerns do participants identify in the development and application of AI technologies?” and lastly, “In what ways do the cultural backgrounds and personal values of participants shape their engagement with AI technologies?”.

Based on previous literature, this study examines individuals' subjective experiences and interpretations of AI bias by analyzing their accounts, considering factors such as personal technology experiences, cultural background, technological literacy, exposure to diverse viewpoints, and trust in AI systems (Jobin et al., 2019, pp. 2-4).

This research is further enhanced by the debates on technology and objectivity. The argument focuses on whether technology has ingrained moral values and biases or if it is fundamentally neutral, like a bench, simply reflecting the intentions and values of its users. The debate at hand is vital to comprehending artificial intelligence (AI). While these systems are frequently presented as unbiased and objective, the reality is that they are designed and developed within complex human contexts, carrying the potential to embody human prejudices and societal biases (Wolff, 2021, p. 1).

Artificial intelligence (AI) has progressed from a theoretical concept to a ubiquitous technology, with significant innovations along the way. AI can be traced back to the mid-twentieth century, when the first ideas about machine intelligence were discussed (Kaplan & Haenlein, 2019, p. 15). Alan Turing developed the Turing Test in the 1950s, laying the groundwork for understanding and developing intelligent machines (Kaplan & Haenlein, 2019, p. 16). Alan Turing outlined the foundational ideas of the future fields of computer science and artificial intelligence (Bostrom, 2016, p. 29). The first artificial intelligence (AI) programs, which were mostly task-specific algorithms, were made possible in the 1950s thanks to the theoretical foundation that Turing established. This period saw a shift from theoretical exploration to practical application, laying the groundwork for the subsequent digital revolution.

Deep learning and neural networks were introduced in the late 20th and early 21st centuries, and this led to a significant advancement in AI capabilities. These advancements made it possible for machines to reach par with or occasionally surpass human performance on difficult tasks like speech and image recognition (Bostrom, 2016, pp. 58-60). The application of AI in a variety of industries, including healthcare (where it aids in diagnostic procedures) and finance (through algorithmic trading), has demonstrated the technology's usefulness as well as its transformative potential (Bostrom, 2016, p. 102).

These developments have had a significant impact on society and sparked discussions about the ethical implications of AI, ranging from privacy issues to the possibility of job displacement as AI systems take on tasks that have historically been performed by humans (Bostrom, 2016, p. 115). The ongoing development of artificial intelligence (AI) offers prospects for increased productivity as well as control and fair benefit distribution challenges.

A major ethical problem with the application of AI technologies is bias in the technology. Recent studies by Pagano et al. (2023) claim that AI bias happens when an algorithm produces results that are consistently biased, frequently because of poor algorithmic design or data input. These biases appear in a variety of applications, such as financial services and facial recognition technologies, where they may reinforce racial and socioeconomic biases already present in society (Pagano et al., 2023, p. 23).

“Systematic prejudice” describes deep-rooted prejudices in AI systems, which frequently result from algorithmic errors or biased training datasets. Strong algorithmic accountability and transparency are vital in technology development because these deeply rooted biases are often undetectable until major harm is observed (Pagano et al., 2023, p. 24).

In order to reduce the gap between those who have access to advanced technologies and those who do not, the “democratization” of AI entails opening up these technologies to a wider audience. The topic of democratization is covered by Gelles et al. (2024, pp. 19840-19848) concerning resource allocation in AI research. They stress the significance of fair access to computing resources, which is essential for promoting diversity and innovation in the industry. The authors contend that although basic physical resources like processing power are important, real democratization also necessitates expanding access to knowledge and data. This promotes a more inclusive development of AI technologies by ensuring that the advantages and decision-making capabilities of AI are not limited to a small group of people but rather are distributed across different demographic groups and geographical areas (Gelles et al., 2024, pp. 19840-19848).

1.2 Academic and Social Relevance

Artificial intelligence (AI) drastically alters several industries, demonstrating its broad applications and possible drawbacks. By evaluating complicated medical data, assisting clinicians, and boosting population health management, artificial intelligence (AI) in healthcare enhances patient outcomes and optimizes delivery (Alowais et al., 2023, p. 4). AI is essential to the development of autonomous vehicles, which will drastically change market dynamics. It also improves operational efficiency, increases equipment availability, lowers maintenance costs, and more in the automotive sector (Javaid et al., 2021, p. 86). Whilst AI influences content filtering and recommendation systems in social media, it also affects consumer behavior and information dissemination in the finance sector through algorithmic trading and risk management (Bogojevic Arsic, 2021, pp. 28–30; Hao & Chen, 2024, p. 1).

A wide range of industries are covered, and many implications of AI technologies are discussed in the vast and eclectic literature on AI bias. Research frequently examines the moral issues and difficulties raised by artificial intelligence (AI), including the potential for biases resulting from faulty algorithmic designs or skewed training data. For example, studies show that AI applications such as facial recognition and hiring algorithms run the risk of racial and gender biases (Parra et al., 2021, p. 1). The literature uses a variety of methods, from theoretical analyses exploring broader socio-ethical implications to empirical studies evaluating the impact of AI in particular applications. This corpus of work offers a crucial context for comprehending how AI technologies can reinforce current societal injustices and guides how to lessen these biases through improved regulation and design. More literature will be further discussed in the theoretical framework.

The extensive research on AI bias frequently overlooks the nuanced interplay between technological design and user interaction, especially in diverse cultural settings. Majority of existing research concentrates on technical solutions for bias, ignoring how these solutions interact with diverse users from various geographical areas (West et al., 2019, p. 18). To close this gap, this research investigates how diverse individuals perceive bias in AI and explores if cultural variations affect people's perceptions of AI bias in technology use. This approach is essential because it tackles the methodological flaw of presuming a consistent user base, which can result in mistakes when creating AI systems that are universally fair (Thilo Hagendorff & Fabi, 2022, p. 8).

In addition to increasing the technology's fairness, addressing AI bias is essential for its ethical integration into daily life and wider societal acceptance. This research ensures that AI systems serve diverse populations fairly by addressing the cultural aspects of AI bias and advancing a more inclusive understanding of AI systems (Roshanaei, 2024, p. 12). Furthermore, policy decisions based on the findings may result in regulations that support accountability and transparency in the development of AI (Olatunji Akinrinola et al., 2024, p. 51). This research is significant from a scientific standpoint because it has the potential to extend and question existing theoretical frameworks on AI ethics, going beyond technical solutions to consider more complex socio-cultural dynamics (Hagendorff, 2020, p. 8).

Given that subjective experiences and nuanced understanding are essential when examining complex issues like AI bias, semi-structured interviews, or qualitative research, are a particularly effective method. A balanced approach is provided by semi-structured interviews, which permit a thorough examination of individual experiences while offering enough structure to methodically address specific research questions. This approach works particularly well for recording participants' in-depth thoughts and emotions, which more rigorous quantitative approaches might overlook. The flexibility that semi-structured interviews offer also makes it easier to explore emergent themes that the researchers had not initially anticipated, adding to the breadth and depth of the data (Braun et al., 2019, pp. 849-850).

Furthermore, this qualitative method enables researchers to modify their questions in response to the discourse, guaranteeing that the discussion stays thorough and pertinent. Semi-structured interviews are a potent tool for obtaining rich, qualitative data that are frequently limited by more structured methodologies because of their adaptable nature. Participants are encouraged to express their opinions more freely in these dynamic and interactive interviews, which offer deeper insights into the personal and cultural contexts that shape their perceptions of AI bias (Kallio et al., 2016, p. 2659).

This thesis is organized into six carefully thought-out chapters that gradually answer the research question. By describing the prevalence and consequences of AI bias, the "Introduction" establishes the context and the applicability of the research question. Chapter two's "Theoretical Framework" summarizes previous research and draws attention to the gaps that this study fills. Chapter three describes the "Methodology" in detail and argues that semi-structured interviews were a suitable choice given the objectives of the study. The 'Results' of the interviews are presented in the fourth chapter, "Discussion" is presented in the fifth chapter which reflects on the results and explores how cultural and individual factors affect how AI bias is perceived. Finally, the sixth and last chapter, the "Conclusion" synthesizes these findings with the literature reviewed and makes recommendations for future research. This narrative approach highlights how each chapter adds to a thorough understanding of AI bias while also guiding the reader through the logical progression of the research.

The methods for detecting and reducing biases are a topic of intense discussion that is currently influencing research on AI bias. Academics are especially worried about the moral ramifications of artificial intelligence (AI) systems that mimic social injustices. These discussions are important because they have a direct bearing on the creation of more equitable AI technologies and the recommendations made for policy (Dignum, 2019, pp. 101-102).

To gain a deeper understanding of the societal effects of AI biases, this research expands upon the theoretical and practical implications by integrating and synthesizing perspectives from various disciplines and looks at how culture may play a role in awareness or perception of bias in AI and adds to the theoretical framework by examining the complex interactions that exist between societal biases and AI technologies. This research explores existing models critically and suggests new frameworks based on participant experiences for the development of ethical AI. Empirical insights are provided into how AI systems may contribute to or mitigate social inequality of awareness and understanding of AI and its biases. Practical implications of the findings are anticipated, with an emphasis on developing algorithms that are both equitable and efficient in order to impact the design of AI systems. Policymaking will be significantly impacted by this, especially in areas like healthcare, criminal justice, and employment where AI has a large influence. Dignum (2019) notes that these contributions are meant to promote an inclusive approach to AI development, guaranteeing that AI technologies meet the various needs of society while addressing potential biases that may surface during their use and implementation (pp. 59-60).

The results of the research support the creation of legislative frameworks that require accountability and transparency in the application of AI. This research also promotes the creation of norms and regulations that stop discrimination and improve justice in AI applications by influencing policy. To address biases, some of the recommendations are, for example, routinely auditing AI systems and putting corrective mechanisms in place when they are found. This could be extremely important in industries like healthcare, where AI-driven decisions can have profound effects on people's lives.

Furthermore, the insights gained from this research can be used to support curriculums in schools that emphasize ethics and AI. The findings contribute to the development of a new generation of technologists who prioritize ethical considerations in AI development by informing educational strategies. The long-term cultural changes within tech companies and the larger tech ecosystem, which promote a more moral and inclusive approach to technology development, depend heavily on the educational impact. These contributions seek to shape a more equitable digital future by advancing academic discourse and bringing about real changes in technology development and deployment practices.

2. Theoretical Framework

2.1 Bias and Fairness in Artificial Intelligence

The core of this research is cognitive bias theory, which explains how inherent cognitive biases influence perceptions and interactions with AI technologies (Tversky & Kahneman, 1974, p. 1124). Recent research, particularly on ChatGPT and other modern AI tools (Ferrara, 2024, p. 5; Jones-Jang & Park, 2023, pp. 1-2; Ray, 2023, pp. 1-2), has shed light on the complexities of AI bias and its implications. For example, while Ferrara (2024, p. 5) focuses on the butterfly effect in AI systems, Jones-Jang and Park (2023, p. 1) investigate how people react to AI failure, revealing automation bias and algorithmic avoidance. Ray (2023, p. 1) also provides a comprehensive review of ChatGPT, covering its history, applications, challenges, and ethical considerations. However, critical analysis is required to compare these findings and draw broad conclusions (Hosseini & Horbach, 2023, p. 2; Rozado, 2023, p. 1). Hosseini and Horbach (2023, p. 3) argue for considerations and recommendations for using ChatGPT, while Rozado (2023, p. 4) investigates the political biases of artificial intelligence. It is critical to define the research focus, separating biases in AI perception from biases inherent in AI systems (Motoki et al., 2023, p. 3). This distinction promotes a more nuanced understanding of AI bias and its societal implications.

Cognitive bias theory has a significant impact on how people interpret and react to AI technologies. Cognitive biases can lead to inaccurate perceptions and decisions (Ganel et al., 2022, p. 1; Hagendorff & Fabi, 2023, p. 3). These biases may influence how people perceive AI systems and their capabilities, potentially leading to anthropomorphism, in which AI is given human-like characteristics, motivations, or feelings (Jones-Jang & Park, 2023, p. 2). This anthropomorphic perception can influence user interactions and expectations, ranging from unfounded trust to excessive fear, and is heavily influenced by individuals' prior technological experiences (Jones-Jang & Park, 2023, p. 1). Furthermore, it is critical to define anthropomorphism at the outset in order to avoid misconceptions about the nature of AI (Li & Suh, 2022, p. 1). Understanding these cognitive biases is crucial for developers, legislators, and educators to successfully navigate the integration of AI into society.

Citizens aware of AI's limitations, for example, may be cautiously optimistic, whereas others who are not may either overestimate AI's potential or completely mistrust it. To avoid misconceptions regarding the nature of artificial intelligence (AI) and to guarantee that discussions about it are based on a precise and comprehensive understanding of the wider implications of anthropomorphism, this term must be clarified as soon as it is introduced (Li & Suh, 2022, p. 1). The findings emphasize how critical it is to comprehend cognitive biases in the context of artificial intelligence. They imply that technological advancement and logical assessment may not be the only factors influencing how the public views artificial intelligence (AI) and, in turn, how AI technologies are accepted and used. Rather, it is also heavily impacted by more profound psychological processes that mold how people interact with technology. For developers, legislators, and educators to successfully traverse the challenging terrain of AI integration into society, this knowledge is essential.

The presence of biases in AI systems has been made clear by empirical evidence, demonstrating how these biases manifest in different domains and have significant societal and ethical ramifications. For example, racial biases in AI algorithms have been discovered in the field of criminal justice, which can result in unfair outcomes when it comes to parole and sentencing decisions. This field has benefited greatly from Kleinberg et al.'s (2016, pp. 1-2) examination of the trade-offs in risk assessment algorithms and their implications for criminal justice fairness. Their findings highlight how algorithmic choices can maintain structural racial inequities.

Gender biases in AI recruitment have sparked concerns about the objectivity and fairness of automated hiring procedures in the workplace. According to a study by Dastin (2022, pp. 296-299), gender biases in AI can result in biased hiring practices and potentially disadvantage qualified applicants based on their gender.

Furthermore, Obermeyer et al. (2019, p. 1), who examined the racial bias in medical diagnostic AI systems, demonstrated how AI bias affects the healthcare industry. According to their research, certain racial groups may experience delayed or inaccurate diagnoses, which could have an impact on quality of life and health outcomes and exacerbate already-existing healthcare disparities.

The wider ethical and societal ramifications of AI biases are also covered by Crawford and Schultz (2014, pp. 93-94), especially considering big data and privacy issues. Their findings highlight the importance of a comprehensive framework for addressing algorithmic bias-related predictive privacy risks. These studies together highlight the broader ethical and societal implications of AI biases, emphasizing the importance of ensuring fair and just applications across multiple industries.

It is imperative to take a dual approach to addressing AI bias, focusing on technological solutions and regulatory frameworks. Advances in machine learning in recent times present encouraging methods for reducing biases. For instance, to guarantee more equitable model results, researchers have created techniques like re-sampling and re-weighting training datasets (Hajian et al., 2016, p. 102). In order to correct for imbalances that might result in biased decisions, these techniques modify the data input. Furthermore, during the model training stage, fairness constraints added to the learning algorithms aid in directly enforcing equity (Friedler et al., 2021, p. 209).

Establishing AI ethics guidelines and putting in place thorough audit systems are critical from a regulatory perspective. According to Jobin et al. (2019, p. 389), these frameworks are designed to guarantee that AI systems follow moral guidelines and that any biases are found and fixed before being put into use. For instance, for AI systems to comply with AI governance standards, the European Union has proposed regulations requiring them to go through stringent bias assessment procedures (Cannarsa, 2021, p. 284).

Furthermore, another crucial aspect of reducing AI bias is transparency. According to Gunning et al. (2019, p. 75), developers are urged to incorporate explainable AI (XAI) systems, which offer users and regulators comprehensible explanations of the decision-making process. To foster trust and enable efficient auditing of AI systems, this transparency is essential.

Additionally, including a variety of stakeholders in the design and implementation of AI systems guarantees that a broad range of viewpoints are taken into account, lowering the possibility of missing any biases (Holstein et al., 2019, p. 117). This involves members of the impacted communities in addition to AI developers and ethicists.

Reducing bias in AI calls for a comprehensive strategy that blends advanced machine learning methods with strong legal frameworks. When combined, these tactics help create AI systems that are not only technologically advanced in terms of technology but also equitable and socially conscious.

Through real-world case studies, the exploration of biases in artificial intelligence (AI) becomes tangibly profound. These examples not only show the direct consequences of AI biases but also provide important new perspectives on how technology and social norms interact.

One real-life case is the use of healthcare algorithms, which have historically disadvantaged African American patients by predicting their healthcare needs based on cost rather than actual health needs. An algorithm that was widely used in the US health system significantly underestimated the health needs of Black patients compared to White patients, according to an influential study by Obermeyer et al. (2019). This was because the algorithm ignored the socioeconomic factors that affect access to care and equated lower healthcare spending with lower healthcare needs. Even though Black patients' levels of illness were comparable to those of other racial groups, this misalignment led to a decrease in the number of Black patients being referred to programs intended for patients with complex medical needs (Obermeyer et al., 2019, p. 447).

Gender bias was found in AI hiring tools in a case study published in the ‘Women's World Banking’ report on algorithmic bias. These biases resulted from algorithms that were trained on historical data and unintentionally favored male applicants. The research underlined the ongoing difficulty of guaranteeing fairness in AI applications by emphasizing the need to update training datasets to incorporate varied viewpoints and test algorithms for bias (Kelly & Mirpourian, 2021, pp. 17–18).

Furthermore, there are significant ethical concerns surrounding the use of AI in law enforcement, especially concerning predictive policing. According to studies, algorithms that forecast criminal activity are frequently trained on skewed police arrest data, which causes minority communities to be disproportionately targeted. This has made tensions between law enforcement and the communities they serve worse and created a vicious cycle of mistrust. To reduce these biases and rebuild confidence in public institutions, it is imperative to ensure algorithmic transparency and incorporate fairness measures (Chen et al., 2018, p. 89).

These case studies not only illustrate the real-world effects of AI biases but also emphasize the pressing need for reforms to legislation as well as advancements in technology. These real-world applications provide valuable insights that underpin the development of more equitable AI systems and contribute to ongoing conversations about the moral application of AI.

2.2 Integration of AI in Daily Life

Recent academic articles offer valuable perspectives on the ways in which diverse demographic groups engage with artificial intelligence (AI) technologies in their day-to-day lives, exposing notable distinctions according to factors such as age, profession, and educational attainment.

As Generation Z (Gen Z) was born during a time of rapid technological advancement, they are especially accustomed to using AI, and they frequently do so in educational settings, making Gen Z "digital natives" (Bhalla et al., 2021, pp. 3-7). They prefer hybrid learning strategies over traditional text-based instruction because they integrate technology and multimedia content. In line with a wider trend, digital natives anticipate that educational institutions will supply cutting-edge technological resources (Seemiller & Grace, 2016, p. 12; Seemiller & Grace, 2017, p. 58). This is reflected in their preference for technology-enhanced learning experiences. This generation exhibits a strong preference for AI-driven tools in learning, which is shaped by their familiarity with digital environments. They also have different expectations and interactions with AI.

On the other hand, older generations—Gen X and Millennials, for example—show distinct patterns of engagement with AI technologies. These groups frequently engage in learning environments with greater caution, appreciating human interaction in addition to technological advancements (Seemiller & Grace, 2017, p. 60). This distinction emphasizes the requirement for AI systems that can adjust to users' differing degrees of comfort and technological familiarity.

Further research indicates that AI is being incorporated into higher education, where institutions are using AI-enabled chatbots and automated tutors to improve student outcomes. Although its use is expanding across a variety of disciplines, a systematic review of AI in higher education finds that it is primarily employed in STEM fields (Page et al., 2021, p. 110). Using AI to personalize instruction and offer scalable, consistent support is the goal of educators; this strategy fits in well with the learning preferences of digital natives (Gough et al., 2017, p. 42).

Furthermore, people who understand AI better and use it more frequently are more likely to adopt it in both their personal and professional lives, according to the relationship between a user's technological literacy and how they interact with AI. This correlation emphasizes how crucial knowledge of AI technologies and education are as factors in adoption and acceptance (Long & Magerko, 2020, p. 34).

These results highlight the necessity for technology developers to create AI systems that are user-friendly and available to a wide range of users. To guarantee that AI technologies are widely used and accepted, they must consider the disparate technological literacies and preferences of various demographic groups. All things considered, the demographic variations in AI interaction demand careful thought to be given to the development and application of AI technologies in order to guarantee that they satisfy the various demands and expectations of users of all ages and educational levels.

Artificial intelligence (AI) has radically changed everyday choices and lives and raised serious concerns about consumer privacy. AI has greatly impacted consumer electronics. In addition to improving user convenience, the use of AI in wearables, smart homes, and personal assistants presents new issues with data security and personal space.

AI-powered smart home technologies provide unprecedented levels of automation and control over house settings, enabling improved security, energy conservation, and personalized living experiences. These systems adjust lighting, heating, and even food shopping based on user habits and preferences (Chui et al., 2023, p. 5). Convenience and privacy are traded off, as the ubiquitous nature of these gadgets prompts questions about the scope of data collecting and the possibility of spying.

AI personal assistants—such as those driven by generative AI—have advanced to handle a variety of duties, including email management and reminder scheduling. They have become a necessary component of daily life due to their capacity to comprehend and anticipate user preferences. However, because of the constant engagement, a lot of information about individual tastes, habits, and even emotional states are collected, which raises privacy concerns. For instance, research indicates that although voice assistants are convenient, there are serious security issues associated with them, such as the possibility of unauthorized data acquisition (Hoy, 2018, p. 82).

Artificial intelligence-enabled wearables evaluate health data in real time, offering insights into potential health hazards, sleep patterns, and fitness levels. Although this ongoing surveillance can encourage people to lead better lives, there are risks involved if private health information is not adequately protected or disseminated without permission. Strict security measures are necessary because of worries about unauthorized access and exploitation of the vast amounts of data these devices collect (Wolfson, 2018, para. 11).

The integration of AI into technologies for consumers undoubtedly assures increased convenience and tailored experiences. It drives product development in industries that have a direct impact on consumers' lives, such as retail, where AI is used to adapt shopping experiences and improve customer support via chatbots and personalized marketing (Moore et al., 2022, p. 5).

However, the widespread use of AI raises substantial ethical concerns, notably in terms of data privacy. Users frequently exchange large amounts of personal information for the convenience that AI provides, sometimes without fully understanding how this data might be used or exploited. There is an urgent need for clear laws to guarantee that corporations use AI properly, protecting consumer rights without limiting innovation (Li et al., 2022, pp. 28-53).

As AI evolves and becomes more deeply integrated into everyday technologies, it is critical that consumers, developers, and policymakers work together to create environments in which technology serves humanity responsibly, prioritizing privacy and security over convenience and efficiency.

2.3 Ethical and Regulatory Considerations

The global landscape of AI regulation includes a variety of methods from the EU, the United States, and China, each impacted by their unique legal cultures, technological capabilities, and political agendas, which shape the development and deployment of ethical AI systems.

The European Union (EU) has positioned itself as a pioneer in comprehensive AI legislation with the adoption of the EU AI Act in early 2023, making the EU the first region to address AI and make laws around it (Butt, 2024, p. 7343). This significant regulation takes a risk-based approach, categorizing AI systems into four levels of danger, ranging from intolerable to low, with each subject to varying degrees of regulatory control. High-risk AI applications, such as those affecting health, safety, or basic rights, are subject to stringent standards such as transparency, data quality, and human oversight in order to mitigate risks and assure fundamental value compliance (Helberger & Diakopoulos, 2022, pp. 1751-1758).

The United States takes a more decentralized approach to AI regulation, focusing on sector-specific guidelines rather than broad legislation (Adamakis, 2024, p. 39). The Federal Trade Commission (FTC) is responsible for overseeing AI applications to ensure that they do not violate consumer laws or lead to misleading activities (Hartzog, 2015, p. 39). The FTC emphasizes the significance of ethical norms, demanding transparency, fairness, and accountability from AI developers and users in order to prevent abuse and defend consumer rights (Selbst & Barocas, 2023, pp. 1029-1044).

China's regulatory approach is distinguished by a twin goal of encouraging AI innovation while retaining strict state supervision (Zeng, 2022, pp. 1-2). Recent laws focus on specific areas, such as deep synthesis and AI-driven recommendation systems. These regulations are intended to reduce harmful content and unethical AI use by forcing service providers to incorporate safeguards that prevent discrimination and defend customer interests (Roberts et al., 2022, pp. 80-81). Furthermore, China maintains rigorous limits on AI uses in public surveillance, requiring that technology such as facial recognition be utilized responsibly and ethically (Hagerty & Rubinov, 2019, pp. 10-17).

These geographical variances reflect a fragmented global regulatory environment in which harmonization is difficult but necessary for controlling the international consequences of AI technologies. The EU's framework serves as a global benchmark for others, aiming for a balanced approach that promotes innovation while protecting the public interest. The United States prioritizes market-driven solutions above comprehensive legislation; therefore, it concentrates on risk mitigation through targeted actions. In contrast, China's model emphasizes the importance of state governance in directing and managing AI development, which is consistent with the country's larger technological and political objectives. Together, these methods reflect an intricate web of global AI governance, with each region's policy informing the larger discussion on how to best balance innovation, ethics, and consumer safety.

AI incorporation into military plans considerably improves capabilities, ranging from autonomous drone surveillance to algorithm-driven decision-making in combat scenarios (Veress, 2022, pp. 87–88). These developments offer improved efficiency and lower hazards for human combatants (Veress, 2022, p. 86). However, they also raise important ethical concerns about delegating fatal judgements to robots (Veress, 2022, p. 94). The prospect of AI systems making life-or-death choices without human interaction has prompted a worldwide debate about the moral consequences and the need for strict international regulatory frameworks (Ulgen, 2022, p. 2).

AI surveillance technology such as facial recognition and predictive policing algorithms are increasingly being used by governments for law enforcement and security objectives (Feldstein, 2019, p. 8). While these tools can greatly improve public safety, they also pose serious risks to privacy and civil liberties (Feldstein, 2019, p. 12). AI surveillance systems can lead to widespread monitoring, frequently without proper control, potentially leading in discriminatory practices and privacy concerns (Feldstein, 2019, pp. 12-13).

The international community has recognized the critical need to develop norms governing the use of AI in combat and surveillance to guarantee that it is used responsibly and ethically and emphasize the significance of human oversight in AI decision-making processes to avoid ethical breaches in conflict scenarios (Gill, 2019, p. 176; United Nations, 2023a, p. 1). The UN's debates on lethal autonomous weapons systems are at the forefront of these efforts, with the goal of developing guidelines that limit the use of AI in specific military operations (United Nations, 2023b, pp. 1-2). Such talks are critical for defining standards that support ethical conduct in the global defense industry (Butcher & Beridze, 2019, p. 93).

The ethical debates surrounding the use of AI in military and surveillance are complex, involving a wide range of issues, from AI system opacity to the potential for bias and error, which could lead to unintended escalation or violations of international law (Morgan et al., 2020, pp. 6-22). There is widespread concern that AI technologies will be repurposed by non-state actors or rogue nations, posing considerable security dangers (Roy, 2024, p. 118). These problems highlight the importance of strong, comprehensive international regulatory measures to adequately regulate the hazards connected with AI-powered military and surveillance applications (Roy, 2024, p. 135).

Establishing international standards for the use of AI in military and surveillance activities is critical. These standards should guide the development and implementation of AI technologies, ensuring that they benefit society without jeopardizing fundamental rights and freedoms. Effective regulation will help balance AI's technological gains with ethical imperatives, creating an environment in which technology breakthroughs contribute to global stability and security (Comunale & Manera, 2024, p. 3; Helkala et al., 2023, p. 1).

2.4 Cultural and Personal Dynamics

Hofstede's (1980, p. 211) cultural dimensions theory provides insightful information about how cultural backgrounds can shape how AI bias is understood and interpreted. According to this theory, cultural norms and values have significant impacts on how people from different societies view and interact with technology, especially AI systems. Moreover, perceptions of AI bias are significantly shaped by age-related factors. The concept of digital ageism has been the subject of recent studies that examine how AI technologies affect and interact with older adults. These studies demonstrate how older populations can be disproportionately impacted by AI biases, raising questions about fair access and representation in AI-driven systems (Chu et al., 2023, p. 1). These age- and culture-related factors highlight how intricate AI perception is. To make sure that AI systems are inclusive and fair for all societal segments, it stresses the necessity of developing AI policies and practices that are sensitive to diverse cultural values and age-related needs.

The literature reveals a wide range of AI biases across domains, including racial and gender bias. Kleinberg et al. (2016, p. 1) found structural racial inequities in criminal justice algorithms. Dastin (2022, pp. 296-299) demonstrates gender biases in AI recruitment, whereas Obermeyer et al. (2019, p. 1) show racial biases in medical diagnostics. Crawford and Schultz (2014, pp. 93-94) emphasize broader societal implications. Chu et al. (2023, p. 1) and Hofstede (1980, p. 211) investigate cultural and age-related factors, demonstrating the complexities of AI perception. While technical expertise can tamper with sensitivity to algorithmic aspects, personal experiences and cultural backgrounds have a significant impact on perceptions of AI bias.

The evolution of perceptions of artificial intelligence (AI) across cultures demonstrates a complex interplay of technology advancements, media portrayals, and public discourse. Initially seen as an idea from science fiction, the integration of AI into everyday life, as described by Gerlich (2023, p. 1), has transformed public perspectives from skepticism to a more sophisticated understanding that balances possible benefits with ethical considerations.

Significant milestones, such as the introduction of machine learning, neural networks, and real-world applications like voice assistants, have all played important roles in changing public perception. These advancements have frequently been followed by media hype, which oscillates between applauding AI's capabilities and fanning anxieties about its ramifications for privacy and employment (de-Lima-Santos & Ceron, 2021, p. 15).

The cultural context has an important bearing on how these perceptions evolve. In Western environments, there is a considerable emphasis on the efficiency and competitive advantage AI offers to industries, whereas in non-Western contexts, the discussion may center on societal and ethical ramifications. This divide is vividly shown in media narratives, which either magnify fears about surveillance and job displacement or glorify AI's capacity to tackle persistent global difficulties (Gerlich, 2023, p. 5).

Furthermore, legislative discussions and regulatory initiatives, particularly those addressing data privacy and algorithmic transparency, have had a significant impact on the public discourse surrounding AI. These conversations are frequently triggered by high-profile instances publicized in the media, resulting in a more cautious and informed public opinion (de-Lima-Santos & Ceron, 2021, p. 20).

To summarize, perceptions of AI are constantly altered by a complex combination of technological advancement, cultural influences, and mediated narratives, reflecting a collective grappling with the enormous effects of AI on human society.

Cultural dimensions have a profound impact on AI system design, particularly Hofstede's (1980, pp. 211-252) theory which emphasizes how cultural values shape user interface choices and customization. Cultural factors such as power distance, individuality versus collectivism, and uncertainty avoidance are critical in shaping how interfaces are constructed and perceived by users from various cultural backgrounds (Barber & Badre, 1998, p. 1). For example, high power distance cultures may prefer interfaces that provide guided navigation and organized options, reflecting their familiarity with hierarchical and authority-based institutions (Swierczek & Bechter, 2010, p. 294).

Furthermore, the individualism-collectivism dimension influences whether an interface encourages individual customization or follows social norms (Li et al., 2009, p. 68). Individualistic cultures may demand more personalized AI experiences that emphasize user liberty and privacy

(Mehmood et al., 2024, p. 14). In contrast, collectivist cultures may prefer interfaces that facilitate group interaction and communal connectivity (Wurtz, 2005, p. 279).

Uncertainty avoidance additionally shapes interface design by determining the level of complexity and amount of information given. High uncertainty avoidance cultures may necessitate clear, simple interfaces that prevent ambiguous user routes (Chadwick-Dias et al., 2002, p. 30).

Furthermore, masculinity and femininity shape the aesthetic and practical aspects of design (Schlageter, 2015, pp. 1-2). More masculine cultures may emphasize performance-oriented aspects, whereas feminine cultures may prioritize usability and aesthetic appeal (Schlageter, 2015, pp. 34-36; Schlageter, 2015, p. 2).

These cultural biases are more than simply theoretical concerns; they have real-world design implications that shape how AI systems are viewed and used in various global contexts, emphasizing the importance of culturally conscious AI development tactics.

2.5 Theoretical Divergence

Artificial intelligence (AI) bias is a complicated subject that has been studied in a variety of academic disciplines, yielding differing theoretical approaches. In computer science, bias is typically viewed in terms of algorithmic design and data integrity (Oladoyinbo et al., 2024, p. 13). Researchers focus on technical elements including data pretreatment, model fairness, and post-processing approaches to reduce bias (Ferrara, 2023, p. 3). They investigate how biases enter training data and propose techniques for developing more egalitarian algorithms by enhancing data quality and diversifying datasets (Ferrara, 2023, p. 3).

Sociologists, on the other hand, see AI prejudice as a reflection of technologically embedded societal inequities (Sartori & Theodorou, 2022, p. 1). This viewpoint emphasizes how AI systems reinforce existing social prejudices in sectors like hiring, lending, and criminal justice (Min, 2023, p. 3815). Sociologists suggest that the inventors of AI, who are mostly from homogeneous groups, unintentionally encode their prejudices into these systems (Zajko, 2022, p. 7). Addressing AI bias requires not only technical answers but also systemic improvements in the diversity of individuals building these tools.

Philosophical discussions about AI bias frequently center on ethical implications and the distribution of blame. Philosophers dispute the moral responsibilities of AI developers and users, proposing that bias reduction should be a collaborative endeavor including multiple stakeholders (Hagendorff, 2020, p. 8). They emphasize the significance of transparency, accountability, and ethical design in AI systems to guarantee they do not perpetuate harm (Ferrara, 2024, p. 4).

These different viewpoints emphasize the complexities of AI bias. While computer scientists focus on technical improvements, sociologists emphasize the importance of addressing underlying societal challenges, and philosophers call for ethical responsibility and more accountability. To address AI prejudice to its fullest, a multidisciplinary approach is required, which combines technical solutions with social and ethical considerations to produce more just and equitable AI systems.

2.6 Future Trends

The combination of quantum computing and artificial intelligence (AI) is expected to alter the AI landscape, providing new capabilities for data processing and analysis (Ahmadi, 2023, p. 18). This integration offers significant improvements in AI efficiency and effectiveness, particularly in complex problem-solving domains where present technologies struggle (How & Cheah, 2024, pp. 290-323). Quantum computing may reduce AI biases by improving algorithmic fairness through faster and more extensive data processing, addressing both present and emerging biases more effectively (Ajani et al., 2024, p. 550).

Simultaneously, developments in Natural Language Processing (NLP) enabled by quantum technologies are predicted to improve AI's grasp of human language, making these systems better at deciphering context and subtlety (Guarasci et al., 2022, pp. 1-2). Quantum NLP (QNLP) is emerging as a hybrid field that applies quantum mechanics to critical aspects of language processing, with the potential to overcome some of the limitations of current NLP systems, such as the need for large datasets and extensive computational resources (Guarasci et al., 2022, pp. 1-3).

Regulatory patterns are also changing, which has important consequences for AI development and deployment. The EU's proactive approach to AI regulation, which includes rigorous standards for transparency and bias reduction in high-risk AI applications, is likely to create a global standard, impacting how AI technologies are managed worldwide (Cloete, 2024, p. 18).

The public perception of AI is evolving because of technological breakthroughs and regulatory changes. As AI systems grow more interwoven into daily life and their effects become more evident, public debates center on ethical implications, trust, and the technology's social benefits and threats.

Overall, the combination of quantum computing, sophisticated NLP, stringent restrictions, and shifting public perceptions is predicted to affect the future of AI, making it more powerful, equitable, and in line with human ideals.

3. Methodology

3.1 Introduction

For this research, qualitative semi-structured interviews of different citizens from all around the world and of different ages, was conducted. These in-depth interviews focused on how citizens perceived and interpreted bias in AI systems and based on their perceptions and experiences, it aimed to uncover the underlying factors that cause these perceptions. This research aims to comprehend the different ways in which biases are identified and the effects these biases have on trust and reliance on AI technologies by examining a wide range of experiences. Furthermore, participants also gave their insights on how they believe AI is affecting their society and what can be done to increase trust and how they view the future of AI.

3.2 Research Design

The chosen methodology is semi-structured interviewing, which allows for flexible but guided conversations that yield rich and nuanced data (Kallio et al., 2016, p. 2955; Osborne & Grant-Smith, 2021, pp. 17-18). This approach is especially well-suited to the exploratory nature of the research, which focuses on the rapidly changing landscape of AI technology (Chennupati, 2024, p. 27). Semi-structured interviews offer a balanced approach because they combine the open-ended nature of unstructured interviews, which allows for the exploration of previously unanticipated insights, with the directive aspects of structured ones, which ensures data consistency and comparability and are especially useful for investigating the multifaceted nature of AI bias. After all, they allow researchers to delve deeply into complex and evolving topics while still providing a structured framework to ensure that specific research themes are addressed (Kallio et al., 2016, p. 2956). Participants are encouraged to freely express their opinions, allowing for in-depth discussions on specific topics (Osborne & Grant-Smith, 2021, p. 6). This methodology is critical for exploring uncharted territory in AI research, as it allows for the capture of diverse and subjective elements present in the perception of AI bias (Kallio et al., 2016, p. 2959).

Furthermore, these interviews allow researchers to gain new insights directly from participants' narratives, resulting in a more complete understanding of their perspectives and attitudes Kallio et al., 2016, p. 2955). They enable interview instructions to be adjusted in response to early findings or evolving research questions, resulting in a more responsive and reflexive research process (Osborne & Grant-Smith, 2021, p. 6).

However, there are certain limits to consider. Semi-structured interviews can produce large amounts of data that are difficult to analyze and interpret (Kakilla, 2021, pp. 1-2). The interviewer's skills, as well as the participant's willingness and ability to articulate their thoughts, have a significant impact on data quality (Kakilla, 2021, pp. 1-2). Compared to quantitative methodologies, this strategy may also require smaller sample sizes, which can affect the generalizability of the findings (Rahman, 2020, p. 105). Nevertheless, the richness and complexity afforded by this qualitative technique are invaluable for examining intricate and subjective topics like views of AI bias (Kakilla, 2021, pp. 1-2).

Semi-structured interviews, as opposed to other qualitative methods such as focus groups or fully unstructured interviews, provide a unique opportunity to gain individual insights without the influence of group dynamics, which is critical when discussing topics such as bias, where personal experiences and perceptions vary greatly (Harrell & Bradley, 2009, pp. 25-33; Luna-Reyes & Andersen, 2003, pp. 280-282). Unlike quantitative surveys, which may confine responses to specified categories, semi-structured interviews can delve further into the 'why' and 'how' questions, revealing the underlying causes for perceptions of AI bias (Carroll & Rothe, 2010, pp. 3480-3484).

This paper's goal is to investigate the multifaceted nature of AI bias and its societal implications by involving people from various cultural backgrounds and levels of technological literacy. The paper's goal of exploring unexplored areas in the field of artificial intelligence and its societal impact is well served by this method, which is perfect for capturing the varied and subjective elements present in the perception of AI bias.

Thus, this methodological decision is justified by the research's objectives to investigate personal and cultural dynamics in perceptions of AI bias, allowing for the flexibility required to collect deep, nuanced insights while ensuring that each interview covers essential components of the research questions.

3.3 Sampling and Data Collection

Convenience sampling will be the primary sampling method used in this research, as participants will be chosen primarily based on their accessibility and willingness to participate through social media applications (Etikan et al., 2016, p. 2). This research benefits from convenience sampling because it is a quick and effective way to collect an immediate sample, which is our data point, from a readily available particular population segment, in this case, citizens who use social media and works well for exploratory research aimed at gaining a preliminary understanding of a specific phenomenon, in this case citizen’s perceptions of bias in AI (Bryman & Bell, 2015; Lopez & Whitehead, 2013, p. 124).

However, the non-random nature of convenience sampling raises issues about the sample's representativeness, potentially limiting the findings' generalizability (Etikan et al., 2016, p. 2). Since participants are chosen based on their accessibility rather than random selection, the sample may not correctly represent the larger population (Etikan et al., 2016, p. 4). This may introduce bias since the sample reflects the characteristics of individuals who are most available or eager to participate rather than a true cross-section of the community (Etikan et al., 2016, p. 2).

Snowball sampling—in which initial participants are invited to suggest additional possible participants from their social networks—is also included (Noy, 2008, p. 330). This approach is especially useful when examining difficult-to-access groups or when the research focuses on features or experiences that are not apparent or known to others (Sadler et al., 2010, p. 370). Snowball sampling can reach a larger and presumably more diverse audience by relying on referrals from initial respondents, gathering a greater range of experiences and viewpoints, which is critical for investigating subjective topics like perceptions of AI bias (Sadler et al., 2010, p. 370). Snowball sampling also goes beyond the initial reach of the researcher's social media network, which solidifies the idea of diversifying the sample (Noy, 2008, p. 330).

However, snowball sampling also has its downsides. It may be biased towards more connected or cooperative groupings, perhaps overrepresenting specific opinions while underrepresenting less connected or cooperative portions of the population (Noy, 2008, p. 335). This strategy is strongly reliant on the initial participants' social networks, which may bias the research findings towards the norms and attitudes prevalent within such networks (Noy, 2008, p. 335).

This research's application of convenience and snowball sampling tries to strike a compromise between data collecting speed and efficiency and reaching a larger and more diversified participant base. Thus, convenience and snowball sampling were used for this research because they are successful at reaching diversified populations quickly and effectively, which is critical for investigations with limited resources and time (Baltar & Brunet, 2012, p. 62).

While these methods provide valuable insights into the perceptions of AI bias, especially within specific subgroups or communities, they require careful consideration and transparent reporting to understand their impact on the research's findings. The research must acknowledge the potential for sampling bias and its implications for the validity and reliability of the conclusions drawn from this methodological approach.

Social media recruitment is also another sampling strategy used in this research, with a focus on citizens using apps like WhatsApp and Instagram. This strategy works especially well for reaching a technologically literate and diverse population, which is important for research that focuses on how people perceive AI bias which they access through technologies such as their phones or computers.

Utilizing social media platforms for recruitment purposes is in line with the communication preferences and habits of the general population, which increases the possibility of participation and engagement (Kaplan & Haenlein, 2010, p. 65).

The target demographic's online behavior and preferred platforms are taken into consideration when selecting specific platforms for the recruitment strategy, which follows a methodical approach (Nuseir, 2020, p. 137). For example, younger adults are big users of Instagram and WhatsApp, therefore they are great platforms for effectively addressing this demographic (Nuseir, 2020, p. 137). Planning entails creating messaging that appeal to possible participants, emphasizing how this research relates to their experiences with technology and AI and how important it is that they contribute to our understanding of AI bias.

Engagement strategies include using direct messaging through social media apps, word-of-mouth (WOM) by asking my peers and direct family members to share the message that I am looking for participants who have used AI before, posting an invitation to participate on my Instagram story and having my friends share it on their story to recruit participants that are not in my direct community and community engagement through relevant social media groups and lastly asking if my participants if they know others who would be interested in participating. These strategies increase the likelihood that my participants will at least be aware of technologies such as AI and have experienced using it before.

Considering social media is dynamic, strategies must be continuously monitored and adapted to ensure that they are effective. Social media platforms offer analytics capabilities that can be very helpful in monitoring participant interaction, reach, and engagement rates (King et al., 2014, p. 244). This information can be used to make real-time adjustments to the recruitment plan. This adaptable strategy guarantees that the hiring procedure stays successful and efficient over the length of the research.

This recruitment technique guarantees widespread participation across varied groups by utilizing social media and upholding stringent ethical standards, so augmenting the research's significance and strengthening the validity of its conclusions.

This research gathered the minimum of 10 interview participants from different backgrounds and different ages (all above 18). The table below shows all the participant's demographics and their pseudonyms.

Table 3.1

Demographics of the Participants

Name	Age	Gender	Place of Birth	Current Residence
Hayley	21	Female	Zagreb, Croatia	Rotterdam, NL
Cam	23	Male	Zwijndrecht, NL	Zwijndrecht, NL
Peter	22	Male	London, UK	Rotterdam, NL
Abhir	23	Male	New Delhi, India	Hanover, DE
Amalia	21	Female	Japan	Rotterdam, NL
Aaron	44	Male	Gudiyattam, India	Hamburg, DE
Hunter	21	Male	Hong Kong	London, UK
Norman	22	Male	India	New York, USA
Jolie	51	Female	Seoul, Korea	Hamburg, DE
Akshan	23	Male	Hong Kong	London, UK

3.4 Operationalization

The operationalization of key constructs in this research is critical to achieving a systematic and comprehensive analysis.

In order to comprehend participants' past and present interactions with diverse technological systems, especially AI, "Integration of AI in Daily Life" will be examined. This includes how often they use AI technologies, how comfortable they are with them, and any significant experiences that may have influenced their opinions. This research can determine how participants' experiences with technology related to their perceptions of AI bias by evaluating their technological backgrounds (Venkatesh & Davis, 2020). To ensure that each construct is appropriately represented and makes a significant contribution to the research objectives, these constructs will be operationalized through carefully crafted interview questions. This includes asking them about their degree of comfort and technological knowledge, which may affect how they view AI and its biases. One question posed to the participants is: “In what situations and on how often do you use AI technologies?” This line of inquiry lays the groundwork for analyzing how technological literacy affects bias recognition by shedding light on the part that individual technology use plays in influencing perceptions of AI bias (Hayden, 1997, p. 218).

The concept ‘Perceptions and Experiences of Bias’ will be operationalized by looking into participants' subjective understandings and experiences with biases in AI systems. This will entail prompting participants to share their thoughts, reflections, and attitudes about situations in which AI systems may exhibit bias, without formal testing. Participants will be asked to recall specific instances of perceived bias in AI-driven decision-making or interactions with biased AI algorithms in their daily lives. Through a series of open-ended questions intended to elicit participants' firsthand experiences with and attitudes towards AI systems, the construct of ‘Perceptions and Experiences of Bias’ will be examined. By eliciting detailed responses from participants about their perceptions and experiences, the interviews can provide a comprehensive understanding of how people conceptualize and respond to AI bias in different contexts. Participants will be asked to share any instances in which they believed biased behavior was demonstrated by an AI system—either in the workplace, in media, or in everyday interactions with AI. For example, a question asked is: "Can you recall an instance where an AI system made a decision/gave you an answer that you felt was unfair or biased? What made you think it was biased?" This question aims to gather subjective perceptions and the reasoning behind these perceptions, providing insights into how individuals recognize and define bias in AI (Lancaster et al., 2023, p. 23).

The concept of “Cultural and Personal Dynamics” is also very important and will be put into practice using demographic questions. This concept acknowledges that cultural context has a major impact on how people view and interact with technology, and consequently AI (Hofstede, 1980, pp. 13-211). In order to operationalize cultural background, participants will be questioned about their childhood, cultural values, and societal standards that they identify with. Inquiries such as “How do your cultural or societal values impact your views on the fairness of AI systems?” will probe how these cultural variables may affect their opinions on technology and artificial intelligence. This makes it possible for this research to map out the ways that culture affects how people perceive AI, evaluating how people from various cultural backgrounds may have varied expectations and opinions about AI technology.

Based on the participants' first responses, each construct will be further explored with follow-up questions to guarantee a thorough grasp of their perspectives and experiences. Participants will be able to elaborate on their first responses to these questions, which will provide more comprehensive data for analysis.

To facilitate the interpretation of each participant's responses, the operationalization process also incorporates demographic questions that give background information about each individual. Their interactions with and impressions of AI may be influenced by the demographic information provided, which includes age, gender, education, and professional history.

This research attempts to collect comprehensive data that reflects the varied and intricate ways in which people perceive and use AI technologies by carefully identifying and investigating these constructs. This methodology guarantees that the research encompasses a wide range of experiences and viewpoints, so enabling a more comprehensive examination of the variables that impact the perceptions of AI bias among diverse groups.

All the recorded data will be sent after the interview is conducted to a secure storage space (EUR Panopto). This storage space ensures any portable devices that have the data from this research will be kept safely.

To transcribe this data, this data will be moved in a password-protected device/cloud service such as One-drive and Google drive, so the recordings can be heard and then hand transcribed and then the recordings will be deleted 2 months after the thesis is approved.

All participants have been given an alias to protect their identities and are aware of how their alias and that their data is being handled with care and confidentiality.

Following these guidelines improves the legitimacy and integrity of the research process while also safeguarding participants.

3.5 Analytical Approach to Perceptions and Actions

A two-tiered analytical methodology is used to systematically separate and analyze perceptions from actions during the interview. The first stage will be 'Perception Analysis'. During this step, participants express their perceptions of AI bias, which are then identified and categorized. It includes comprehensive coding of interview transcripts to extract themes on how biases are perceived in various socio-technical contexts. The second stage is 'Action Analysis'. This step investigates whether and how these perceptions lead to specific actions or behavioral changes. It investigates any reported changes in AI technology use, policy advocacy, or professional practice changes among participants. This method seeks to elucidate the impact of subjective interpretations of AI on practical decision-making and behavior by explicitly distinguishing between perceptions and actions. This distinction enables an examination of the direct and indirect effects of perceived biases on individual (and group) actions, resulting in a more nuanced understanding of the socio-technical dynamics at play.

3.6 Data Analysis

Although the methodology is strong, it still needs close attention to detail to be maintained during the management and analysis of the data to ensure validity and integrity. Thematic analysis is a structured and adaptable method for analyzing qualitative data, as described by Braun and Clarke (2006, p. 80), includes several essential steps to guarantee the validity and reliability of the research findings. This makes it an ideal method for the research on perceptions of AI bias.

A meticulous and methodical approach to coding and theme development is necessary to attain rigor in thematic analysis. To obtain a comprehensive grasp of the content, the researcher first thoroughly familiarizes themselves with the data by delving into the intricacies and subtleties of the responses (Naeem et al., 2023, p. 2). This stage is crucial for identifying significant and recurring patterns in the data and essential for producing preliminary codes that are based on real facts rather than assumptions (Bingham, 2023, pp. 3-4).

Initial codes are then generated (Bingham, 2023, p. 6). This entails methodically labelling particular data segments that seem relevant to the research questions (Bingham, 2023, pp. 5-6). The process of coding needs to be methodical and thorough, addressing every facet of the data that is pertinent to the research questions. Using a constant comparative method, where data segments are continuously compared with each other to refine the codes, each data segment related to the research focus is coded during this step (Bingham, 2023, pp. 5-6).

The next stage is to look for themes by grouping the initial codes into potential themes and gathering all the data relevant to each theme (Naeem et al., 2023, p. 4). This stage makes it possible to find broader patterns of meaning that emerge throughout the dataset.

After that, these themes are examined and improved upon to make sure they make sense and appropriately depict the “dataset”, which is the transcripts from our participant interviews. This entails carefully evaluating how well the themes relate to the coded extracts as well as the full dataset.

Finally, the themes are identified and defined, entailing a thorough examination of each theme’s significance and contribution to the comprehension of the research question. These themes are given names that accurately reflect the underlying data once they have been refined and validated. After that, the final report is created, which tells the story of the data concerning the research question by fusing the analytical narrative with vivid and compelling data extracts and describes how the themes play a role in theory development and broader knowledge. To make sure that the research findings are substantial and contribute towards the existing knowledge, this integration of theme identification is essential.

Thematic analysis provides a deep and nuanced comprehension of the data while also ensuring that the findings are solid, reliable, and attained ethically.

Triangulation methods are used to confirm the results even more. This entails cross-checking the consistency of the results using a variety of data sources, researchers, theories, and even methodologies. Triangulation emphasizes the convergence of data from several perspectives, which increases the research's validity (Farquhar et al., 2020, p. 166).

This systematic approach to thematic analysis will enable a detailed and nuanced understanding in answering the research question and the complementary questions and explore how citizens shape perceptions of AI bias.

For this thematic analysis, as the researcher is a visual learner, an iPad was used to display all the transcripts of each participant, it was easy have the transcript up to make direct edits on the transcript, by highlighting, color coding and making comments, and having the split view function to have my drawing apps to take additional notes.

3.7 Credibility & Ethical Considerations

When conducting semi-structured interviews, various ethical considerations must be considered to protect participants and maintain the research's integrity. Ethical considerations involve acquiring informed consent, controlling power dynamics, maintaining confidentiality, and minimizing any potential harm to participants.

When it comes to social media recruitment, ethical considerations are crucial, particularly regarding permission and privacy. Informed consent is a fundamental ethical criterion in qualitative research (Xu et al., 2020, p. 2). It is essential to communicate the goals of this research, the voluntary nature of involvement, and the intended use of the data in an open and transparent manner (Rogelberg, 2002, p. 39). Participants will be made aware that their participation is private and that they are free to stop at any moment without facing any repercussions. Additionally, before moving further, it is necessary to guarantee that participants have had any queries adequately answered and that they are completely aware of the rules of participation in order to obtain informed consent digitally (De Sutter et al., 2020, p. 2).

Maintaining data integrity and confidentiality is critical in the field of qualitative research, especially when working with technology like as artificial intelligence (AI). A strong adherence to informed consent procedures, cautious treatment of personal data during recruiting, and secure storage are all essential components of a data management plan. The collected data will be handled with care - interviews will be conducted through recording apps such as voice memos and zoom. Sensitive information is protected from potential breaches and unauthorized access via data encryption and secure storage technologies. This degree of security is crucial, especially when handling data that could contain sensitive personal information or unique identifiers.

Each participant will receive a consent form, with participants getting informational sheets outlining their rights and the parameters of the research; it is advised that they read it before the interview to understand their rights as participants and to consent to this kind of research. Consent forms are stored securely with the data it pertains to, ensuring that it can be verified if necessary (Van den Eynden et al., 2011, p. 23).

Minimizing the amount of data collected is also a matter of ethics. To lessen the chance of jeopardizing participant privacy, only the necessary information is acquired for this research.

Power dynamics in semi-structured interviews can have a substantial impact on data collection (Karatsareas, 2022, pp. 101-102; Nunkoosing, 2005, p. 699). Researchers have inherent power since they control the interview procedure and the data's subsequent interpretation (Karatsareas, 2022, pp. 101-102; Nunkoosing, 2005, p. 699). It is critical to address this by creating a culture of mutual respect and collaboration. Researchers should be able to recognize and address power disparities, allowing participants to communicate their own opinions and feelings with comfort and ease (Nunkoosing, 2005, p. 699). This entails being sensitive to nonverbal indications and tailoring the interview method to the participant's comfort level.

While semi-structured interviews are generally low risk, they can include conversations about sensitive issues that may cause participant's distress. Researchers must be prepared to deal with emotional responses provide appropriate support and not use any speech types to demean participants for their perceptions, whilst trying to not insert their ideas (Mirza et al., 2023, p. 443). As this research discusses experiences with emerging technology (AI), it tends to lean towards less emotional and intimate questions about the participants themselves and more on their usage and understanding of AI and its biases. Following the interview, participants should be debriefed to address any concerns and explain any misunderstandings. Participant well-being should be a priority throughout the research process (Mirza et al., 2023, p. 443).

By systematically addressing these ethical considerations, researchers can conduct semi-structured interviews that protect participant rights while producing valid and trustworthy data. Ethical rigor not only improves research quality, but it also fosters community trust and increases the acceptability of research findings.

When using semi-structured interviews as a research method, different limitations must be considered, as well as the researcher's reflexivity. These considerations are critical for understanding the potential biases and limits that may influence the research findings.

A major limitation of semi-structured interviews is the possibility of variability in the data acquired (Nunkoosing, 2005, p. 704). The flexibility of the interview approach might result in inconsistencies between interviews, since different participants may provide differing amounts of detail, and interviewers may investigate some topics more than others depending on the interview dynamics (Nunkoosing, 2005, p. 704).

As Karatsareas (2022, p. 102) points out, social desirability bias is a possible limitation when conducting semi-structured interviews. This bias happens when participants give replies that they believe are socially acceptable or favorable, rather than their real opinions (Karatsareas, 2022, p. 102). The presence of this bias can skew research findings, resulting in inaccurate understandings of participant attitudes and behaviors. According to Karatsareas (2022, p. 102), interviewers must be aware of this tendency and use techniques such as framing questions neutrally and creating a nonjudgmental interview environment to reduce the influence of social desirability and increase the authenticity of the data collected.

Being reflexive means that the researcher is conscious of how their actions affect the course and results of their research (Nunkoosing, 2005, p. 704). It is imperative for researchers to critically examine how their personal experiences, prejudices, and relationships with subjects influence the information gathered and analyzed (Nunkoosing, 2005, p. 704). This self-awareness is vital for reducing bias and increasing the reliability of research findings (Nunkoosing, 2005, p. 704).

The interview questions and data interpretation may be influenced by the researcher's technological background and prior AI experiences, which could result in an unconscious bias that supports particular narratives or points of view (Aquino et al., 2023, p. 5). Maintaining objectivity and integrity in the research process requires acknowledging and correcting these biases through reflexive techniques.

Conducting morally good and scientifically valid qualitative research requires exercising reflexivity and accepting the limitations of semi-structured interviews. Through a critical analysis of the impact of their own prejudices and the research methodology, scientists can reduce the possibility of errors and improve the reliability of their results.

4. Results

This chapter summarizes the results of the qualitative analysis of interviews with participants from various backgrounds. Each theme identified will provide direct quotes from participants to demonstrate the variety of perspectives and experiences with AI technologies. This factual reporting aims to create the groundwork for future discussion and conclusion.

Addressing the primary research question, “How do individuals perceive bias in AI technologies?”, the findings are divided into four themes corresponding to the four complementary sub-questions. The complementary sub-questions are: “How do concerns about data integrity and algorithmic transparency influence perceptions of fairness and trust in AI systems?”, “What are the common patterns of AI utilization among participants in their personal and professional activities?”, “What ethical concerns do participants identify in the development and application of AI technologies?”, and “In what ways do the cultural backgrounds and personal values of participants shape their engagement with AI technologies?”.

These questions seek to identify the underlying patterns and themes that emerge from personal narratives, as well as how they relate to established theories and debates in AI ethics. Each theme sheds light on the various facets of AI interaction and bias, contributing to the creation of a thorough portrait of how users view and interact with AI biases.

To extract important insights from the interviews, the analysis used thematic coding, concentrating on the participants' individual interactions with AI technologies, perspectives on the impact of AI on society, and firsthand experiences. This section attempts to provide a thorough understanding of the intricate relationships between algorithmic decision-making and human cognition by connecting these empirical findings to the theoretical framework developed in previous chapters. The results are organized to answer the main research question with the guiding sub-questions, in turn, guaranteeing a logical progression that builds upon the information gathered to support the main thesis of this research.

The sections that follow discuss these themes under the headings “Bias and Fairness in Artificial Intelligence”, “Patterns of AI Utilization”, “Ethical Concerns in AI”, and “Cultural and Personal Influences on AI Engagement”. The table below summarizes the identified themes and corresponding sub-themes that guide the discussion of the results.

Table 4.1

Themes and Sub-themes

Themes	Sub-Themes/Axial Codes
Bias and Fairness in Artificial Intelligence	Perceptions and Experiences of Bias Data Integrity and Algorithmic Transparency
Integration of AI in Daily Life	AI Utilization Patterns
Ethical and Regulatory Considerations	Ethical and Societal Considerations Trust and Reliability in AI Systems Regulatory Frameworks and Governance
Cultural and Personal Dynamics	Cultural and Individual Influences Future Projections and Concerns

4.1 Bias and Fairness in Artificial Intelligence

For 'Perceptions and Experiences of Bias', all the participants had their perceptions on what they perceived bias as, depending on if they had been exposed to AI systems before.

Akshan (23) said "I feel that bias in AI can only be produced when that training data is deemed to be unfair or skewed negatively to multiple different factors.… The bias stems from things like discrimination or racial profiling", similarly, Aaron (44), who works in the aerospace and healthcare industry, has further insights into the AI field, and mentioned that "some use cases in our industry is due to the lack of training or maybe the training data which is biased towards a certain aspect…the bias could come from training, …algorithm, ...cognitive,... the human is making the models", both participants alluding towards their perception leaning towards the fact that they highlight that the training data is one of the main issues in the appearance of biases in AI systems.

On that same note, Aaron (44) goes on to say "We have to customize a lot of, you know, the in the aerospace…It's due to the availability of these customized versus non-customized areas. So, the prediction algorithms are more favoring". Akshan (23) also mentioned "The training data… like all these norms and values can get into the systems, but it's via the initial data set…there are subjective interpretations during data curation and the lack of varied viewpoints in development teams", further explaining how he believes the inherent bias comes from this initial training data set and that just goes to show that the development of AI is a continual cycle that all stakeholders could be involved in as it is everchanging.

Peter (22), mentions how “it’s a prejudice that the AI has against certain groups, either of people, certain categories, ...humans, ...strategies for a sales company, products, even nature biomes …Really, just any type of category the AI has a bias towards.”, another respondent who shared a roughly similar perspective was Hunter (21), who said that “AI bias would just be an AI system having a preference for maybe a type of answer or a preference for certain facts or figures or something like or the way they phrase things even”, both participants here share a common understanding to how the bias in AI systems can be towards just about anything and that the bias is not contained towards one aspect.

Norman (23) highlighted how “it comes from like machine learning, right? You fill it with a lot of information and then it learns, kind of like the human brain, right?”, similarly to that, Hayley (21) also said “it’s basically how people have biases, AI has biases as well within certain groups of people or societies…certain history that it kind of has wrong information about because of the creators' biases”, both participants recognized how when humans are the ones to make such AI systems, the AI systems tends to reflect how the human brain creates bias and how these AI systems just reflect unconscious biases that may have submerged in the creation and feedback processes that AI systems go through.

Whereas Jolie (51), Cam (23), Akshan (23) and Amalia (21) shared that they have used AI systems before but have not experienced bias when getting answers from generative AI sources like ‘ChatGPT’. Jolie (51) mentioned, “Personally, I don't think of any instances that I’ve experienced that, but I guess I read enough and I understand enough”. Similarly, Cam (23) said “I personally never had AI write something that’s inappropriate…But I normally use it for creating things. Yeah. And writing business texts”. Akshan (23) added “Personally, like, I’ve never experienced that sort of bias with AI, but I studied it… in my undergraduate degree for law”, which shows that in his experiences with AI systems, bias was not so apparent nor frequent, until he had to dive deeper into the topic of AI by studying it, which further promotes the idea that AI is not a common thing people tend to know about unless presented to them. Amalia (21) said “Honestly speaking I never felt anything that like jumped out of me like something like that that really left me” hinting towards the idea that if general citizens use the AI tools available to them online for “simple tasks”, the chances of bias occurring is less due to the nature of the prompts and what the AI is being used for.

On the other hand, Abhir (23) studies architecture and gave his perspective on what he thinks AI bias for him in his AI usage context looks like; “you get an answer which is, I would say, politically right. It’s not as per your liking, but the answer is quite what you would get in a very normal Google language or like an interview language”. Hinting towards a sort of idea that there is one tone of voice that makes these systems neutral and that is by being general and not specific, to which he also mentions “it felt like the answer wasn’t really designed as per my wish, but it was designed as more like a general answer”, further solidifying his perception on the fact that the AI systems are not designed to be super niche in all topics and areas and work as a general generative algorithmic technology. Sometimes Abhir (23) asks AI for its input on his ideas, one of the things he mentioned was that in his experiences with AI systems, the AI’s “wouldn’t say a very detailed answer or a very personalized answer, but I did not really expect this generic answer” further solidifying the idea that his perception of the AI system is that these systems are molded to fit everyone’s needs.

Abhir (23) also added “, I could have just gotten these answers if I were to write it in Google or ask another architect”. Another insight given by Aaron (44) was that “The human judgment is playing a major role still because due to complexity…, the model predicted by AI and versus the human judgment on the same areas, we discover anomalies… and the lack of justification of this prediction”. Both perceptions add onto the discourse that some of the participants agree that the human judgement sometimes is necessary or even at times better at getting the “job” done.

Hunter (21) discussed how “it is very important that companies are transparent because this allows for the general public to trust the company more”, reflecting how the previous literature mentioned that the public tends to trust things which they are aware and educated on, this also being topics like how AI systems work, whereas in comparison Jolie (51) said “I don’t see companies doing things out of the goodness of their heart. I see it actually as something that’s more regulatory and necessary because of GDPR rules” arguing that there could be enforcement from regulatory bodies in the future to further explain how these AI systems work as AI developers usually tend to be private companies rather than public ones “The company’s goal is to make money, right?... Therefore, they need to be competitive, and the whole concept of being competitive means that you’re not going to share everything about what you do and why you do things”.

Akshan (23) added, “Along with the guidelines being proper, there should also just be a level of transparency between developers, government officials, all like the higher-ups”, encouraging all the relevant potential stakeholders to collaborate and protect their citizens.

Norman (23) argued that if companies were to let out the stage of their development processes to the public, it could foster more trust among the users who understand technical sorts of language; “this technological knowledge even if they do publish it, only a very technological mind will understand these documents and what they mean. So, I think it will still be used quite a bit” but still mentioning that it is important to garner trust by revealing issues or biases their systems may face.

On the other side of the argument, Amalia (21) said “I don’t think the public needs to be engaged at all. Like, they don’t need to know”, Amalia (21) further explained that “I think majority of the population won't even understand it", sharing similar opinions to Norman (23) that the general population who are not technologically literate may not be able to understand development reports of AI systems. To add on Akshan (23) also said "I don't think the initial building blocks are necessary. I think once you have your prototype the first test is never going to be perfect. So, I think once you perfect it, then it's okay to show it", also added a similar idea that the initial steps do not seem to be that concerning to most users and only till the AI system is complemented, then it makes sense to ask people directly for their opinion on how the system is working.

4.2 Integration of AI in Daily Life

For 'Integration of AI in Daily Life', participants shared how they used AI in their lives. Jolie (51) added "I'm using AI all the time, like even writing a text to my friends…Predicting what I might want to say, … I actually use the suggestions that they put out there, which makes life easier" but at the same time mentioned that "my college degree was mathematics but I actually starting doing a degree with computer science" further alluding to fact that her exposure to learning about computer science at a younger age helped her add AI systems to her life and helped her with understanding how some of these AI systems work.

Abhir (23) also explained, "First use is just for me when I use, at times, uh, when I use, um, AI for architectural purposes… the person level … I do try to search for upcoming opportunities and how can I excel". Both these participants mentioned how they use AI in their daily lives for various tasks, sometimes seeing the AI as an advisor for their individual contexts.

Cam (23) mentioned how he has integrated AI into his daily life, as he owns his own retail store, he utilized AI to complete tasks for him and mentioned that "it gave me such good logos, good designs, everything accurate, the product information, everything was so good that I was like, why would I even go through the process. It will cost more. It will take more time", alluding towards the idea that in a business, efficiency and costs are aspects entrepreneurs prioritize.

4.3 Ethical and Regulatory Considerations

A central theme is ethical and regulatory concerns, where participants voiced worries about how AI would affect social norms, employment, and privacy. Akshan (23) said "It's critical to achieve a balance between, innovation and concerns with ethics" further going on to explain that "The risk of prejudice and discrimination or bias has to be properly addressed…the line kind of is drawn with prioritizing justice, transparency, and accountability in AI development”. Aaron (44) similarly added “The society feedback or whatever the means of engagement has to be an integral part of this AI models”.

Jolie (51) also added “Our daughter, she knows ChatGPT, and she knows how to use it, but she says she would never use it because she thinks that the quality of the information that comes out is not trustworthy”, this is how Jolie was informed through her environment and thus her overall trust seemed to be low. These views seem to be contrasting each other but it does prove previous literature in the idea that older people tend to be less optimistic in the world of technology.

For ‘Regulatory Frameworks and Governance’, participants shared their different opinions. Cam (23) said “I feel like there should be some control and regulation…If I’m very good at sales and you’re a really good creator and marketer and I’m just someone pretending to be using AI and then getting your work… It’s not fair” reflecting how in business contexts, the adaptation of AI being used over real humans can be disheartening for the person who decided to not use AI systems and that regulatory bodies should pose something to protect people who have job insecurity due to AI systems being able to do their job quicker and faster, many of the other participants also shared a similar opinion to this.

Hayley (21) also added “For the jobs part that you know they kind of don’t make it too big of a thing which I think it’s unfortunately not going to happen, and I think that ideally again it would be good” mentioning job security and how the government and regulatory bodies should look out for that, though she is not the most hopeful about it.

Peter (22) proposed a different type of solution to regulating AI, “you need a third sector…bubble that is meant to review, refine, deliver products…this is going to be used not just for the people of now but for generations to come”. Similarly, Aaron (44) mentioned “So it could be good to have a kind of regulatory, a kind of specific to the AI, and then what level of AI and the certification of this AI”.

4.4 Cultural and Personal Dynamics

All participants shared where they were from and discussed if their cultural values and personal views tend to shape their perception of AI and its biases. Amalia (21) mentioned “I think AI is created to be equal, fair…right by the law, open, more open-minded structure” and continued to explain that “because I grew up in such an open-minded and accepting backlight environment, I think, aligns with a lot of my morals, values”, thus reflecting on the idea that Amalia (21)’s perception on how she trusts AI had somewhat been shaped by how she was raised in Asia.

Individuals' interactions with AI are heavily influenced by the global and local environments into which they are embedded. Peter (22) said “I like to know why I have certain ways of answering certain things based on my cultural background, not in a judgmental way. I’m just curious… Fairness, justice, equality, and equity are very important to me", Peter (22) grew up in different places and was set in an international setting at a young age, and goes on to explain how the idea of fairness was important to him as he picked those values from his cultural mix of where he grew up and that it did affect the way he viewed ideas of equity, equality, fairness and trust.

Norman (23) says "I'm always questioning things…I guess just aware and knowledgeable about the world, so that gets me asking more questions…Maybe someone who's raised in just one city in their little town will not be asking". This goes to show that having exposure to different cultures could make someone more open to understanding how AI operates as the curious side of them will want to find out.

5. Discussion

The Discussion section seeks to understand the prior findings, placing them within the larger research framework on AI biases. This analysis will examine the consequences of these views on policy, technology design, and societal effect, comparing them to existing research and theoretical insights.

5.1 Bias and Fairness in Artificial Intelligence

The participants' perceptions and experiences with AI biases are examined in this theme, which raises questions about justice and the moral implications of AI systems. This theme is essential to comprehend the ramifications of biased AI since it closely corresponds with research questions about the moral dilemmas raised by AI technologies and the demand for just AI solutions from a societal point of view. Therefore, this section aims to answer the first complementary sub question "How do concerns about data integrity and algorithmic transparency influence perceptions of fairness and trust in AI systems".

Fairness and trust in AI systems are greatly hindered by concerns about algorithmic transparency and data integrity (Lee, 2018, p. 1; Oladoyinbo et al., 2024, p. 5). As they directly affect users' confidence in AI decisions and the ethical implications of these technologies, these concerns (that the participants had) are crucial in establishing the trustworthiness of AI technology (Kizilcec, 2016, p. 2393).

All participants did agree with the idea that if they understood (and if not already, get educated on) how these AI systems work, they would have a better understanding of how it works thus being able to trust it more, e.g. Jolie (51) said: "that's kind of a big part that I feel like there's a lot of education and a lot of understanding one another and it just doesn't exist and so I feel like that would play a role in bias or not". Building trust in AI systems requires transparency, which includes being open about how data is utilized, and decisions are made (Schmidt et al., 2020, p. 261). If people can verify the fairness of the methods and comprehend the processes involved, users of transparent AI systems are more likely to trust the results (Li et al., 2022, p. 41). This understanding is essential, particularly when AI affects important facets of people's lives including work, healthcare, and legal decisions (Li et al., 2022, p. 41).

With the perceptions that are shared in the results sections and existing literature, to answer the complementary sub-question guiding this section "How do concerns about data integrity and algorithmic transparency influence perceptions of fairness and trust in AI systems", perceptions of data integrity and algorithmic transparency in AI systems are significantly shaped by concerns about algorithmic transparency and data integrity. Since it shapes how people view and trust AI systems, ensuring transparency and integrity is both a technological challenge and an ethical one. The participants all explained how transparency in AI operations is essential for fostering trust, particularly in areas where decisions have a big impact on people's lives, such as healthcare and law. It enables users to comprehend and validate the procedures that go into these judgements. This is hand with existing literature which explains that when users can observe and comprehend the decision-making process and feel that the data being utilized is reliable and treated ethically, they are more inclined to trust artificial intelligence (Kizilcec, 2016, p. 1).

These are the general perceptions and ideas of AI bias shared amongst some of the participants depending on their exposure. Whilst some of it aligns with the previous literature presented, it is highly dependent on what contexts these AI systems are being used for that reflects bias.

For 'Data Integrity and Algorithmic Transparency', all participants shared their perception of how they think the most efficient way is for honesty between the AI developers and the consumers. Perspectives ranged from ideas of most participants agreeing that AI developers should be more engaged with the public to foster trust between the consumers and the machines, to some participants arguing that governments and regulatory bodies could enforce transparency if the developers do not do it first some participants just disagreeing with the idea that type of information needs to be shared.

Additionally, data integrity guarantees the accuracy, relevance, and objectivity of the information used by AI systems. The fairness of AI applications can be undermined by compromising data integrity, which might result in AI outputs that either reinforce preexisting prejudices or introduce new ones. Users may become less trusting of these systems as a result, believing them to be unfair or untrustworthy (Oladoyinbo et al., 2024, p. 4).

Open communication about the operation of AI systems, the use of data, and the process of decision-making is essential to an ethical AI commitment. This aligns with the findings from previous literature mentioning how transparency, people are better able to understand AI procedures and generate knowledgeable opinions regarding the dependability and equity of these systems (Schmidt et al., 2020, p. 261).

AI developers and companies are encouraged to actively engage with the public and regulatory organizations to promote a greater knowledge of AI technologies since this will help to assure fairness and foster confidence. This interaction can promote more acceptance of these technologies and reassure consumers about the moral application of AI as expressed in the results section.

In conclusion, resolving issues with algorithmic transparency and data integrity is essential to preserving AI systems' ability to be trusted and equitable. To guarantee the responsible and ethical application of AI technologies, these endeavors necessitate constant communication among developers, users, and regulatory bodies.

5.2 Integration of AI in Daily Life

An important theme that emerges is the integration of artificial intelligence (AI) in daily life, emphasizing the ways in which AI technologies are regularly used in both personal and professional contexts. Participants talked about a variety of uses for artificial intelligence (AI), from its essential applications in professional settings like data analysis and decision support, to its use for routine tasks like scheduling and querying. This theme highlights the widespread influence of artificial intelligence (AI) on people's daily routines and professional obligations, thereby directly addressing research inquiries concerning AI's effects on daily life and efficiency. In this theme, the second guiding sub-question will be answered: "What are the common patterns of AI utilization among participants in their personal and professional activities?".

To answer this question, whilst looking at 'AI Utilization Patterns' that the participants described, recent academic work highlights the significant shaping of demographic characteristics, including age, education, and professional background, on the integration of artificial intelligence (AI) into daily life.

With their upbringing in the midst of a technological boom, Generation Z in particular makes heavy use of AI in learning environments. They also show a preference for hybrid learning, which incorporates multimedia content, a trend that is changing expectations and resources in education (Bhalla et al., 2021, pp. 3-7; Seemiller & Grace, 2016, p. 12; Seemiller & Grace, 2017, p. 58). In contrast, older generations like Millennials and Gen X are more cautious when it comes to AI and value human engagement in addition to technology (Seemiller & Grace, 2017, p. 60). This was reflected through the older and younger participants with their levels of trust.

All the participants, especially those who have discussed AI in educational settings such as Hunter (21), Akshan (23), Hayley (21), Abhir (23), Peter (22), Amalia (21) and Norman (22) all mentioned how discussions during class or previous academic experiences encouraged them to know about these AI systems and utilize the power of AI in their daily lives. Hayley (21) for example mentioned how in her course "We had a paper where we were very strictly told to use it and even like reflected and everything, so I use it", showing how education institutions are possibly adapting with the technological advancements of AI. This again goes with the previous literature in the levels of trust and integration of AI in daily life.

Further analysis reveals that the STEM fields are the main areas in higher education where AI is being adopted. Here, it is being used to improve student engagement and personalize instruction, which is in line with the learning preferences of the younger, tech-savvy generations (Gough et al., 2017, p. 42; Page et al., 2021, p. 110). This is consistent with a larger trend in which people's engagement with artificial intelligence is greatly influenced by their technological literacy, which in turn affects their propensity to incorporate these technologies into both personal and professional spheres (Long & Magerko, 2020, p. 34).

Whilst AI is widely used in consumer technology like wearables, smart homes, and personal assistants, it also raises serious privacy issues. Privacy issues are a major component of ethical concerns, including questions regarding how AI technology stores and uses personal information. This extends to the capabilities of AI applications to perform invasive surveillance, notably in consumer electronics, where the gadgets can capture substantial personal information, raising concerns about the potential exploitation of such data. Jolie (51) added, “It’s kind of hidden in the usage of all of our phones, our Alexa…”. As AI technologies become increasingly integrated into everyday life, striking a balance between improved user experience and data security threats is a crucial issue that will only grow in complexity (Chui et al., 2023, p. 5; Hoy, 2018, p. 82; Wolfson, 2018). To guarantee that AI innovations prioritize user safety and privacy while promoting innovation, developers, users, and legislators must work together as AI advances (Li et al., 2022, pp. 28-53; Moore et al., 2022, p. 5).

These diverse interactions highlight the need for AI systems to be flexible and available to a wide range of users, guaranteeing that the advantages of AI are shared fairly while attending to the specific requirements and concerns of various user groups.

5.3 Ethical and Regulatory Considerations

This theme is essential for looking into the ethical conundrums raised by AI technologies, their governance, and how existing and new regulatory frameworks attempt to address them. It draws attention to the ongoing discussion about the necessity of strong, open regulations that can control the creation and application of moral AI systems. This theme is complemented by the sub-question: “What ethical concerns do participants identify in the development and application of AI technologies?” to which all participants discussed the ethical concerns they have. The concerns bought up highlight the importance of ongoing communication among all stakeholders in AI—developers, users, regulators, and the general public—to ensure that AI technologies are developed and implemented in accordance with ethical standards and societal norms. The overriding subject is the importance of education and transparency regarding how AI systems work, their applications, and the broader ramifications for society. This is in line with Kizilcec (2016, p. 2393) findings that establishing trust between the AI and user is key.

For ‘Ethical and Societal Considerations’, participants expressed concern that AI systems could perpetuate existing societal biases or bring new types of discrimination. Such concerns are consistent with the findings of Johnson et al. (2019, p. 510), who discovered that AI systems frequently mimic societal and cultural prejudices encoded in training data. This ability to influence broad societal norms and individual opportunities emphasizes the importance of incorporating fairness, transparency, and accountability into AI development—themes also emphasized by Ferrara (2023, p. 3), who highlighted the critical role of ethical guidelines in AI governance.

Furthermore, participants understood the significance of user interaction in machine learning processes in mitigating bias risks. This feedback loop, in which user experiences assist modify AI systems, supports Gerlich's (2023, p. 502) claim that participatory design in AI development can improve system fairness and inclusivity. By incorporating users—who are frequently the direct recipients of AI biases—in the development process, AI systems can be better scrutinized for unintentional discriminatory impacts and corrected in accordance with ethical standards.

For 'Trust and Reliability in AI Systems', many participants reported broad trust in AI systems, while there was some heterogeneity impacted by demographic factors such as age and context of AI use. This is consistent with the findings of Lee and Coughlin (2014, p. 10), who stated that elderly people are more skeptical about AI technology due to a lack of knowledge and innate distrust in digital solutions. In contrast, younger users frequently rely on AI for knowledge and decision-making support. Additionally, Cotten et al. (2013, p. 3) address how people's familiarity with technology affects their level of trust, positing that people are more inclined to trust AI in non-critical or comfortable situations, like when they use generative AI for straightforward questions. For example, Hayley (21) stated, "Instead of Googling something, just asking ChatGPT, like I do, and a lot of my friends who are international students in Rotterdam do", demonstrating a trend in which younger demographics are not only more trusting of AI but also prefer it over traditional information sources.

According to Tai (2020, p. 339), people are more likely to trust AI in familiar or non-critical circumstances, such as utilizing generative programs like ChatGPT for simple queries. This finding suggests that the context in which AI is used also greatly impacts trust levels. This demonstrates a comprehensive grasp of the ways in which various use cases may influence users' perceptions of the dependability and trustworthiness of AI systems, indicating that user experiences and contextual familiarity are important factors in forming user trust.

For 'Regulatory Frameworks and Governance', many perspectives were shared. A major concern was how AI might affect jobs and societal standards. Participants expressed concern about job displacement caused by AI automation, as well as broader societal transformations that may arise from growing reliance on technology. This raised concerns that AI may change existing societal roles and relationships, potentially leading to a redefinition of jobs and social engagements. Some participants advocated for a strong regulatory framework to effectively govern the development and deployment of AI systems. Without strong control, AI technology may be exploited, compounding its negative effects. The debates emphasized the importance of international standards and cooperation to ensure that AI technologies are properly managed across borders, reflecting the global nature of technology and its implications.

The prospect of AI systems making life-or-death choices without enough human oversight was especially concerning. Participants advocated for strict ethical norms and oversight to prevent potential misuse and guarantee that AI applications in these sectors are carried out properly.

5.4 Cultural and Personal Dynamics

The subject of cultural and personal dynamics is how people's experiences and interactions with AI are shaped by their personal histories and cultural contexts. The participants' personal tales illustrated how their interactions with AI technologies are influenced by their unique experiences and cultural prejudices. This theme adds value to the research by shedding light on the various ways people from various cultural backgrounds view and engage with AI, which is directly related to research questions concerning the influence of cultural factors on AI adoption and perception.

Cultural origins and personal beliefs have a complex impact on how people perceive and interact with AI technologies. Drawing on Hofstede's (1980, p. 211) cultural aspects theory, it is clear that cultural norms and values shape how people from various societies perceive and use AI technologies.

This section aims to answer the question "In what ways do the cultural backgrounds and personal values of participants shape their engagement with AI technologies?", some participants noticed direct effects of their upbringing in trusting these AI systems, but the rest of the participants may have not had culture as a main factor for shaping their perception but other factors, as previously discussed.

Participants' understanding and adoption of AI technologies are heavily shaped by their cultural backgrounds aligning with Hofstede's (1980, pp. 211-252) findings. Participants from various cultural backgrounds offer distinct viewpoints on their interactions with AI, driven by societal conventions and personal experiences. For example, Amalia (21) sees AI as matching her ideals of openness and fairness, which reflect her background in an open-minded cultural milieu. This suggests that cultural circumstances influence not only personal beliefs but also perceptions of technology's function and fairness.

Engagement with AI also differs depending on how cultures perceive technology. Hayley (21) also mentioned "In Croatia, ChatGPT isn't used as much. I feel like I don't have a lot of friends here that use it…I'm more skeptical towards it because of that because it's very like not a thing here" but then proceeded to give insights on "in regards to Rotterdam…where I Google everything through ChatGPT" giving an interesting comparison that shows that depending on where Hayley is, the culture and the way the culture views AI affects if Hayley decides to use the AI or not. Hayley (21) shows a geographical variance in trust and use of AI, with Croatia having less familiarity and skepticism about AI than

Rotterdam. This is consistent with Hofstede's (1980, pp. 211-252) findings that societal structures and prevalent attitudes towards technology might impact the acceptance and integration of new technologies into daily life.

For instance, Norman (23) uses his own experiences of living in several continents—including Asia, Europe, and America—to demonstrate how cultural diversity affects viewpoint. He describes how his exposure to a variety of cultures has increased his curiosity and openness. This is consistent with research by (Gelles et al., 2024, p. 19845), who found that exposure to many cultures can greatly improve cognitive flexibility and openness—qualities essential for critically interacting with sophisticated technology like artificial intelligence.

People from different cultural origins, like Peter (22) and Norman (23), frequently highlight values like justice, fairness, and curiosity. These values affect how they interact with AI systems. Their experiences confirm the findings of (Gelles et al., 2024, p. 19845) about the contribution of cultural variety to the promotion of ethical awareness in technology use, underscoring the significance of diverse cultural insights in constructing a thorough understanding of AI and its ethical implications.

Perspectives on AI are also shaped by cultural backgrounds and professional experiences, as industry professional Aaron (44), who supports openness in AI systems, points out. He states, "In fact, as an individual, as well as working in the industry, I would expect that this AI model can explain how the decision has been made, how the result has arrived". This need for explainability aligns with the demands of the industry as a whole and supports research by (Hagendorff & Fabi, 2023, p. 112), who contend that user trust in AI operations is largely dependent on openness in AI operations. Aaron's demand for transparent AI operations underscores the junction of professional expertise and the need for accountability in AI technologies, and it represents the rising consensus—discussed in the literature—that comprehension of AI processes is critical to user confidence.

These diverse cultural interactions suggest that AI development must consider a broad spectrum of cultural values to ensure inclusivity and fairness. Designers and policymakers should incorporate these cultural insights to tailor AI technologies that are accessible and acceptable across different cultural contexts, ensuring that AI systems respect and reflect the diversity of user backgrounds.

5.7 Chapter Summary

The primary relevance of the study is the critical importance of developing fair and accountable AI systems since there is significant pervasion of AI technologies in all sectors. About this, the findings reveal that respondents are very cautious about the potential for bias within AI algorithms to perpetuate existing social inequalities. This has, in response, elicited much advocacy by stakeholders to establish robust frameworks that will see to it that AI systems are advanced not only in technology but with ethical standards that promote fairness and accountability. Further analysis indicates that, technological solutions to bias are needed, combined with transparent regulatory practices that might be independently audited. This dual approach is in line with the overarching demand in society for AI systems to be both practical and just: to serve the public good without widening social divides. It is further suggested that including ethical considerations at the early design stage might well act as a way to 'pre-address' many of these concerns about bias and fairness in the AI system. Human beings, together with technologists, policymakers, and members of the public, must work in conjunction to ensure that AI is deployed ethically. This entails developers being continuously educated on AI ethics, adhering rigorously to guidelines, and the general public involving themselves in discussions regarding the implications of AI in society. Looking at these arguments, it is evident that the way toward fair and accountable AI requires a commitment from all stakeholders. Culture can be created to make people more responsible and ethical in a way that would mitigate most of the risks associated with AI biases and improve the trustworthiness of AI applications.

6. Conclusion

This thesis seeks to investigate the various perceptions of bias in artificial intelligence (AI) among diverse individuals. Centered on the key research question, "How do different individuals perceive bias in AI?", the present research delves into the deep understanding of AI's impact on our daily lives and societal structures. To supplement this core inquiry, four essential sub-questions were posed to provide depth and broader context: "How do concerns about data integrity and algorithmic transparency influence perceptions of fairness and trust in AI systems?", "What are the common patterns of AI utilization among participants in their personal and professional activities?", "What ethical concerns do participants identify in the development and application of AI technologies?", lastly, "In what ways do the cultural backgrounds and personal values of participants shape their engagement with AI technologies?". The responses to these questions have allowed for a thorough investigation of the varied connections between humans and AI systems, revealing major differences in trust, reliance, and ethical considerations affected by personal experiences and social ideals.

The findings suggest that concerns about data integrity and algorithmic openness have a significant impact on views of justice and trust in AI systems. Participants reported a significant desire for transparency in order to trust AI choices, echoing broader calls for accountability in AI operations. This requirement for openness is viewed as a basis for establishing AI systems' trustworthiness, particularly in high-stakes areas such as healthcare and legal decisions (Lee, 2018, p. 1; Oladoyinbo et al., 2024, p. 5).

The research indicated various patterns of AI use, with differences mostly affected by demographic characteristics such as age and occupational background. Younger individuals, who were generally more technologically savvy, incorporated AI more easily into their personal and professional lives. In contrast, older generations engaged cautiously, emphasizing a combination of human interaction and technical assistance (Seemiller & Grace, 2016, p. 12; Seemiller & Grace, 2017, pp. 58-60).

Participants raised substantial ethical concerns about the development and implementation of AI technology. Key topics included AI's propensity to perpetuate current socioeconomic injustices, as well as the vital need for ethical frameworks to govern AI development. These problems highlight the societal and moral components of AI technologies, which necessitate strict ethical standards and strong regulatory frameworks to prevent biases (Pagano et al., 2023, p. 34).

This research underscores the indispensable role of cultural diversity in shaping perceptions and the development of artificial intelligence systems. By examining how diverse cultural backgrounds influence user interactions and biases towards AI, the study reveals that inclusively designed AI systems are crucial. These systems must not only address the functional needs of users but also respect and integrate their cultural values and norms. This approach enhances both the fairness and effectiveness of AI technologies, ensuring that they are perceived as trustworthy and are genuinely beneficial across varied global contexts. Recognizing the profound impact of cultural diversity on AI perceptions guides critical advancements in AI development, advocating for a design philosophy that upholds equity and inclusivity at its core (Hofstede, 1980, pp. 211-252).

Furthermore, by reorienting the attention from the strictly technical components of AI to the user-centered perceptions of bias, this research innovates within the field of bias perception studies. According to this unique method, algorithmic fairness should not be the only metric used to quantify bias in AI; users from different cultural and personal backgrounds should also be considered when assessing the system's bias. An enlarged viewpoint like this might result in more all-encompassing approaches to AI development and bias reduction, highlighting the necessity of a multifaceted strategy to comprehend and deal with AI prejudice.

This research aims to contributes to the conversation about how cultural norms and values influence how people interact with technology by incorporating findings from cultural studies. It suggests that user perceptions are essential to comprehending the whole impact of technology on society, offering a useful framework for future research into the cultural influences on other technical interactions. It also contributes to the advancement of scholarly discourse and lays the groundwork for subsequent investigations aimed at dissecting the intricate connections among culture, perception, and technology advancement.

The research's practical implications offer valuable insights for many stakeholders engaged in the creation, advancement, and management of artificial intelligence systems. Firstly, this research's conclusions about the various perspectives on the application of AI highlight the necessity of designing AI systems with an inclusive methodology that takes cultural and individual variances into account. This can help developers design adaptable AI systems that satisfy the functional needs of various user groups while also conforming to their cultural norms and ethical expectations, building user acceptability and trust.

Additionally, the effects of algorithmic transparency and data integrity on fairness and trust provide important discussions for policy improvements in AI usage. These insights might be used by legislators to create more precise rules that guarantee AI systems are open about how they utilize data and make decisions. Regulations of this kind would aid in reducing prejudices and boosting public confidence in AI applications, both of which are necessary for the technology's broader adoption and moral incorporation into society.

This research also emphasizes how critical it is to improve AI education and literacy. Users may be better equipped to interact with AI systems critically and successfully if educational initiatives improve knowledge of both the technical and social aspects of the technology. This is especially true in a time when artificial intelligence is influencing every aspect of life, necessitating critical and knowledgeable consumers in order to responsibly evolve AI technologies.

Thus, these findings imply that corporate responsibility is now more important than ever in the development of AI. Businesses should use these findings to hone their tactics so that their AI products are accepted by a wide range of users and are not biased. Companies can better match their products with the societal values and expectations of their customers by putting user trust and ethical considerations first. This could result in increased adoption rates and customer satisfaction.

In essence, this research's practical implications emphasize the value of a multi-stakeholder approach in the advancement of AI technology, while also serving as a bridge between theoretical research and practical application. This approach guarantees that the development of AI is not only technologically sophisticated but also morally and socially responsible.

This research adds to scholarly discussion and provides concrete suggestions for AI development and application by combining these theoretical ideas with real-world considerations. It implies that more moral, just, and practical technological solutions can result from bridging the gap between how AI is built and how consumers perceive it.

This section offers a solid framework for comprehending this research's broad ramifications, laying the groundwork for further investigation and application of these discoveries in theoretical research and real-world AI development.

The research conducted here has made considerable strides in our knowledge of how bias in AI is perceived in various cultural and individual circumstances however it is imperative to acknowledge the inherent constraints of the research design and methods as they may have impacted the findings.

One of the main limitations of this research is the small sample size, which was limited due to resource and time restrictions. Although the qualitative approach allowed for in-depth, insightful studies of individual experiences, the results may not be fully generalizable to a larger population. Future studies could solve this constraint by using a larger, more diverse sample with a broader range of demographics and geographical areas.

The reliance on semi-structured interviews, while useful for acquiring rich, detailed data, has drawbacks. The data gathered is highly subjective, relying mainly on individual opinions that may miss wider, more universally relevant findings. Furthermore, the nature of personal reporting may add biases such as recall bias or social desirability bias, in which participants adapt their responses to what they perceive is anticipated or acceptable.

The scope of this research was limited to perceptions of AI bias; hence it did not empirically test AI systems for actual biases. This emphasis on perceptions, while useful, does not provide a direct assessment of the AI systems themselves. Future research could combine empirical testing and perceptual studies to create a more comprehensive understanding of AI biases.

Given that AI technology is always changing, the findings of this research may have limited longevity. AI systems and their societal implications are subject to change, therefore some of the findings from this research may become less relevant over time. Continued research is required to stay up with technological breakthroughs and their effects on society.

Despite efforts to ensure neutrality and objectivity, there is always the possibility of researcher bias in data interpretation in qualitative research. Reflexivity was used to reduce this risk, but some subjective interpretations are unavoidable.

Recognizing these limitations does not diminish the value of this research, but rather provides avenues for future investigation and development. Future research can build on the groundwork set by this thesis, overcoming these constraints and investigating other facets of AI's impact on society.

Based on this research's findings and limitations, various areas for future research could investigate and address the intricacies of AI bias. Increasing the variety of the sample in research projects is one important topic. Future research should broaden the scope and diversity of this research's sample to include a more diverse range of cultural, professional, and demographic backgrounds. This extension would improve the findings' generalizability while also offering a deeper comprehension of the diverse ways in which different groups interpret AI bias.

Furthermore, considering the rapid pace at which AI technology is developing, longitudinal research may offer important new perspectives on how people's views of AI bias evolve—particularly as they grow more accustomed to using these tools in their daily lives. This research may provide useful information for creating AI systems that are more focused on the needs of users by tracking changes in perception and trust.

Future research could take these results even further by involving even more diversified populations and mixing methodologies to ensure further robustness; this study may also include experimental designs to see the real-world impact of AI biases. This comprehensive approach will strengthen the understanding of the sociotechnical interplay of AI systems while supporting the development of policy frameworks to advance transparency and inclusivity across AI applications.

These are some potential recommendations that future research, on the topic of AI bias, could incorporate to build upon the foundation established by the current research by resolving its shortcomings and deepening its understanding of AI bias. Future research can help design AI systems that are both socially conscious and technologically sophisticated by examining these areas, which will ultimately result in more equal outcomes for all user groups.

By cultivating a more profound comprehension of these processes, this research facilitates the development of socially conscious AI systems. It backs the formulation of regulations that promote openness and justice in AI research—which is fundamental to the equitable advancement of technology.

The experience of conducting this research has been stimulating and transformative, offering invaluable insights into the topic of AI bias and my personal development as a researcher. This research process helped me gain a better grasp of the nuances of AI technologies and how they affect society. Interacting with individuals from diverse backgrounds has expanded my perspective and underlined the relevance of cultural, individual, and professional aspects playing a role in perception when conducting AI research.

One of the most important takeaways from this research is the value of interdisciplinary teamwork. By combining concepts from computer science, sociology, and psychology, I obtained a more complete understanding of AI bias and its ramifications. This multidisciplinary approach not only expanded on this research's findings but also highlighted the interdependence of numerous fields in solving complex societal concerns.

This research journey has highlighted the importance of reflexivity in the research process. As a researcher, I've constantly focused on my own biases, preconceptions, and positionality, recognizing the impact they might have on this research's results. This reflexivity has allowed me to approach the research with humility and openness, fostering a deeper appreciation for the diverse perspectives of this research's participants.

Furthermore, this research experience helped me grow personally and professionally in a variety of ways. It improved my analytical and critical thinking skills, helping me to confidently traverse complicated research topics and data. Additionally, it has improved my capacity to successfully convey research findings to varied audiences, both orally and in writing.

Overall, this research has been extremely rewarding and informative. It has fueled my interest in understanding the societal implications of developing technology and motivated me to continue researching this intriguing topic. As I end this research, I bring with me not just the knowledge gained, but also a renewed dedication to making a meaningful contribution to the progress of knowledge and the improvement of society.

As I conclude this research, I am reminded of the enormous impact that AI technologies have on our lives and the critical need to overcome the biases inherent in them. This research emphasizes the crucial relevance of supporting transparency, equity, and accountability in AI development and deployment by shedding light on the multidimensional perspectives of AI bias and its implications for people from various backgrounds. May this research be a catalyst for positive change, conversation, and the development of a future where technology supports mankind responsibly and ethically as we negotiate the difficulties of an increasingly AI-driven society. Together, let’s work to create an AI that upholds the principles of justice, fairness, and inclusivity so that technological advancements enhance rather than detract from the human experience.

References

Adamakis, E. (2024). A Comparative Analysis of the EU and US Artificial Intelligence (AI) Regulation Regimes. *Ihu.edu.gr*. https://repository.ihu.edu.gr//xmlui/handle/11544/30384

Ahmadi, A. (2023). Quantum Computing and Artificial Intelligence: The Synergy of Two Revolutionary Technologies. *Asian Journal of Electrical Sciences*, *12*(2), 15–27. https://doi.org/10.51983/ajes-2023.12.2.4118

Ajani, S. N., Khobragade, P., Dhone, M., Ganguly, B., Shelke, N., & Parati, N. (2024). Advancements in Computing: Emerging Trends in Computational Science with Next-Generation Computing. *International Journal of Intelligent Systems and Applications in Engineering*, *12*(7s), 546–559. https://ijisae.org/index.php/IJISAE/article/view/4159

Alowais, S. A., Alghamdi, S. S., Alsuhebany, N., Alqahtani, T., Abdulrahman Alshaya, Almohareb, S. N., Atheer Aldairem, Alrashed, M., Khalid Bin Saleh, Badreldin, H. A., Yami, A., Shmeylan Al Harbi, & Albekairy, A. M. (2023). Revolutionizing healthcare: the role of artificial intelligence in clinical practice. *BMC Medical Education*, *23*(1). https://doi.org/10.1186/s12909-023-04698-z

Aquino, Y. S. J., Carter, S. M., Houssami, N., Braunack-Mayer, A., Win, K. T., Degeling, C., Wang, L., & Rogers, W. A. (2023). Practical, epistemic and normative implications of algorithmic bias in healthcare artificial intelligence: a qualitative study of multidisciplinary expert perspectives. *Journal of Medical Ethics*. https://doi.org/10.1136/jme-2022-108850

Baltar, F., & Brunet, I. (2012). Social research 2.0: virtual snowball sampling method using Facebook. *Internet Research*, *22*(1), 57–74. https://doi.org/10.1108/10662241211199960

Barber, W., & Badre, A. (1998, June). Culturability: The merging of culture and usability. In *Proceedings of the 4th Conference on Human Factors and the Web* (Vol. 7, No. 4, pp. 1-10). https://shorturl.at/Q6rnY

Bhalla, R., Tiwari, P., Chowdhary, N. (2021). Digital Natives Leading the World: Paragons and Values of Generation Z. In: Stylos, N., Rahimi, R., Okumus, B., Williams, S. (eds) Generation Z Marketing and Management in Tourism and Hospitality. Palgrave Macmillan, Cham. https://doi.org/10.1007/978-3-030-70695-1_1

Bingham, A. J. (2023). From Data Management to Actionable Findings: A Five-Phase Process of Qualitative Data Analysis. *International Journal of Qualitative Methods*, *22*. https://journals.sagepub.com/doi/full/10.1177/16094069231183620

Bogojevic Arsic, V. (2021). Challenges of Financial Risk Management: AI Applications. *Management:Journal of Sustainable Business and Management Solutions in Emerging Economies*, *26*(3). https://doi.org/10.7595/management.fon.2021.0015

Bostrom, N. (2016). The Control Problem. Excerpts from*Superintelligence: Paths, Dangers, Strategies*. *Science Fiction and Philosophy: From Time Travel to Superintelligence*, 308–330. https://doi.org/10.1002/9781118922590.ch23

Braun, V., & Clarke, V. (2006). Using Thematic Analysis in Psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. Tandfonline. https://doi.org/10.1191/1478088706qp063oa

Braun, V., Clarke, V., Hayfield, N., & Terry, G. (2019). Thematic Analysis. In: Liamputtong P, ed. *Handbook of research methods in health social sciences*, 843-860. Springer. https://shorturl.at/sxTVp

Bryman, A., & Bell, E. (2015). *Business Research Methods* (4th ed.). Oxford University Press. https://shorturl.at/ziT0B

Butcher, J., & Beridze, I. (2019). What is the State of Artificial Intelligence Governance Globally? *The RUSI Journal*, *164*(5-6), 88–96. https://doi.org/10.1080/03071847.2019.1694260

Butt, J. S. (2024). Analytical Study of the World's First EU Artificial Intelligence (AI) Act, 2024. *International Journal of Research Publication and Reviews*, *5*(3), 7343–7364. https://doi.org/10.55248/gengpi.5.0324.0914

Cannarsa, M. (2021). Ethics guidelines for trustworthy AI. *The Cambridge handbook of lawyering in the digital age*, 283-97. Chicago. https://doi.org/10.1017/9781108936040

Chadwick-Dias, A., McNulty, M., & Tullis, T. (2002). Web usability and age. *ACM SIGCAPH Computers and the Physically Handicapped*, *73-74*, 30–37. https://doi.org/10.1145/960201.957212

Chen, I., Johansson, F. D., & Sontag, D. (2018). Why is my classifier discriminatory?. *Advances in neural information processing systems*, *31*. Chicago. https://proceedings.neurips.cc/paper/2018/file/1f1baa5b8edac74eb4eaa329f14a0361-Paper.pdf

Chennupati, A. (2024). The evolution of AI: What does the future hold in the next two years. *World Journal of Advanced Engineering Technology and Sciences*, *12*(1), 022–028. https://doi.org/10.30574/wjaets.2024.12.1.0176

Chu, C. H., Donato-Woodger, S., Khan, S. S., Nyrup, R., Leslie, K., Lyn, A., Shi, T., Bianchi, A., Rahimi, S. A., & Grenier, A. (2023). Age-related bias and artificial intelligence: a scoping review. *Humanities and Social Sciences Communications*, *10*(1), 1–17. https://doi.org/10.1057/s41599-023-01999-y

Chui, M., Yee, L., Hall, B., & Singla, A. (2023). The state of AI in 2023: Generative AI's breakout year. Chicago. http://dln.jaipuria.ac.in:8080/jspui/bitstream/123456789/14348/1/The-state-of-organizations-2023%20-%20McKinsey%20Report.pdf

Cloete, F. (2024). Governing Artificial Intelligence (AI) and Other Technologies in the Digital Era. *Administratio Publica*, *32*(1), 1-30.Chicago

Comunale, M., & Manera, A. (2024). The economic impacts and the regulation of AI: A review of the academic literature and policy actions. IMF Working Paper No. 2024/065. International Monetary Fund. https://www.imf.org/-/media/Files/Publications/WP/2024/English/wpiea2024065-print-pdf.ashx

Cotten, S. R., Anderson, W. A., & McCullough, B. M. (2013). Impact of Internet Use on Loneliness and Contact with Others Among Older Adults: Cross-Sectional Analysis. Journal of Medical Internet Research, 15(2), e39. https://doi.org/10.2196/jmir.2306

Crawford, K., & Schultz, J. (2014). Big data and due process: toward framework to redress predictive privacy harms. *Boston College Law Review, 55(1),* 93-128. https://lira.bc.edu/files/pdf?fileid=89d8ef18-a073-4fc4-b5d7-eec70979e081

Dastin, J. (2022). Amazon scraps secret AI recruiting tool that showed bias against women*. Ethics of Data and Analytics, 296–299. https://doi.org/10.1201/9781003278290-44

De Sutter, E., Zaçe, D., Boccia, S., Di Pietro, M. L., Geerts, D., Borry, P., & Huys, I. (2020). Implementation of Electronic Informed Consent in Biomedical Research and Stakeholders' Perspectives: Systematic Review. *Journal of Medical Internet Research*, *22*(10), e19129. https://doi.org/10.2196/19129

de-Lima-Santos, M.-F., & Ceron, W. (2021). Artificial Intelligence in News Media: Current Perceptions and Future Outlook. *Journalism and Media*, *3*(1), 13–26. https://doi.org/10.3390/journalmedia3010002

Dignum, V. (2019). *Responsible artificial intelligence: how to develop and use AI in a responsible way* (Vol. 1). Cham: Springer. Chicago. https://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=2557435

Etikan, I., Musa, S. A., & Alkassim, R. S. (2016). Comparison of Convenience Sampling and Purposive Sampling. *American Journal of Theoretical and Applied Statistics*, *5*(1), 1–4. https://doi.org/10.11648/j.ajtas.20160501.11

Farquhar, J., Michels, N., & Robson, J. (2020). Triangulation in Industrial Qualitative Case Study research: Widening the Scope. *Industrial Marketing Management*, *87*(1), 160–170. https://eprints.bournemouth.ac.uk/33433/3/triangulation%20pre%20print.pdf

Feldstein, S. (2019). *The global expansion of AI surveillance* (Vol. 17, No. 9). Washington, DC: Carnegie Endowment for International Peace. Chicago. https://blog.fdik.org/2019-09/WP-Feldstein-AISurveillance_final1.pdf

Ferrara, E. (2023). Fairness and Bias in Artificial Intelligence: A Brief Survey of Sources, Impacts, and Mitigation Strategies. *Sci*, *6*(1), 3–3. https://doi.org/10.3390/sci6010003

Ferrara, E. (2024). The Butterfly Effect in artificial intelligence systems: Implications for AI bias and fairness. *Machine Learning with Applications*, *15*, 100525. https://doi.org/10.1016/j.mlwa.2024.100525

Friedler, S. A., Scheidegger, C., & Venkatasubramanian, S. (2021). The (Im)possibility of fairness. *Communications of the ACM*, *64*(4), 136–143. https://doi.org/10.1145/3433949

Ganel, T., Sofer, C., & Goodale, M. A. (2022). Biases in human perception of facial age are present and more exaggerated in current AI technology. *Scientific Reports*, *12*(1). https://doi.org/10.1038/s41598-022-27009-w

Gelles, R., Kinoshita, V., Musser, M., & Dunham, J. (2024). Resource Democratization: Is Compute the Binding Constraint on AI Research? *Proceedings of the AAAI Conference on Artificial Intelligence*, *38*(18), 19840–19848. https://doi.org/10.1609/aaai.v38i18.29959

Gerlich, M. (2023). Perceptions and Acceptance of Artificial Intelligence: A Multi-Dimensional Study. *Social Sciences*, *12*(9), 502. https://doi.org/10.3390/socsci12090502

Gill, A. S. (2019). Artificial Intelligence and International Security: The Long View. *Ethics & International Affairs*. 33(2), 169-179. https://doi.org/10.1017/S0892679419000145

Gough, D., Oliver, S., & Thomas, J. (2017). *An introduction to systematic reviews* (2nd ed.). Sage Publications Ltd. https://www.google.nl/books/edition/An_Introduction_to_Systematic_Reviews/41sCDgAAQBAJ?hl=en&gbpv=1&dq=Gough,+D.,+Oliver,+S.,+%26+Thomas,+J.+(2017).+An+introduction+to+systematic+reviews+(2nd+edition).+SAGE.&printsec=frontcover

Guarasci, R., De Pietro, G., & Esposito, M. (2022). Quantum Natural Language Processing: Challenges and Opportunities. *Applied Sciences*, *12*(11), 5651. https://doi.org/10.3390/app12115651

Gunning, D., Stefik, M., Choi, J., Miller, T., Stumpf, S., & Yang, G.-Z. (2019). XAI—Explainable artificial intelligence. *Science Robotics*, *4*(37), eaay7120. https://doi.org/10.1126/scirobotics.aay7120

Hagendorff, T. (2020). The Ethics of AI Ethics: An Evaluation of Guidelines. *Minds and Machines*, *30*(1), 99–120. https://doi.org/10.1007/s11023-020-09517-8

Hagendorff, T., & Fabi, S. (2023). Why we need biased AI: How including cognitive biases can enhance AI systems. *Journal of Experimental & Theoretical Artificial Intelligence*, 1–14. https://doi.org/10.1080/0952813x.2023.2178517

Hagerty, A., & Rubinov, I. (2019, July 18). *Global AI Ethics: A Review of the Social Impacts and Ethical Implications of Artificial Intelligence*. ArXiv.org. https://doi.org/10.48550/arXiv.1907.07892

Hajian, S., Bonchi, F., & Castillo, C. (2016). Algorithmic Bias. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '16*. https://doi.org/10.1145/2939672.2945386

Hao, T., & Chen, X. (2024). *The Influence of AI and Recommender Systems on Media Consumption Patterns: Perception from Content Creators in Chinese Media Industry*. https://www.researchgate.net/publication/378772114_The_Influence_of_AI_and_Recommender_Systems_on_Media_Consumption_Patterns_Perception_from_Content_Creators_in_Chinese_Media_Industry

Harrell, M. C., & Bradley, M. A. (2009). *Data Collection Methods: Semi-Structured Interviews and Focus Groups*. RAND Corporation. United States. https://policycommons.net/artifacts/4829747/data-collection-methods/5666469/ on 30 Jun 2024. CID: 20.500.12592/zngrv2.

Hartzog, W. (2015). Unfair and deceptive robots. *Maryland Law Review, 74(4),* 785-832. https://heinonline.org/HOL/LandingPage?handle=hein.journals/mllr74&div=41&id=&page=

Hayden, M. A. (1997). Diversity Factors Related To Technological Literacy. *Bulletin of Science, Technology & Society*, *17*(4), 218–226. https://doi.org/10.1177/027046769701700452

Helberger, N., Diakopoulos, N. (2022). The European AI Act and How It Matters for Research into AI in Media and Journalism. *Digital Journalism*, *11*(9), 1–10. https://doi.org/10.1080/21670811.2022.2082505

Helkala, K., Lucas, G., & Barrett, E. (2023). Editorial: Ethical challenges in AI-enhanced military operations. *Frontiers in Big Data*, *6*. https://doi.org/10.3389/fdata.2023.1229252

Hofstede, G. (1980). *Culture's consequences: International Differences in work-related Values*. SAGE Publications, Newbury Park. https://books.google.nl/books?hl=en&lr=&id=Cayp_Um4O9gC&oi=fnd&pg=PA13&dq=Hofstede,+G.+(1980).+Culture%E2%80%99s+consequences:+International+Differences+in+work-related+Values.+Sage+Publications.&ots=V6CHIyQJI6&sig=1S77mwo_bLXeyYeTMjNC_xiDNYs&redir_esc=y#v=onepage&q&f=false

Holstein, K., Wortman Vaughan, J., Daumé, H., Dudik, M., & Wallach, H. (2019). Improving Fairness in Machine Learning Systems. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems - CHI '19*. https://doi.org/10.1145/3290605.3300830

Hosseini, M., & Horbach, S. P. J. M. (2023). *Fighting reviewer fatigue or amplifying bias? Considerations and recommendations for use of ChatGPT and other large language models in scholarly peer review*. *8*(1). https://doi.org/10.1186/s41073-023-00133-5

How, M.-L., & Cheah, S.-M. (2024). Forging the Future: Strategic Approaches to Quantum AI Integration for Industry Transformation. *AI*, *5*(1), 290–323. https://doi.org/10.3390/ai5010015

Hoy, M. B. (2018). Alexa, Siri, Cortana, and More: An Introduction to Voice Assistants. *Medical Reference Services Quarterly*, *37*(1), 81–88. https://doi.org/10.1080/02763869.2018.1404391

Javaid, M., Haleem, A., Singh, R. P., & Suman, R. (2021). Artificial Intelligence Applications for Industry 4.0: A Literature-Based Study. *Journal of Industrial Integration and Management*, *7*(1), 1–29. https://doi.org/10.1142/s2424862221300040

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, *1*(9), 389–399. https://www.nature.com/articles/s42256-019-0088-2

Johnson, K., Pasquale, F., & Chapman, J. (2019). Artificial intelligence, machine learning, and bias in finance: Toward responsible innovation. Fordham Law Review, 88(2), 499. Available at: https://ir.lawnet.fordham.edu/flr/vol88/iss2/5

Jones-Jang, S. M., & Park, Y. J. (2023). How do people react to AI failure? Automation bias, algorithmic aversion, and perceived controllability. *Journal of Computer-Mediated Communication*, *28*(1). https://doi.org/10.1093/jcmc/zmac029

Kakilla , C. (2021). Strengths and Weaknesses of Semi-Structured Interviews in Qualitative Research: a Critical Essay. *Strengths and Weaknesses of Semi-Structured Interviews in Qualitative Research: A Critical Essay*, *1*(1). https://doi.org/10.20944/preprints202106.0491.v1

Kallio, H., Pietilä, A.-M., Johnson, M., & Kangasniemi, M. (2016). Systematic methodological review: Developing a framework for a qualitative semi-structured interview guide. *Journal of Advanced Nursing*, *72*(12), 2954–2965. https://shorturl.at/ei6GW

Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of Social Media. *Business Horizons*, *53*(1), 59–68. https://doi.org/10.1016/j.bushor.2009.09.003

Kaplan, A., & Haenlein, M. (2019). Siri, Siri, in my hand: Who’s the fairest in the land? On the interpretations, illustrations, and implications of Artificial Intelligence. *Business Horizons*, *62*(1), 15–25. https://doi.org/10.1016/j.bushor.2018.08.004

Karatsareas, P. (2022). Semi-structured interviews. *Research methods in language attitudes*, 99-113.Chicago. https://doi.org/10.1017/9781108867788.010

Kelly, S., & Mirpourian, M. (2021). Algorithmic Bias, Financial Inclusion, and Gender. *Women’s World Banking*. https://www.womensworldbanking.org/wpcontent/uploads/2021/02/2021_Algorithmic_Bias_Report.pdf

King, D. B., O’Rourke, N., & DeLongis, A. (2014). Social media recruitment and online data collection: A beginner’s guide and best practices for accessing low-prevalence and hard-to-reach populations. *Canadian Psychology/Psychologie Canadienne*, *55*(4), 240–249. https://doi.org/10.1037/a0038087

Kizilcec, R. (2016). How Much Information? Effects of Transparency on Trust in an Algorithmic Interface. *Proceedings of the CHI Conference on Human Factors in Computing Systems*, 2390–2395. https://doi.org/10.1145/2858036.2858402

Kleinberg, J., Sendhil Mullainathan, & Raghavan, M. (2016). Inherent Trade-Offs in the Fair Determination of Risk Scores. *ArXiv (Cornell University)*. https://doi.org/10.48550/arxiv.1609.05807

Kraus, S., Jones, P., Kailer, N., Weinmann, A., Banegas, N. C., & Tierno, N. R. (2021). Digital Transformation: An Overview of the Current State of the Art of Research. *SAGE Open*, *11*(3), 215824402110475. Sagepub. https://doi.org/10.1177/21582440211047576

Lancaster, C., Schulenberg, K., Flathmann, C., McNeese, N. J., & Freeman, G. (2023). ”It’s Everybody’s Role to Speak Up... But Not Everyone Will”: Understanding AI Professionals’ Perceptions of Accountability for AI Bias Mitigation. *ACM Journal on Responsible Computing*, *1*(1), 1–30. https://doi.org/10.1145/3632121

Lee, C., & Coughlin, J. F. (2014). PERSPECTIVE: Older Adults’ Adoption of Technology: An Integrated Approach to Identifying Determinants and Barriers. Journal of Product Innovation Management, 32(5), 747–759. https://doi.org/10.1111/jpim.12176

Lee, M. K. (2018). Understanding perception of algorithmic decisions: Fairness, trust, and emotion in response to algorithmic management. *Big Data & Society*, *5*(1), 205395171875668. https://doi.org/10.1177/2053951718756684

Li, F., Ruijs, N., & Lu, Y. (2022). Ethics & AI: A Systematic Review on Ethical Concerns and Related Strategies for Designing with AI in Healthcare. *AI*, *4*(1), 28–53. https://doi.org/10.3390/ai4010003

Li, M., & Suh, A. (2022). Anthropomorphism in AI-enabled technology: A literature review. *Electronic Markets*. https://doi.org/10.1007/s12525-022-00591-7

Li, X., Hess, T. J., McNab, A. L., & Yu, Y. (2009). Culture and acceptance of global web sites. *ACM SIGMIS Database: The DATABASE for Advances in Information Systems*, *40*(4), 49–74. https://doi.org/10.1145/1644953.1644959

Long, D., & Magerko, B. (2020). What is AI Literacy? Competencies and Design Considerations. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. https://doi.org/10.1145/3313831.3376727

Lopez, V., & Whitehead, D. (2013). *(PDF) Sampling data and data collection in qualitative research*. ResearchGate.

https://www.researchgate.net/publication/255950308_Sampling_data_and_data_collection_in_qualitative_research

Luna-Reyes, L. F., & Andersen, D. L. (2003). Collecting and analyzing qualitative data for system dynamics: methods and models. *System Dynamics Review*, *19*(4), 271–296. https://doi.org/10.1002/sdr.280

Mehmood, K., Katrien Verleye, Arne De Keyser, & Lariviere, B. (2024). The transformative potential of AI-enabled personalization across cultures. *Journal of Services Marketing/□the □Journal of Services Marketing*. https://doi.org/10.1108/jsm-08-2023-0286

Min, A. (2023). Artifical Intelligence and Bias: Challenges, Implications, and Remedies. *Journal of Social Research*, *2*(11), 3808–3817. https://doi.org/10.55324/josr.v2i11.1477

Mirza, H., Mirza, C., & Bellalem, F. (2023). Ethical considerations in qualitative research: Summary guidelines for novice social science researchers. Social Studies and Research Journal, 11(1), 441-449. https://www.researchgate.net/publication/370838199_Ethical_Considerations_in_Qualitative_Research_Summary_Guidelines_for_Novice_Social_Science_Researchers

Moore, S., Bulmer, S., & Elms, J. (2022). The social significance of AI in retail on customer experience and shopping practices. *Journal of Retailing and Consumer Services*, *64*(1), 102755. https://doi.org/10.1016/j.jretconser.2021.102755

Morgan, F. E., Boudreaux, B., Lohn, A. J., Ashby, M., Curriden, C., Klima, K., & Grossman, D. (2020). Military applications of artificial intelligence. *Santa Monica: RAND Corporation*.

Motoki, F. Y. S., Neto, V. P., & Rodrigues, V. (2023). More Human than Human: Measuring ChatGPT Political Bias. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4372349

Naeem, M., Ozuem, W., Howell, K. E., & Ranfagni, S. (2023). A Step-by-Step Process of Thematic Analysis to Develop a Conceptual Model in Qualitative Research. *International Journal of Qualitative Methods*, *22*(1), 1–18. Sagepub. https://doi.org/10.1177/16094069231205789

Noy, C. (2008). Sampling Knowledge: The Hermeneutics of Snowball Sampling in Qualitative Research. *International Journal of Social Research Methodology*, *11*(4), 327–344. https://www.tandfonline.com/doi/abs/10.1080/13645570701401305

Nunkoosing, K. (2005). The Problems With Interviews. *Qualitative Health Research*, *15*(5), 698–706. https://doi.org/10.1177/1049732304273903

Nuseir, M. T. (2020). Is advertising on social media effective An empirical study on the growth of advertisements on the Big Four (Facebook, Twitter, Instagram, WhatsApp). *International Journal of Procurement Management*, *13*(1), 134. https://doi.org/10.1504/ijpm.2020.105191

Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, *366*(6464), 447–453. https://doi.org/10.1126/science.aax2342

Oladoyinbo, Tunboson Oyewale and Olabanji, Samuel Oladiipo and Olaniyi, Oluwaseun Oladeji and Adebiyi, Olubukola Omolara and Okunleye, Olalekan J. and Ismaila Alao, Adegbenga, Exploring the Challenges of Artificial Intelligence in Data Integrity and its Influence on Social Dynamics (January 13, 2024). Asian Journal of Advanced Research and Reports, Volume 18, Issue 2, Page 1-23, 2024, Available at SSRN: https://ssrn.com/abstract=4693987

Olatunji Akinrinola, Chinwe Chinazo Okoye, Onyeka Chrisanctus Ofodile, & Chinonye Esther Ugochukwu. (2024). Navigating and reviewing ethical dilemmas in AI development: Strategies for transparency, fairness, and accountability. *GSC Advanced Research and Reviews*, *18*(3), 050–058. https://doi.org/10.30574/gscarr.2024.18.3.0088

Osborne, N., & Grant-Smith, D. (2021). In-Depth Interviewing. *Methods in Urban Analysis*, 105–125. https://doi.org/10.1007/978-981-16-1677-8_7

Pagano, T. P., Loureiro, R. B., Lisboa, F. V. N., Peixoto, R. M., Guimarães, G. A. S., Cruz, G. O. R., Araujo, M. M., Santos, L. L., Cruz, M. A. S., Oliveira, E. L. S., Winkler, I., & Nascimento, E. G. S. (2023). Bias and Unfairness in Machine Learning Models: A Systematic Review on Datasets, Tools, Fairness Metrics, and Identification and Mitigation Methods. *Big Data and Cognitive Computing*, *7*(1), 15. https://doi.org/10.3390/bdcc7010015

Page, M. J., Moher, D., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., & McGuinness, L. A. (2021). PRISMA 2020 Explanation and elaboration: Updated Guidance and Exemplars for Reporting Systematic Reviews. *BMJ*, *372*(160). NCBI. https://doi.org/10.1136/bmj.n160

Parra, C. M., Gupta, M., & Dennehy, D. (2021). Likelihood of Questioning AI-based Recommendations Due to Perceived Racial/Gender Bias. *IEEE Transactions on Technology and Society*, *3*(1), 1–1. https://doi.org/10.1109/tts.2021.3120303

Rahman, S. (2020). The Advantages and Disadvantages of Using Qualitative and Quantitative Approaches and Methods in Language “Testing and Assessment” Research: a Literature Review. *Journal of Education and Learning*, *6*(1), 102–112. https://doi.org/10.5539/jel.v6n1p102

Ray, P. P. (2023). ChatGPT: a Comprehensive Review on background, applications, Key challenges, bias, ethics, Limitations and Future Scope. *Internet of Things and Cyber-Physical Systems*, *3*(1), 121–154. https://doi.org/10.1016/j.iotcps.2023.04.003

Roberts, H., Cowls, J., Hine, E., Morley, J., Wang, V., Taddeo, M., & Floridi, L. (2022). Governing artificial intelligence in China and the European Union: Comparing aims and promoting ethical outcomes. *The Information Society*, 1–19. https://doi.org/10.1080/01972243.2022.2124565

Rogelberg, S. G. (2002). *Handbook of research methods in industrial and organizational psychology*. Oxford: Blackwell Publishers.

Roshanaei, M. (2024). Towards best practices for mitigating artificial intelligence implicit bias in shaping diversity, inclusion and equity in higher education. *Education and Information Technologies*. https://doi.org/10.1007/s10639-024-12605-2

Roy, K. (2024). *Artificial Intelligence, Ethics and the Future of Warfare*. Taylor & Francis. https://doi-org.eur.idm.oclc.org/10.4324/9781003421849

Rozado, D. (2023). The Political Biases of ChatGPT. *Social Sciences*, *12*(3), 148. https://doi.org/10.3390/socsci12030148

Sadler, G. R., Lee, H.-C., Lim, R. S.-H., & Fullerton, J. (2010). Research Article: Recruitment of hard-to-reach Population Subgroups via Adaptations of the Snowball Sampling Strategy. *Nursing & Health Sciences*, *12*(3), 369–374. https://doi.org/10.1111/j.1442-2018.2010.00541.x

Sartori, L., & Theodorou, A. (2022). A sociotechnical perspective for the future of AI: narratives, inequalities, and human control. *Ethics and Information Technology*, *24*(1). https://doi.org/10.1007/s10676-022-09624-3

Schlageter, M. (2015). Culture of Responsibility: An Analysis of Corporate Social Responsibility Within Hofstede's Masculinity-Femininity Framework. *Unr.edu*. http://hdl.handle.net/11714/649

Schmidt, P., Biessmann, F., & Teubner, T. (2020). Transparency and trust in artificial intelligence systems. *Journal of Decision Systems*, *29*(4), 260–278. https://doi.org/10.1080/12460125.2020.1819094

Seemiller, C., & Grace, M. (2016). *Generation Z Goes to College.* Josey Bass, San Francisco, CA. https://www.apuaf.org/wp-content/uploads/2017/03/Book-review-Generation-Z-goes-to-college.pdf

Seemiller, C., & Grace, M. (2017). Generation Z: Educating and Engaging the next Generation of Students. *About Campus*, *22*(3), 21–26. https://doi.org/10.1002/abc.21293

Selbst, A. D., Barocas, Solon. (2023). Unfair artificial intelligence: how ftc intervention can overcome the limitations of discrimination law. *University of Pennsylvania Law Review, 171(4),* 1023-1094.

Suresh, H., & Guttag, J. (2021). A Framework for Understanding Sources of Harm throughout the Machine Learning Life Cycle. Equity and Access in Algorithms, Mechanisms, and Optimization. https://doi.org/10.1145/3465416.3483305

Swierczek, F.W., Bechter, C. (2010). Cultural Features of e-Learning. In: Spector, J., Ifenthaler, D., Isaias, P., Kinshuk, Sampson, D. (eds) Learning and Instruction in the Digital Age. Springer, Boston, MA. https://doi.org/10.1007/978-1-4419-1551-1_18

Tai, M. C.-T. (2020). The impact of artificial intelligence on human society and bioethics. *Tzu Chi Medical Journal*, *32*(4), 339–343. National Library of Medicine. https://doi.org/10.4103/tcmj.tcmj_71_20

Thilo Hagendorff, & Fabi, S. (2022). Methodological reflections for AI alignment research using human feedback. *ArXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2301.06859

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and Biases. Science, 185(4157), 1124–1131. https://doi.org/10.1126/science.185.4157.1124

Ulgen, O. (2022). AI and the Crisis of the Self: Protecting Human Dignity as Status and Respectful Treatment. In *The Frontlines of Artificial Intelligence Ethics* (pp. 9-33). Routledge. Chicago. https://shorturl.at/w8SQ3

United Nations. (2023a). *First Committee Approves New Resolution on Lethal Autonomous Weapons, as Speaker Warns 'An Algorithm Must Not Be in Full Control of Decisions Involving Killing'*. [Press release]. Retrieved from https://press.un.org/en/2023/gadis3731.doc.htm

United Nations (2023b). A/C.1/78/L.56. UN Doc N23/302/66. Retrieved from https://undocs.org/A/C.1/78/L.56

Van den Eynden, V., Corti, L., Woollard, M., Bishop, L., & Horton, L. (2011). Managing and sharing data; a best practice guide for researchers. https://repository.essex.ac.uk/2156/1/managingsharing.pdf

Venkatesh, V., & Davis, F. D. (2020). A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies. *Management Science*, *46*(2), 186–204. https://doi.org/10.1287/mnsc.46.2.186.11926

Veress, C. B. (2022). The Limits of Protection of Human Rights in Warfare Led by AI. Acta Univ. Sapientiae, Legal Studies, 11(2), 81–97. https://acta.sapientia.ro/content/docs/112_05.pdf

West, S.M., Whittaker, M. and Crawford, K. (2019). Discriminating Systems: Gender, Race and Power in AI. AI Now Institute. https://ainowinstitute.org/wp-content/uploads/2023/04/discriminatingsystems.pdf

Wolff, J. (2021). How Is Technology Changing the World, and How Should the World Change Technology? *Global Perspectives*, *2*(1). https://doi.org/10.1525/gp.2021.27353

Wolfson, S. (2018). Amazon's Alexa recorded private conversation and sent it to random contact. *The Guardian*, *24*. Chicago. https://www.theguardian.com/technology/2018/may/24/amazon-alexa-recorded-conversation

Wurtz, E. (2005). Intercultural Communication on Web sites: A Cross-Cultural Analysis of Web sites from High-Context Cultures and Low-Context Cultures. *Journal of Computer-Mediated Communication*, *11*(1), 274–299. https://doi.org/10.1111/j.1083-6101.2006.tb00313.x

Xu, A., Baysari, M. T., Stocker, S. L., Leow, L. J., Day, R. O., & Carland, J. E. (2020). Researchers' Views on, and Experiences with, the Requirement to Obtain Informed Consent in Research Involving Human participants: a Qualitative Study. *BMC Medical Ethics*, *21*(1). https://link.springer.com/article/10.1186/s12910-020-00538-7

Zajko, M. (2022). Artificial intelligence, algorithms, and social inequality: Sociological contributions to contemporary debates. *Sociology Compass*, *16*(3). https://doi.org/10.1111/soc4.12962

Zeng, J. (2022). *Artificial Intelligence with Chinese Characteristics*. Springer Nature. https://link.springer.com/content/pdf/10.1007/978-981-19-0722-7.pdf

Appendix A - Coding Tree

Themes	Sub-themes/Axial Codes	Open Codes
Bias and Fairness in Artificial Intelligence	Perceptions and Experiences of Bias	Definition and Understanding of AI Bias, Instances of AI Bias in Daily Use, Awareness of Bias Effects, Personal Encounters with AI Bias, Impact of Bias on User Trust, Discrepancies in AI responses, AI generating culturally insensitive content, Unexpected AI behavior, AI reinforcing stereotypes
	Data Integrity and Algorithmic Transparency	Concerns over data source credibility, Lack of explanation in algorithm decisions, Transparency in AI decision-making, Issues with AI understanding emotions
Integration of AI in Daily Life	AI Utilization Patterns	Using AI for everyday queries, AI assistance in academic research, Professional reliance on AI for data analysis, AI integration in routine tasks, Use in Academic Settings,

		Application in Professional Environments, Personal Use and Reliance on AI, Personal and Daily Use
Ethical and Regulatory Considerations	Ethical and Societal Considerations	AI impact on employment, Ethical dilemmas in AI, Societal changes due to AI, AI influence on human behavior, Ethical Considerations in AI Development, Transparency and Accountability in AI System
	Trust and Reliability in AI Systems	Reliability of AI in critical situations, Trust issues with AI accuracy, Dependency on AI, Personal skepticism towards AI
	Regulatory Frameworks and Governance	Calls for AI-specific legislation, Government regulation impact, Corporate accountability, Public policy and AI ethics
Cultural and Personal Dynamics	Cultural and Individual Influences	Personal anecdotes of AI misunderstandings, Cultural biases in AI, Individual differences in AI interaction, Influence of background on AI acceptance, Impact of Cultural Background on AI Perception,

		Variation in AI Usage by Geographic Location, Impact of Cultural Background on AI Interaction,
	Future Projections and Concerns	Expectations for AI Evolution, Concerns About Bias and Reliability in Future AI, Optimism and Concerns for AI Development

Appendix B - Interview Guide

Opening (flexible):

Hi, my name is Hrishita Pramanik, and I am currently conducting research for my master's thesis in Media & Business at Erasmus University Rotterdam, focusing on perceptions of bias in artificial intelligence. By the end of this interview, I hope to gain deeper insights into how different individuals perceive bias in AI.

- The interview will be around 45 to 60 minutes long and will only be used for research purposes, is this okay for you?
- Do you consent to me recording the interview for academic purposes?
- Would you allow me to identify you by your first name? Or do you want to remain anonymous?
- Have you had an opportunity to review the consent form?
- Do you have any questions so far?

At any given time, you are free to withdraw from this interview. If at any point you feel uncomfortable or do not want to answer a question, please feel free to tell me.

Demographic Questions:

Personal

1. What name would you like me to call you during this interview?
2. Can you share your age or the age group you belong to?
3. Gender
4. Where were you born, and where do you currently reside?
5. What is your current occupation or field of study?
6. How would you describe your cultural background?
7. Can you describe your level of familiarity and usage of AI technologies in daily life?
8. What has been your exposure to technology and AI systems, both professionally and personally?

Introduction to AI Bias

1. Before we delve into specific questions, could you share your understanding of AI bias? Feel free to explain in your own words.

2. Based on your initial explanation, how familiar do you feel with the concept of AI bias? (Follow-up with a basic definition of AI bias to ensure a common starting point for further discussion.)

Suresh, H., & Guttag, J. V. (2021). A Framework for Understanding Sources of Harm throughout the Machine Learning Life Cycle.

Paraphrased Definition for Interviews:

"AI bias refers to systematic and unfair discrepancies that arise in the development and deployment of artificial intelligence systems. These discrepancies can negatively impact certain groups of people, often mirroring existing prejudices in society. This form of bias can stem from various sources, including the data used to train AI models, the design of the algorithms themselves, or the contexts in which the AI is applied."

First Block: Experiences with AI

1. Could you share any personal encounters with AI where you noticed something that could be considered biased or unfair? What was your reaction?
2. Thinking about bias in AI, what signals or indicators make you recognize something as biased within AI systems?
3. Reflecting on your own background and experiences, how do you believe these aspects change your views on AI and its fairness?

Second Block: Understanding of AI Bias

1. In your opinion, how do the technical aspects of AI development, like the choice of data or algorithm design, contribute to the presence of bias in AI applications?
2. Considering societal norms and values, can you discuss how you think they might seep into AI systems and manifest as bias?
3. Have conversations about AI and its issues, including bias, come up in discussions with your peers or within your community? What perspectives were shared?
    a. In what ways have you noticed media or popular culture discussing AI bias? Do you think these discussions influence public perception?

Third Block: Trust and Ethical Considerations

1. From your perspective, what actions or measures should be taken to mitigate bias in AI, considering aspects like development, deployment, and oversight?
2. What role do you believe AI developers and companies should play in engaging with the public on issues of bias and fairness in AI systems?

    a. In your opinion, what are the most effective ways for companies to be transparent about the potential biases in their AI systems?
3. What role do you think government and regulatory bodies should play in managing and mitigating AI bias?

Fourth Block: Personal Dynamics

4. How do you weigh the benefits of AI technologies against the potential for bias and harm? Where should we draw the line?
5. How optimistic are you about the ability of technology and society to address and overcome issues of bias in AI? Why?
6. Are there any emerging technologies or approaches that you find promising?
7. Is there a particular aspect of AI bias that you think deserves more attention or research than it currently receives?
8. Reflecting on our conversation, has discussing these topics changed your perspective on AI and bias in any way?

Closing the Interview

1. Would you like to share any additional thoughts or insights about your experiences or opinions related to AI bias that we haven't covered before we conclude this interview?
2. Do you have any inquiries or clarifications regarding the topic or the interview process?
3. Shall I send you a copy of the completed research findings once it is finished?

Thank you for taking the time to participate in my research. Your contribution is greatly appreciated, and please don't hesitate to contact me if you ever need any assistance in the future.