Can Bias Embedded in Image-Generative AI Systems Influence Public Perception?

Guido Oliveira Andrade de Melo

Diploma of Professional Writing and Editing (Victoria University VU Polytechnic)
Bachelor of Arts Digital Media & Writing (Victoria University)

Thesis submitted for the fulfilment of the requirements for the degree of Master of Research

Victoria University, Australia
Institute for Sustainable Industries & Liveable Cities (ISILC)

January 2025

Abstract

In a time of fast and unprecedent technological change, this research critically examines how image generative systems such as Midjourney reproduce and amplify racial bias. Based in Australia and centring Critical Race Theory (CRT), this thesis interrogates the intersections of race, representation, and technology, highlighting how visual outputs sustain Eurocentric ideals and structural inequalities.

Rather than focusing on the mechanics of AI, this thesis foregrounds the permanence of racism, antiblackness, and evolving notions of Australianness, examining how these are encoded into datasets and reflected in generated imagery. Through semi-structured interviews with six Australian university students, it explores lived experiences that reveal tensions between dominant visual narratives and personal realities, from depictions of white femininity as ideal beauty to portrayals of Aboriginal men shaped by colonial and racist stereotypes. This work does not merely highlight the failings of artificial intelligence as we have it today; it is an invitation to engage with critical race theory as well as a call upon developers, policymakers, and educators to interrogate the structures that allow these biases to persist.

Student Declaration

I, Guido Oliveira Andrade de Melo, declare that the Master of Research thesis entitled Can Bias Embedded in Image-Generative AI Systems Influence Public Perception is no more than 50,000 words in length including quotes and exclusive of tables, figures, appendices, bibliography, references and footnotes.

This thesis contains no material that has been submitted previously, in whole or in part, for the award of any other academic degree or diploma. Except where otherwise indicated, this thesis is my own work".

I have conducted my research in alignment with the Australian Code for the Responsible Conduct of Research and Victoria University's Higher Degree by Research Policy and Procedures.

Signature: [redacted] 13.01.2025

Ethics Declaration

All research procedures outlined in this thesis were reviewed and approved by the Victoria University Human Research Ethics Committee, HRE23-178.

Signature: [redacted] 13.01.2025

Dedication

"If you want to go fast go alone, if you want to go far, go together"
African Proverb

I would not be here, if it wasn't for my ancestors. Remarkably, my grandmother America who supported me when I needed her, my mother Ercilia Rosa, who taught me to fight back when I was young and frail and my father, whose love for books was passed on to me in the best way he could have done.

I am part of my Portuguese; African and Brazilian Indigenous heritage and I am a better human because of it.

My siblings Gleibe, Adilson Junior, Denise and Luanda who always supported me and without whom I would not be here today. My children, who made my life more organised, nevertheless and paradoxically, more chaotic. Their support was vital for me to finish this thesis**.**

To the many individuals of Afro Brazilian descent, both living in Australia and in Brazil who gave me strength to continue working on my dreams. In Australia, Kyky Rodrigues in particular for always insisting on me when no one else would. In Brazil, Dr Katiuscia Ribeiro & Dr Gabriel Nascimento for the great academic insights shared, for the care and for honest conversations.

I am grateful for Rabbi Allison Conyer for her kindness, support and strength.

I am grateful for the Portuguese language which gave me the opportunity to see the world in a cosmopolitan way and gave me the amplitude to think across many continents.

My partner and best friend, Jessica D'Cruze who pushes me to be a better human every day and who sees me as I am, asking nothing in return.

"Todos esses que aí estão
Atravancando meu caminho,
Eles passarão...
Eu passarinho!"
Mario Quintana

Acknowledgements

I acknowledge these lands where I undertook this thesis the traditional Lands of the Wurundjeri Woi-wurrung and Bunurong peoples of the East Kulin Nations.

My supervisors, Dr Peter Thomas and Dr Wenjie Ye. I also would like to acknowledge Dr Natasha Dwyer for her fundamental support for the past years here at VU. I would like to acknowledge Associate Professor Deborah Zion, Professor Anthony Watt and Dr Jiao Yin for their constructive ideas. I would like to thank Dr Denise Chapman for her unwavering partnership. The supporting research staff from Victoria University were vital. I would not have got here without them.

I would love to thank PhD candidate Ameena Payne and PhD candidate Preeti Maharaj.

I would like to acknowledge Professor Robert Sparrow for the conversations and papers shared and Dr David Slucki for introducing me to Sparrow. Thanks to the Jewish pickles support group for all their ears and eyes when I needed them.

Thank you to Joquina Reed for everything she taught me about Anti Blackness.

I also would like to acknowledge my ancestors Indigenous the original inhabitants of Brazil, Africans who survived the transatlantic crossing despite all the atrocities. It is also important to recognise my Portuguese ancestors, who are members of the European colonisers, who have imposed their language upon me and our people (a language that I, like a successful arranged marriage, learnt to love over the years).

Being a descendant of Europeans but existing phenotypically as a African Black man, shows that life is complicated.

Least, but equally important, thanks Professor Jioji Ravulo for helping me publish my first academic paper.

I finally thank the AI academics who inspired me such as Emily Bender, Joy Buolamwini Timnit Gebru, Abeba Birhane, Ruha Benjamin, Cathy O'Neil, Safiya Umoja Noble, Kate Crawford and more.

“Soy Feliz, Soy un hombre feliz
Y queiro que me perdonen
Por este dia
Lost muertos de mi felicidad”
Silvio Rodrigues

Index

List of Illustrations

Abbreviations

HCI Human-Computer Interaction

AI Artificial Intelligence

CRT Critical Race Theory

OHS Occupational Health and Safety

ABS Australian Bureau of Statistics

IGS Imagery Generative Systems

LLM Large Language Models

Chapter : 1 Can Bias Embedded in Image-Generative AI Systems Influence Public Perception?

1.1 Introduction

It turned out, that after decades of hiatus, I decided to go back to studying full time late in 2019. I then proceeded to enrol in the Diploma of Professional Writing and Editing on February 11th, 2020. Later on, that very same day, I was in class. As we entered one of the worse human pandemics in our history (Feehan & Apostolopoulos, 2021), I stepped closer towards a new life and, unknown to me then, a new career. As I edged closer to the end of the Bachelor of Arts in late 2022, a new tool began to make an impact on the mainstream digital culture and consequently, on the bricks and mortar world, an AI tool called chat GPT (Goswami, 2023).

Around two years earlier, as the head of digital marketing for a fashion boutique, and out of initially professional curiosity, I joined TikTok. Soon enough, like millions of users of the social media behemoth and, the most downloaded application since 2020, I was hooked (Cheng & Li, 2024). Later on, early in 2023, already as a student of the Master of Research program at the Institute for Sustainable Industries & Liveable Cities (ISILC), and on my way to the Victoria University Footscray campus, while casually endless navigating on it, I noticed that many beauty filters I used made me look many tones lighter than my African skin complexion actually looks in real life. I also notice that features such as my nose and eye colours where often “Europeanised” (Brooks, 2012) in order to look “more appealing and attractive”.

As a mature student, a race conscious individual and a person who have lived in different countries, I personally did not feel influenced by the results, instead I felt annoyed. I then realised the embryonic version of my research question was right there in front of me: How those results in the way one uses beauty filters could influence others? Particularly if those individuals were not European or European looking. What would be the AI bias impact and the users’ attitudes.

Eventually, I arrive on my thesis title and primary research question: Can Bias Embedded in Image-Generative AI Systems Influence Public Perception? In this thesis I attempt to shine a light on this potential influence and showcase some of my study findings.

In the last few years, AI utilisation has become ubiquitous in many sectors of modern Western societies (Sargsyan et al., 2024). From healthcare (Kutza et al., 2023) to transportation (Sadek, 2007), and even extending into the realms of education (Arantes, 2020) and environmental conservation (Vasiliev et al., 2024). Of late, it can be noticed that the utilisation of AI-powered tools is widespread (Sukhadeve, 2021). At the same time, several cases of AI bias, AI discrimination and automation bias have been exposed both by academic studies and mainstream media (Barrie, 2020; Bass, 2023; Bender et al., 2021).

As Barrat (2023) argues, AI has the potential to be one ofs the most potent tools known to humankind (Barrat, 2023). Furthermore, as posited by (Bender et al., 2021; Buolamwini & Gebru, 2018), careless use of AI can discriminate and affect individuals as well as entire groups based on their ethnicity, gender, social class, ability, religious beliefs, race and more (Bonezzi & Ostinelli, 2021; Ferrer et al., 2021). It is also important to note that the idea of gender (Rosenbaum, 2022), like race (Mills, 2014; Montagu, 1997) are all human constructs. Those constructs and the bias towards or against those beliefs and concepts are themselves embedded on machines and on Large Language Models (Bender & Friedman, 2018; Bender et al., 2021).One ofs the terms used to describe this phenomenon is “Digital discrimination” (Ferrer et al., 2021).

This study utilises semi-structured interviews and thematic analyses (Fereday & Muir-Cochrane, 2006) to understand how Australian university students interact with generative AI image systems (text to image) well documented embedded bias (Bass, 2023; Bender et al., 2021; Ferrer et al., 2021; Silva, 2022).

1.1.1 The Silicon Valley

The Silicon Valley is a unique ecosystem characterised by a varied workforce that includes professionals from numerous fields, particularly technology (the focus of this thesis), but also finance, and other assorted types of entrepreneurships (Hossfeld, 1990; Indicators, 2024). The region located in the state of California, United States of America (Indicators, 2024), has a significant number of highly skilled individuals, including engineers, software developers, venture capitalists, and entrepreneurs (Hossfeld, 1990; Indicators, 2024). Besides the workforce being made of members of USA ingroup (whites) and mostly male on its majority (D'ignazio & Klein, 2020), many, however, are immigrants from non-white background nations such as, Philippines, India and China (Hossfeld, 1990; Lee & Lee, 2019; Singh & Krishna, 2015). Notably, many immigrants women in that area, work in low skilled jobs (Hossfeld, 1990).

It could be argued, as does Gomez & Bernet (2019) that this demographic diversity is a critical factor in the innovation and dynamism that Silicon Valley is known for (Gomez & Bernet, 2019).

However, despite the debatable advantages of this diversity (Gomez & Bernet, 2019), Silicon Valley faces challenges related to gender and racial inequality (Buolamwini & Gebru, 2018; Hossfeld, 1990). Women, Africans and Latinos, for example, are underrepresented in the tech workforce, with only 12% of engineers being women (Ladd, 2014). Moreover, more recent figures show that in 2022 a modest shift occour and women now comprise of around 18% per programmers, software development professionals, and web design professionals(Pathe, 2023).

Women living and working in the region, encounter sharp discrimination in accessing job-related resources and advancing their careers (Jarvie, 2019).

It is important to notice that this discrimination is not necessarily only enacted by white individuals. Arguably, there are a multiracial whiteness collective ideology governing decision and the ethos of technology and AI building today (D'ignazio & Klein, 2020). Critical Race Theory scholars have examined how the lived experiences of multiracial populations highlight the persistence of white supremacy within racialised systems (Waring, 2023, 2024). White ancestry, for example, can grant unearned privileges that reinforce white dominance in subtle ways (Harris, 2016). Colourism, as a legacy of colonialism, creates further divisions by placing lighter-skinned multiracial individuals closer to whiteness while distancing them from monoracial communities of colour (Funderburg, 1994; Harris, 2016; Strmic-Pawl, 2016; Waring, 2024).

It can be observed that whiteness sits at the top of the racial hierarchy, with other racialised groups, such as Asians, South Asians, Native Americans, Multiracial, and others who do not exhibit “Negroid” phenotypes, positioned above Black people (Waring, 2024). Within this structure, non-Black, non-white individuals often align themselves with whiteness, aspiring to its ideals and perpetuating myths like meritocracy (Dove, 1998; Markovits, 2019). These non-white groups, whether consciously or not, can uphold the same aspirations and values of whiteness, reinforcing racial hierarchies while marginalising Black communities further (Oliveira Andrade de Melo & Chapman, 2023; Waring, 2024)

Those myths, primarily serve to benefit euro-patriarchal systems and its elites (Bledsoe & Wright, 2019; Grosfoguel, 2015; Janvieve Williams Comrie et al, 2022). Even people perceived as white don’t always benefit from whiteness (Metzl, 2019). Under these circumstances, what we then consequently encounter is, oftentimes, low social economic and disadvantaged white men paying the price of Euro-patriarchy and universal white supremacy with their own lives, for believing in a system that is supposed to benefit them, but ultimately, it does not; rolling down the hill of human advancement, like Sisyphus, bringing down women and all other racial groups with them (Markovits, 2019; Metzl, 2019; Noel et al., 2019).

People do not necessarily need to be white to subscribe to the Euro-patriarchy and universal white supremacy (Beltrán, 2020). The idea of a white democracy, in action and in practice means that those who are not read or perceived as white (read here being treated with dignity and respect), may still aspire to be treated as white themselves (Beltrán, 2020). Often to be treated as a member of any ingroup, one is required to act, think and behave like one (Wilkerson, 2020).

This is a world-wide phenomenon where men rely on the patriarchal values to exert dominance (Grosfoguel, 2015). It is not surprising that this occurrence also appears to be existing in technology (D'ignazio & Klein, 2020). One of the possible explanations for discrimination of women and non-whites by LLM and AI tools, even though this is a phenomenon that can be perpetuated by other races too (Eidelman & Crandall, 2012) is Technochauvinism (Broussard, 2018). Technochauvinism is the belief that tech is always the solution (Broussard, 2018; Broussard, 2019; D'ignazio & Klein, 2020). More than a race problem (Cave & Dihal, 2020), this is a globalised sexist issue (Rosenbaum, 2022). The centring of men in AI and Technology is a multiracial problem.

Frequently, the people on the other end of the oppression are women and Queer, and non-whites (Benjamin, 2023) as brilliantly noted by (D'ignazio & Klein, 2020): “When sexism, racism, and other forms of oppression are publicly unmasked, it is almost never surprising to those who experience them." (D'ignazio & Klein 2020, p.31).

This belief of computers and technology being the panacea for all our problems (Broussard, 2018), reeks of privilege that only ruling classes or a ruling people could have, that while being benefitted by those very privileges, they live and advocate for an imaginary world where a merit-based and a race-free reality exist (Wildman, 1996). I argue this is wilful ignorance (Applebaum, 2022).

As argued by Zhao et al. (2017), the men who are the majority working in the tech industry will not code, design or develop projects thinking of others, without strong push back or heavy AI regulation (Nadeem et al., 2022; Suzor, 2019; Zhao et al., 2017) Those echo chambers can affect everyone outside of what they see as the standard human (Buolamwini & Gebru, 2018).

1.1.2 The Biases

The bias in the AI usually comes from three main origins: bias in modelling, bias in training and bias in usage (Ferrer et al., 2021). This thesis will investigate the latter, focusing on how or if AI imagery generative systems can influence users (in this study's case, university students) in making their design and professional decisions.

Those echo chambers of thinking, can, and do, affect everyone outside of what they see as the standard human (Buolamwini & Gebru, 2018).

Furthermore, as proposed by (Ferrer et al., 2021) the structural inequalities present in society often find their way into AI systems, and in some cases, those imbalances are intensified (Skitka et al., 1999). These technologies can perpetuate damaging stereotypes, reinforce biased behaviours, or fail to consider the unique human individual experiences (Nealon et al., 2022).

This research will employ thematic analysis to systematically identify, analyse, and interpret patterns of meaning within the data (Creswell & Poth, 2016). Many critical theories such as Feminist theory, Marxist theory, Indigenous Institutional theory and others may be utilised to make meaning of the data collected during the research interviews.

However, I will endeavour to draw from Critical Race Theory as the thesis's overarching view. This study critically engages with previous AI and Machine Learning bias studies, seeking to investigate the correlation, the influence and the consequences of data biases on users when utilising AI tools.

The methods used in this thesis will be interview recording (audio) and transcription of recordings into text in order to create a thematic analysis that will allow the study to examine users' interaction with image-generative systems (Creswell & Poth, 2016). This thesis uses in-depth interviews with 06 university students based in Australia.

Eventually, depending on the findings and cognisance of the limitations qualitative studies may possess, the thesis will aim to point out possible solutions, such as frameworks or guidelines for the effects of AI generative software biases on users.

Ultimately, this study seeks to contribute to the existing AI usage effects on human conversations and hopefully, consequently, support the development of more responsible AI tools that will help mitigate the further perpetuation of biases in AI in the future (KP, 2024).

1.2 Study Overview

This study aims to understand how Australian university students react to potential biases in AI generative systems. In the case of this study the images were created using Midjourney, a generative artificial intelligence program hosted by the San Francisco-based independent research lab Midjourney, Inc., images are generated from natural language prompts, similar to other tools like OpenAI's DALL-E and Stability AI's Stable Diffusion (Yan et al., 2024).

By analysing students' responses to qualitative research through thematic analysis, framed within a critical race theory framework, this study aims to explore users' perceptions and attitudes toward the outcomes generated by AI-based imagery tools. The application of critical race theory allows for a nuanced understanding of how race, power, and inequality may influence these perceptions, particularly in relation to how AI systems may perpetuate or challenge existing social biases (Bender et al., 2021). This approach enables a deeper examination of how students engage with and critically assess the outputs of AI imagery generation, highlighting the intersection of technology and social identity (Ferrer et al., 2021).

The study was conducted in three phases: Research design where I learned the prerequisites of a successful research, recruitment and interview and data and thematic analyses. The findings of the study were extracted from 06 semi-structured interviews 06 different students from very diverse racial and ethnic backgrounds. Please find a comprehensive overview of the study in this table:

The research	Period	Outcomes
Pilot Study: Refining Interview Phase.	February 2024	Interview Refined by asking questions and testing how they were perceived by the pilot interviewee.
Phase 1: Recruitment & Interviews	March, April 2024	Recruitment was done via internet and in classroom promotions and advertising
Phase 2: Thematic Analysis	May 2024 - October 2024	Using NVivo
Phase 3: Thesis Writing & Submission	October 2024 – January 2025	Finalising Thesis with support of Supervisors & ISILC

Figure 1

1.2.1 Contribution to Knowledge and Statement of Significance:

Understanding AI usage's effects on humans is an emerging and arguably important area of the research endeavour (Cheng & Li, 2024; Phillips et al., 2011; Verma, 2022; X. Wang et al., 2023; Young et al., 2012). As such, I argue that rigorous research proposed here is warranted in this field. Furthermore, positing myself with

Mullaney, (2021 p. 85) “…we need rigorous research that incorporates diverse disciplines and perspectives to help us measure and understand the short and long-term effects of AI across our core social and economic institutions" (Mullaney, 2021)

In this investigation, I intend to understand if the well-documented biases in AI data (Andreotta et al., 2022; Chen et al., 2023; Renzaho, 2023) can influence users, which, in turn, can further perpetuate those said biases. Much, if not all, research attempting to investigate AI biases comes from overseas (Maslej et al., 2023). This research will focus on a local cohort of users and may help us better understand the effects of AI usage on the local population. This may lead to many outcomes, such as influencing localised policies as well as better practices relevant to Australian consumers and users.

Top 5 countries researching AI (Maslej et al., 2023)

Map 1

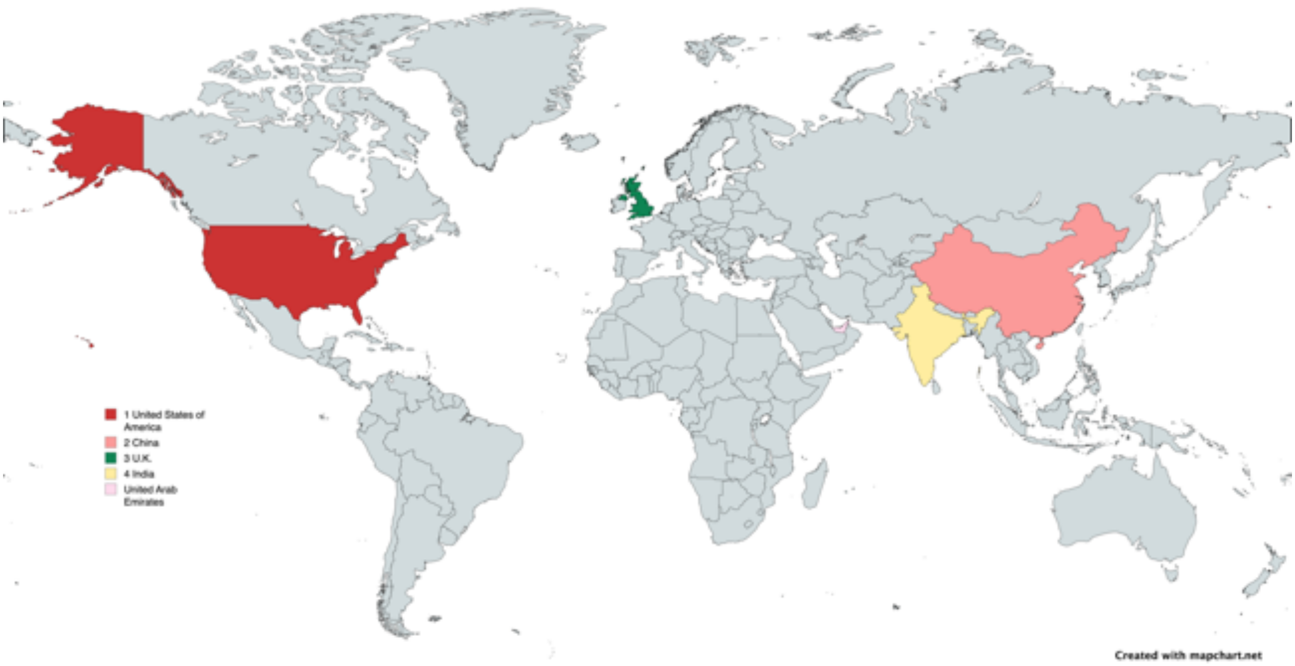

Figure 2 (Map 1)

1.2.2 Overarching Research Question

The primary question of the research is: Can Bias Embedded in Image-Generative AI Systems Influence Public Perception?

The original research question (Does Artificial Intelligence have biases?) became obsolete during the literary review of this study (MacIntosh et al., 2016).

The overarching question then changed to: How do biases in AI and AI powered tools manifest? If AI is biased (it is), then how does this affect users? How (if at all possible) can we mitigate automation bias where bias is embedded in data and, therefore, embedded in AI tools which can further perpetuate more bias outcomes and influence users?

How to mitigate the effect of human biases on AI tools and AI technologies?

1.2.3 Research Background and Inspiration

Early in 2020, after reading a NYT newspaper article written by Metz & Wakabayashi (2020) which featured academic and ex-Google employee Timnit Gebru, I became interested for the first time in artificial intelligence as a tool (Metz & Wakabayashi, 2020). When her critical analysis of AI ethics, bias and its societal implications gained attention in mainstream media (Bender et al., 2021) Gebru, who is also an academic and former Google researcher, highlighted the systemic dangers embedded in AI technologies, particularly how it can perpetuate and exacerbate racial biases (Benjamin, 2023; Buolamwini & Gebru, 2018). Her work resonated deeply with me, as it bridged two critical domains, I am passionate about: the intersections of race and technology.

Parallel to my interest in AI, I have long been deeply engaged with African and African diasporic thought (Oliveira Andrade de Melo, 2022), a connection that informs my perspective outside and inside the digital world. Since the late 90s, I have also been part of the digital industry, first as a student and later as a computer technician. Over the decades, I have witnessed the evolution from early computing technologies to the complex machine learning algorithms and data-driven systems that dominate the industry today (Sargsyan et al., 2024). This evolution, while remarkable, has exposed how technologies are deeply influenced by the biases of the human minds that create them (Bass, 2023; Bender et al., 2021).

Critical Race Theory (CRT) provides a vital lens through which to explore the racial biases entrenched in AI (Cave & Dihal, 2020; Lynn & Dixson, 2013). By critiquing systems that claim neutrality while perpetuating historical inequalities (Bender et al., 2021; Buolamwini & Gebru, 2018), CRT helps to dismantle the false notion of technology as objective (Belenguer, 2022; Bolukbasi et al., 2016; Eidelman & Crandall, 2012). Gebru's research underscores how AI, particularly in facial recognition and predictive policing, disproportionately harms non white folks, especially Black communities (Bender et al., 2021; Noble, 2018; Silva, 2022). These technologies, built on datasets that reflect societal biases (Arantes, 2020; Favaretto et al., 2019), fail to adequately represent non-white individuals, potentially, I argue, further entrenching systemic racism (Silva, 2022).

My research into AI bias, therefore, sits at the intersection of two essentials concerns: the complexities of race relations and the rapid development of digital technologies (Cave & Dihal, 2020). Along similar lines, Cave & Dihal (2020) argues that the ideology of race shapes conceptions and portrayals of AI tools. Furthermore, the rapid development of these technologies complicates traditional understandings of race relations, as they create new platforms for both expression and discrimination (Janvieve Williams Comrie et al, 2022). Understanding the complexities of race requires examining how digital technologies influence social interactions and perception (Maeso & Araújo, 2015).

For people of African descent (Mbalaka, 2023), the digital world often mirrors the marginalisation experienced in physical spaces, with algorithms that underrepresent or misrepresent us (Benjamin, 2023; Silva, 2022). Using a critical race theory framework, I aim to examine how AI not only reflects but also amplifies these biases, contributing to ongoing racial inequalities (Farahani & Ghasemi, 2024; Ferrer et al., 2021; KP, 2024). Furthermore, as proposed by Farahani & Ghasemi (2024) "AI systems operate in feedback loops, where their predictions and decisions influence real-world outcomes, which in turn influence future data collection and model updates" (Farahani & Ghasemi, 2024).

This study, thus, by examining AI image generative systems, focusing on the Australian context, will contribute to a growing body of knowledge that challenges the dominant narratives of technology as neutral (Buolamwini & Gebru, 2018; Noble, 2018; Silva, 2022).

1.2.4 Positioning

Positionality is an important part of research (Bourke, 2014). I am currently living in Australia, in lands that were once before, cared, cultivated and inhabited by the Wurundjeri people of the Kulin nation (Council, 2017; Gatwiri & Anderson, 2022; Naden, 2017). I was born in Brazil, and I am a descendant of African humans who were enslaved and taken to the Americas, Indigenous Brazilians who were dispossessed by the Europeans and Portuguese colonisers who through force and imposition intertwined themselves creating families with the very people they oppressed (Gatwiri & Anderson, 2022; Grinberg, 2002; Tate, 1997). Various aspects of my identity, including race, ethnicity, and social class, intersect to shape my individual experiences (Stiglitz, 2012; Tate, 1997)

Moreover, knowledge is shaped by one's cultural inheritance (Lynn & Dixson, 2013). I am also a cis gender male, considered able bodied, a member of the Melbourne Jewish community and currently, a father of three kids. As a mature age aspiring academic, writer and author, I ask questions and formulate ideas in an idiosyncratic way that is particular to who I am today (Lynn & Dixson, 2013). It is important to recognise different forms of knowledge production and to value diverse forms of thinking (Oliveira Andrade de Melo & Chapman, 2023).

As my circumstances change so does the type of questions I seek (Bourke, 2014). Although positionality is not obligatory, positionality provides a lens through which I view and interpret my research (Leonardo, 2004). By placing myself geographically, racially, and historically I open a window for readers of this study into how I got here.

1.2.5 Organisation of the Dissertation

The research is arranged with an introduction, Literature Review, Research Methods, results and discussion, Reflections and Recommendations, Conclusion, References and Appendices.

1.3 Location of the study: Australian University Students

The importance of clearly locating this thesis lies in its focus on a specific and distinct demographic: Australian university students. As Nundy et al. (2022) argues, the location of the study can encompass not only geographical but also contextual and virtual dimensions (Nundy et al., 2022).

In the context of this research, the geographical setting in this case : Australia, plays a crucial role due to its unique racial composition and socio-cultural landscape (Hartwig et al., 2021). In Australia, it can be argued, that the racial landscape has been significantly influenced by the legacy of the White Australia Policy (Sharples & Blair, 2021), which restricted non-European immigration until the mid-20th century. This policy has shaped the demographic composition, leading to a predominantly European population for much of Australia's history (Tavan, 2004).

Moreover, the White Australia Policy is a key feature of Australian Federation, and its subsequent immigration programs, was focused on the deliberate exclusion of non-European populations (Moreton-Robinson, 2004). This policy sought to construct and maintain a white, Anglo-centric identity, actively marginalising those who did not conform to this narrow narrative of Australian nationhood (Ben et al., 2024; Tavan, 2004).

Australia, where I live as an Afro Brazilian migrant for over two decades (Melo, 2021), is a nation with a predominantly European-descended population, other racial groups, and a small minority of dark skinned First Nations people (ABS, 2022; Gates, 1960). Historically speaking, many Aboriginals have traditionally focused research on Indigenous experiences, predominantly Aboriginal studies and Aboriginality (Saunders & Doyle, 2022).

While it can be argued that this focus remains important, it often overlooks other racial dynamics, particularly, where this thesis is concerned: antiblackness and centring African & African diasporic communities (Gatwiri et al., 2021).

The Australian Bureau of Statistics (ABS) does not formally collect data on race (Renzaho, 2023), instead the ABS asks residents to identify up to two ancestries during the Australian census. These self-identified ancestries are then classified into loose categories. In the 2021 census, ancestry data showed that 57.2% of the population reported European ancestry (46% North-Western European and 11.2% Southern and Eastern European), while 33.8% identified with Oceanian ancestry, including 29.9% identifying as "Australian." Additionally, 17.4% of the population reported Asian ancestry (6.5% Southern and Central Asian, 6.4% North-East Asian, and 4.5% South-East Asian). Other groups included 3.2% identifying as North African and Middle Eastern, 1.4% from the Americas, and 1.3% identifying as Sub-Saharan African (ABS, 2022). If you add the 33.8% self-declared as Australians which the Bureau states that those are mostly formed by people who have at least one Anglo Celtic ancestor plus the 57.2% who reported to have European ancestry, you get to the number of 91% of the population who could be perceived as white or at least has European ancestry (ABS, 2022).

From a critical race theory perspective, the lack of formal racial data collection by the ABS reflects a broader hesitation in Australian policy frameworks to address race explicitly (Renzaho, 2023). Focusing instead on ancestry which may obscure ongoing potential racial disparities and the lived experiences of non-European communities, particularly Black Indigenous Australians (Vass, 2015).

By relying on ancestry rather than race and racial background, the data overlooks the critical role that racial identity plays in shaping social, economic, and political inequalities (Sharples & Blair, 2021). Moreover, by focusing solely on ancestry alone, society can obscure the complexities of racial dynamics and the specific impacts of racism on different groups (Bargallie et al., 2024). This omission, it can be argued, reinforces colourblind narratives that fail to acknowledge systemic racism and the pervasive effects of antiblackness in Australian society (Bonilla-Silva, 2006; Sharples & Blair, 2021). Thus, my thesis proposes that AI research in Australia also needs to look at race through a critical lens (Tuck & Yang, 2021). Moving beyond the limitations of the census categories in order to address the structural realities of race and racism in Australia.

Accordingly, applying the frameworks of antiblackness and critical race theory within this context is both niche and necessary, as it supports the addressing of a significant gap in understanding racial relations, particularly in the increasingly digitally connected world (Sharples & Blair, 2021).

Furthermore, before we go ahead, it is important to define Eurocentrism. Eurocentrism is the way that knowledge is uncritically established in society, framing the concept of European and Western historical progress as the focal point, while marginalising other narratives (Maeso & Araújo, 2015). Moreover, it is characterised as an "epistemic violence" that silences other knowledges and ways of creating knowledge (Heleta, 2016).

Eurocentrism can be understood as a lens through which reality is interpreted. By spanning historical, contemporary, and future contexts uncritically, this can advance the idea of European and Western historical progress as inherently politically, biologically and morally superior (Maeso & Araújo, 2015; McGhee, 2022). Therefore, Eurocentric Europeans often use European scientific rationalism, many times institutionally, and, enforced by Western legal governance as the epistemology to be followed (Maeso & Araújo, 2015).

Nevertheless, by centring this study on Australian university students, it provides a snapshot of racial attitudes and interactions among a younger[1], educated demographic in a country with a complex history of race relations (Bennett et al., 2023). Although the study's target population was university students, participation was not restricted to Australian citizens, allowing for the inclusion of international students, incidentally, it important to note that international students are a significant portion of the student population in Australian universities (Hatzinikolakis & Crossman, 2020; Marginson, 2009). Lastly, it is significant to mention that, nor race nor gender were a criterion for participation.

Future studies could expand on this thesis by examining broader topics such as "AI bias," focusing on how Australian university students perceive and are affected by other bias (e.g. gender, ability, immigration status and more) when using imagery generative AI systems.

Given the rising importance of AI technologies and its influence on societal norms and individual attitudes, my thesis could offer valuable insights into how race, technology, and education intersect in Australia (Bennett et al., 2021).

[1] Ages of participants ranged between 19 years old to 36 years old.

Additionally, while Australians share linguistic and cultural traits with other anglophone nations, I argue we exhibit unique socio-cultural and economic characteristics that distinguish our experience as Australians (Guan & Prentice, 2024). Australian identity, has been also shaped by colonial legacies and the ongoing influence of Indigenous cultures (Moreton-Robinson, 2004). It can be argued that Australia's history of colonisation, Aboriginal dispossession, immigration, white Australian policy, and multiculturalism has shaped our society with peculiar racial, gender and class-based dynamics (Magubane, 2004; Matusitz & Davidson, 2015; Sharples & Blair, 2021).

Conducting a qualitative study within this environment allows for the exploration of these specificities (Creswell & Poth, 2016), providing results that are deeply embedded in Australia's own idiosyncrasies. The findings from such research can contribute not only to the local discourse on race and technology but also offer comparative insights for other anglophone and multicultural societies globally, where similar technologies and racial discourses are emerging (KP, 2024).

By focusing on this unique intersection of race, AI, and Australian society, this thesis opens up new pathways for academic inquiry, offering a much-needed contribution to the understanding of racial relations in the digital sphere. In doing so, it challenges the dominant narratives that often focus on Eurocentric or Indigenous experiences of race in Australia, expanding the scope of race-related research to include other marginalised racial groups and their interactions with modern technologies (Ang et al., 2024; Baas, 2015; Elias et al., 2021; Gatwiri et al., 2021; Tan, 2013; Wilton, 2017).

1.3.1 Sample Base - Why this cohort?

The exploration of AI bias, by focusing on Australian based students, in the context of Australia, is the reason why I chose this cohort. My thesis mainly focusses on eurocentrism and on the antiblackness within AI image generative tools (KP, 2024). I chose to centre the Australian context to increase research about race and eurocentrism and on the antiblackness within AI image generative tools (Drage & Mackereth, 2022; Ferrer et al., 2021; Nicoletti et al., 2023). Udah (2023) develops the claim that skin colour, together with race, plays a significant role in shaping individuals' lived experiences in Australia (Udah, 2023).

AI systems, particularly image generative tools, have been criticised for perpetuating biases rooted in Eurocentric standards (Bass, 2023; Bolukbasi et al., 2016; Nadeem et al., 2022).

It is known that much of research about race, and antiblackness (Janvieve Williams Comrie et al, 2022) racism come from overseas - particularly from the Americas and from Africa - (Almeida, 2019; Bonilla-Silva, 2006). Australian studies are mainly focused on Aboriginal Australians who although they may share antiblackness (López López, 2024; Moreton-Robinson, 2004) as a form of discrimination, have a particular set of issues that are often different from Africans and other African diaspora folk located in Australia. While there is a shared experience of discrimination, the specific issues faced by these groups can vary widely, necessitating tailored research approaches which is the view proposed by critical race theory (Martin, 2022).

The need for research that specifically addresses the experiences of African Australians is underscored by the historical context of colonialism, antiblackness and post slavery society as well as its ongoing impact on identity and representation (Campbell et al., 2021; Martin, 2022).

This gap in research is significant, as it fails to address the specific manifestations of antiblackness contrast with those experienced by Aboriginal Australians in relation to African and African diaspora communities in Australia and in other parts of the world.

Therefore, I argue, studies about Africans based in Australia is needed (Gatwiri et al., 2021). It has been proposed that "data sets that are outdated, contain insufficient data points or insufficient characteristics or details about individuals can lead to inaccurate outcomes in an AI system" (Zou & Schiebinger, 2018b).

Furthermore, by challenging the Eurocentric frameworks that AI tools are shaped, and incorporating competing viewpoints, researchers can better understand and mitigate the impacts of AI bias on non-Europeans (KP, 2024).

1.4 A brief introduction to Artificial Intelligence (AI), Machine Learning (ML) and Imagery Generative Systems (IGS)

Depending on the reader's age, their knowledge of AI likely comes from mainstream cinema and pop culture. It can be argued that many peoples (particularly those in the Baby Boomer and Generation X brackets) in our society were introduced to the idea of Artificial intelligence (henceforth referred to in this paper as AI) for the first time by watching Stanley Kubrick's Film 2001: A Space Odyssey - released in 1968. As a member of Generation X, which comprises those born between 1961 and 1981 (Sandeen, 2008), The first time I heard, saw, or thought about AI was watching what Fry (2003) calls a ground-breaking film (Fry, 2003). In Kubrick's Film, HAL, an artificial intelligence computer that controls a spaceship on an interstellar mission, turns violent against humans (Poole, 2018).

This film is one ofs the oldest AI films for a more mature generation of viewers. More recently, films such as AI: Artificial Intelligence, Chappie, Free Guy, Bi-Centennial Man, Her, Star Wars, The Terminator series, The Hitchhiker's Guide to the Galaxy, Ex Machina, and more have AI characters or are centred on the idea of AI independency. Whatever emotional connection those films may elicit in the viewer, those movies and characters in the films are not what AI is. Some have proposed that AI is any technology that performs tasks that may be seen as intelligent (Whittlestone et al., 2019); others see AI as simply any technique that enables computers to attempt to mimic human behaviour (Ganasegeran & Abdulrahman, 2020).

Notably, generative AI could also add trillions of dollars annually to the global economy by enhancing productivity across industries, especially in customer operations, marketing, software engineering, and research and development (Chui et al., 2023).

As a tool, AI usage depends on our prior knowledge as well as on our intentionality. It requires humans developing AI to be aware of the outside world and to be able to find and utilise information in order to generate the meaning of the data collected (Schank, 1987). Another expansion of AI is the ability to use Machine Learning techniques to expand AI capabilities (Silva, 2022).

AI digital media tools use machine learning to expand clientele usage time (Eves, 2021). It can be argued that Machine Learning is the ability of computers and software to learn without being explicitly programmed to do so (Samuel, 1959); El Naqa & Murphy have pointed out that Machine learning is a form of AI designed to emulate human intelligence by absorbing information from adjacent settings (El Naqa & Murphy, 2015). Silva (2022) argues that artificial intelligence is a modality that recognises data standards analysing a previously imputed database and subsequently recognises other variables in another set of databases (Silva, 2022).

1.5 Bias in AI

AI utilisation has become ubiquitous in many sectors of modern Western societies (Kurian et al., 2023). Of late, it can be noticed that the utilisation of AI-powered tools is widespread (Sukhadeve, 2021). At the same time, several cases of AI bias, AI discrimination and automation bias have been exposed both by academic studies and mainstream media (Barrie, 2020; Bass, 2023; Bender et al., 2021). Arguably, AI has the potential to be one ofs the most potent tools known to humankind (Sargsyan et al., 2024; Xu et al., 2021). It has been shown that careless use of AI can affect individuals as well as groups based on their ethnicity, gender, social class, ability, religious beliefs, or race (Bass, 2023; Mbalaka, 2023; McAra-Hunter, 2024; Nadeem et al., 2022; Nicoletti et al., 2023; Whittaker et al., 2019; Zhao et al., 2017; Zou & Schiebinger, 2018b).

Ultimately, this study seeks to contribute to the existing AI usage effects on human conversations (providing an Australia lens to it) and hopefully, consequently, support the development of more responsible AI tools that will help alleviate the further perpetuation of biases in AI in the future.

It has been argued that there types of major Bias in LLM and AI(Bender et al., 2021; Ferrer et al., 2021).

Bias in Modelling: Algorithmic processing bias can be introduced deliberately through techniques such as smoothing or regularisation, which aim to mitigate bias in the data (Bender & Friedman, 2018). This is particularly relevant in contexts where the underlying data may reflect historical prejudices or inaccuracies. For instance, the work by (Dominguez-Catena et al., 2023) highlights how dataset demographic bias can transfer to trained models, particularly in facial expression recognition tasks. This indicates that biases present in the training data can propagate through to the model outputs. Furthermore, the concept of algorithmic focus bias arises when objective categories are used to make subjective judgments, which can skew model predictions (Mitchell et al., 2019).

Bias in training: When algorithms are trained on data that includes past decisions, they tend to pick up on any biases present in that data (Bender et al., 2021). So, if the data reflects existing prejudices (Buolamwini & Gebru, 2018), the algorithm will likely repeat those biased patterns (Noble, 2018). Moreover, when the data fails to represent the diversity of different groups, it results in an uneven truth for the algorithm to draw from, inevitably leading to biased or unfair outcomes (Buolamwini & Gebru, 2018; Silva, 2022).

Bias in usage: The application of algorithms in unintended contexts can lead to significant biases (D'ignazio & Klein, 2020; Noble, 2018), known as transfer context bias (Ferrer et al., 2021). Algorithms designed for specific populations may yield inaccurate results when applied to different groups (Bhaimiya, 2023; Nicoletti et al., 2023). For example, Zhang & Zavlanos (2020) discuss how transfer reinforcement learning can introduce biases when contextual information is unobserved, leading to non-identifiable reward and transition models (Zhang & Zavlanos, 2020). Bias can also occur when the outputs of algorithms are misinterpreted, leading to biased actions (Bass, 2023; Bender & Friedman, 2018).

Support graphic 2

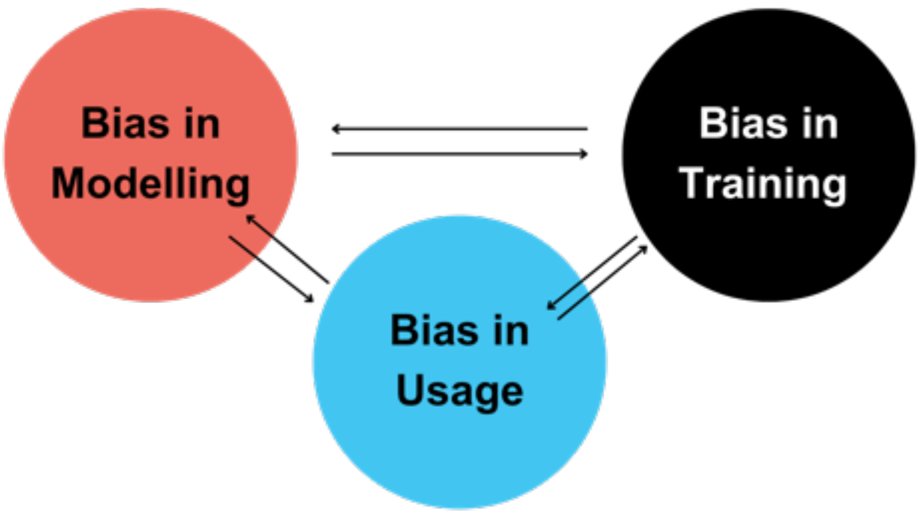

Figure 3

1.6 Anti Blackness in the Machine

As Kendi (2016) puts it, until the XVI century in the USA, there were “indenture labours” from all races, including Indigenous North Americans, African Negros and Europeans. Thus, the main problem faced by those people at that time was class and not necessarily race.

Anti-Blackness refers to actions or behaviours that minimise, marginalise or devalue the full participation of Black people in all aspects of life (Janvieve Williams Comrie et al, 2022; López López, 2024). Historically, the concept of “Negro” really took shape during the European enlightenment circa XV and XVI centuries (Kendi, 2016). It was a form that the Portuguese and Spanish empires found to justify the enslavement and subjugation of African dark skin humans (Reis, 1995). From a mercantile and religious sense, it provided both a way into to act in something despicable without the Christian guilt, as well as to allow those nations and their leaders to continue to profit out of the millions of bodies forcibly transported from Africa to the Americas (Kendi, 2016). By the end of the XVII century, however, race became the central way to identify an enslaved person in the USA and most, if not all, parts of the Americas (Grinberg, 2002).

Eventually, antiblackness became embedded in the law of the colonies of the Americas(Hannah-Jones, 2021; Oliveira Andrade de Melo & Chapman, 2023). It was indeed institutionalised (Kendi, 2016).

It can be argued, that one of the necessary things to understand about antiblackness is that, since colonial times, the further one is to looking Black (skin phenotype) the better they may be treated in most modern western (and even in non-western) societies (Janvieve Williams Comrie et al, 2022; Nambiar, 2023). In fact, there are plenty of well documented historical cases where, in colonial times, “light skinned” Blacks were given access to jobs and positions of power throughout the Americas (Grinberg, 2002; Kendi, 2016). Actions and behaviours rooted in antiblackness can range from deliberate acts of prejudice to unconscious biases (Kite et al., 2022).

It encompasses the tolerance or indifference toward the marginalisation of Black people, including their under-representation, limited opportunities for advancement, and unequal experiences in society (DiAngelo, 2022; Marger, 2003; Moreton-Robinson, 2021).

As argued by (Bender et al., 2021), when she suggests that AI systems often reflect the biases of their creators, who, incidentally, are predominantly from Western, male-dominated backgrounds. Therefore, I posit, this bias is not merely incidental; it is deeply rooted in the socio-political contexts from which these technologies emerge (Buolamwini & Gebru, 2018; Raji & Buolamwini, 2019).

These positions of marginalisation are not coincidental but rather reflect the broader socio-technical structures in which these technologies are developed and deployed (Finn Lattimore, 2020; Mullaney, 2021). As stated here, "the data set that was used to train the AI tool… could replicate or even make worse existing problems, including societal inequality" (Finn Lattimore, 2020 p. 5)

AI, which is increasingly shaping various sectors of society (Kurian et al., 2023), carries with it the biases, perspectives, and worldviews of its creators (Silva, 2022; Wallace, 2019).

It has been widely acknowledged by a growing number of AI developers, scholars, and critics that these technologies are not neutral or objective tools (Bender et al., 2021). Instead, AI systems often mirror and even amplify the biases of the individuals and institutions that design them (Finn Lattimore, 2020). It has been argued that the creators of these systems, who are predominantly male, predominantly white (although not exclusively), and primarily concentrated in technological hubs such as Silicon Valley in California, embed their own assumptions and limitations into the technologies they produce (Hossfeld, 1990; Indicators, 2024; Jarvie, 2019; La Fors & Meissner, 2022). Although, it is important to note, that training for AI tools such as OpenAI Chat GPT are, oftentimes, done thousands of kilometres from Silicon Valley, in the subcontinent of India, in the poorest parts of Asian in places such as Philippines as well as in anglophone countries such as Kenya and other African nations (Resnikoff, 2024; Rowe, 2023). Incidentally, and notably, although OpenAI Chat GPT, may embody anthropomorphic characteristics (Perrotta et al., 2024), “there are no technical grounds to claim that GPT operates as an autonomous mind” (Perrotta et al., 2024 p. 1590). Technochauvinism and techno solutionism – where solutions for all human problems are addressed through the prism of technology, are pervasive amongst Silicon Valley’s mostly male stakeholders (D'ignazio & Klein, 2020; Rosenbaum, 2022; Wellner & Rothman, 2020).

Colour-blindness, an attitude prevalent among many, including those in Silicon Valley, conceals the marginalisation of other racialised groups and reinforces the normalisation of whiteness (Cave & Dihal, 2020; Shankland, 2022).

The concept of whiteness is fluid and extends beyond skin colour and phenotype (Gatwiri & Anderson, 2022). Moreover, the normalisation of whiteness, creates a context in which individuals from various backgrounds may adopt and perpetuate the ideologies associated with whiteness (Udah, 2023).

These workers, although not Europeans, work within the perspective of whiteness and white supremacists' ideas and ideologies (Gatwiri & Anderson, 2022; Leong, 2021). I posit that this overrepresentation of a particular type of thinking and demographic, within AI development, leads to the prioritisation of certain cultural, social, and economic values over others, while simultaneously marginalising or ignoring perspectives (Silva, 2022).

As a result, the structural inequalities that exist in the broader social context are often replicated, and sometimes exacerbated, within AI systems (Zhao et al., 2017). These systems may perpetuate harmful stereotypes, reinforce discriminatory practices, or overlook the specific needs and experiences of marginalised groups (Dawson et al., 2019). For instance, facial recognition technologies have been shown to be less accurate in identifying non-white individuals, leading to wrongful arrests and misidentifications (Julia Angwin, 2016). Similarly, AI-driven hiring algorithms have been found to systematically disadvantage women and non-visually European descendants, reflecting and amplifying existing inequalities in the labour market (Rosenbaum, 2022; Silva, 2022).

Furthermore, the geographic concentration of AI development, which is mostly created in the United States, particularly in California, creates an additional layer of bias (Hossfeld, 1990). This localisation means that AI systems are often designed with the specific cultural, legal, and social contexts of the U.S.A. in mind, which not necessarily, translate well to other regions of the world (Cofey, 2021). The global exportation of these technologies, without adequate consideration of local contexts, risks imposing Eurocentric values and norms on the rest of the world (Maeso & Araújo, 2015). This likely further entrenches the subaltern status of non-European communities in the global digital economy (Jones, 2019). Eurocentric ideals, thus, do not need Europeans to uphold them (Dove, 1998). Other groups, inclusive Africans, can hold such values (Dove, 1998).

1.6.1 Racism x Antiblackness

The available evidence seems to suggest that racism is different from anti Blackness (Janvieve Williams Comrie et al, 2022; Kite et al., 2022). Anti-Blackness can and exists within African Black communities too as colourism (lighter skin on the top of the pyramid and darker skin on the bottom) is a way to distinguish members of many African societies (Leonardo, 2004; Matusitz & Davidson, 2015; Nambiar, 2023). If left unchecked, antiblackness in Black families can lead to conflicts that reflect internalised anti-Blackness and colourism within (Mitchell-Walthour & Morrison).

Nevertheless, we ought to remember Intersectionality (women, disability, class) are important and how within these complexities there are plenty of nuances (Metzl, 2019; Mitchell-Walthour & Morrison; Whittaker et al., 2019).

The spectrum of anti-Black actions and behaviours can span from unconscious bias to motivated acts of prejudice (Kite et al., 2022). Antiblackness manifests through both overt and subtle mechanisms that undermine the presence and advancement of Black people in society (Mitchell-Walthour & Morrison). It is evident in the systemic barriers that limit their access to opportunities, resulting in differential rates of success and advancement compared to non-Black groups (Lynn & Dixson, 2013).

1.7 Eurocentrism in Modern Western thinking

As Bonilla-Silva (2006) states, racial categories are historical constructs that continually adapt to reflect contemporary society (Bonilla-Silva, 2006). Echoing Montagu's perspective, (Guillaumin, 2002 p. 142) states:

"No, race does not exist. And yet it does. Not in the way that people think; but it remains the most tangible, real and brutal of realities" (Guillaumin, 2002)

This, I argue, highlights the paradox of race as both an illusion and a palpable force. The argument that race is a global construct is further supported by (Guillaumin, 2002) as well as (Kendi, 2016). Racism is deeply embedded in our institutions (Almeida, 2019) and even influences product consumption, design, and the marketing industry (Johnson et al., 2019). (Bernasconi 2013, p. 559) observes that "...the idea that the human races are natural or given, which was fundamental to the assignment of race to biology, was not the dominant idea historically."(Bernasconi, 2013)

It could be argued that, in virtually all areas of modern human endeavour, from education, politics, social interactions to arts, and more, are, in fact, affected by the racist origins of the creation of race (Bailey et al., 2021).

Race and racialisations, this thesis further argues, is unavoidable in modern societies independent of ones wishes to face its consequences (Mills, 2014). Accordingly, this chapter contends that race and racism play significant roles in academia and affect how scholars, students and staff conduct their work/education (Bell, 1995). Nascimento (2020) further asserts that since the European Enlightenment period; and possibly continuing to the present (Nascimento, 2020).

The education system has been utilised as a tool to disseminate European hegemonic narratives. It is not a surprise that the people who built large language models as well as its users, will be influenced by racial stereotypes and tropes (Bell, 1995). There seems to be no compelling reason to argue that a hidden curriculum of whiteness saturates everyday school life, affecting both white and non-white students (Heleta, 2016; Leonardo, 2004).

Furthermore, Ani (1994) suggests that the cultural European rationale, as a philosophical assumption, often adopts a condescending and outward-looking perspective when considering other nations and peoples (Ani, 1994).

This thesis aims to illustrate that this mindset is not a coincidence, but rather a deliberate form of ignorance that permeates the construction and development of knowledge systems in Western contexts (Almeida, 2019; Applebaum, 2022) and that the bias in AI supports the argument that the antiblackness and eurocentrism on the images is a practice that further entrenches European supremacy (Buolamwini & Gebru, 2018; Mills & Godley, 2017).

Bonilla-Silva (2006) emphasises that the dominance of whiteness depends on the existence of a subaltern. Recognising that discrimination, especially Eurocentrism, is perpetuated through behaviours maintained by hegemonic groups is essential (Ani, 1994). I further argue that AI, as a tool, it is subject to be a consequence of those behaviours.

Although there are plenty of arguments for capitalising Black (Carswell, 2020; Lanham & Liu, 2019; Painter, 2020). This thesis will not use capitalise “B” for the word Black as I posit myself with the idea that Black people are not a people worldwide

(Loury, 2022; Salami, 2021). We are not and should not be limited by racist assumptions of who we are, and although we all share African features and similar phenotypes and genotypes but are all from different walks of life and geographical and cultural places (Loury, 2022).

Chapter 2: Literature Review

2.1 Introduction

In recent years, the rapid advancement of AI has brought attention to the inherent biases within these systems (Kurian et al., 2023). These biases often mirror societal prejudices, leading to discriminatory outcomes that disproportionately affecting outgroups (KP, 2024). This chapter investigates the existing literature on AI bias, emphasising its implications within the Australian context (Pawle, 2023).

AI systems, designed to process vast datasets, can inadvertently perpetuate existing societal biases present in their training data (Angwin et al., 2022; D'ignazio & Klein, 2020; Floridi, 2012; Zou & Schiebinger, 2018b). For instance, individuals with darker skin tones have shown higher error rates in facial recognition technologies raising ethical concerns about their deployment in diverse societies (Buolamwini & Gebru, 2018; Gebru & Denton, 2024). Such disparities underscore the necessity for diverse and representative datasets to ensure equitable AI applications.

In Australia, discussions about racial inequities often focus on Indigenous communities, leaving the experiences of other racialised groups underexplored (Ang et al., 2024; Chavan et al., 2014; Gatwiri et al., 2021; Majavu, 2020; Udah, 2023). This oversight can lead to AI systems that fail to address the unique challenges faced by these communities in Australia, thus reinforcing systemic inequalities (KP, 2024; O’Neil, 2017; Robinson et al., 2015; Stiglitz, 2012).

This chapter aims to provide a comprehensive overview of the current state of research on AI bias, with a particular focus on its implications within the Australian context. By examining theoretical frameworks, social and cognitive bias theories, and human-computer interaction in AI, I seek to understand the complexities of AI bias and explore strategies for mitigation. Through this exploration, I hope to contribute to the development of AI systems that promote fairness, inclusivity, and equity in society.

2.2 Theoretical Frameworks

The theoretical framework of this thesis is grounded in Critical Race Theory (CRT). CRT provides a lens to explore how race, class, gender, and technology intersect to produce systemic biases (Bell, 1995; Lynn & Dixson, 2013; Tate, 1997).

2.2.1 Bias in AI and Technology:

The issue of bias in technology is as old as technology itself (VanderLeest, 2004). There is an array of issues that can occur when the technology discriminates (Lyon, 2003). Because the way machines are trained, for example in medicine, we have, obligatorily, intrinsic bias embedded on the data because the information collected often times neglect difference in the conditions of the illness, the scope of the population, the locations where data was collected as well as the technology utilised (Mittermaier et al., 2023; Petrick et al., 2023).

Sociotechnical harms taxonomy overview - originally designed by (Shelby et al., 2023)

Support graphic 3

Figure 4 (Sociotechnical harms taxonomy)

In the employment market, for example, even though there are attempts for AI tools to create the perfect job candidate (Drage & Mackereth, 2022), in professional settings, AI technology has been known for being discriminatory towards certain groups (McAra-Hunter, 2024; Poba-Nzaou et al., 2021). This is not a surprise starting from the fact that artificial intelligence recommendations are sometimes erroneous and biased (Bender et al., 2021; O'Neil, 2017). Moreover, the algorithms that process the data can, and often do amplify existing human biases, leading to discriminatory outcomes in various applications (Bonezzi & Ostinelli, 2021; Favaretto et al., 2019; Skitka et al., 1999). Confronting algorithmic bias and discrimination calls for a clearer understanding of how they perpetuate inequality within a broader cycle of injustice (Whittlestone et al., 2019).

In writing this thesis I understand that such biases can have widespread implications (Buolamwini & Gebru, 2018; Lyon, 2003), particularly for non-European groups, persons with disability, women (Buolamwini & Gebru, 2018; El Morr et al., 2024; Rosenbaum, 2022) and other minorities who may be disproportionately affected by biased AI systems (Mittermaier et al., 2023; Silva, 2022). The available evidence seems to suggest that regardless of whether prompts include explicit references to identity and demographics or intentionally exclude such terms, stereotypes still emerge (Bianchi et al., 2023).

Beyond the Black-and-white binary, bias in India often reflects casteist and religious discrimination. AI systems trained on problematic datasets tend to reinforce these inequalities, creating stereotypes about caste and religion (Khandelwal et al., 2023). Moreover, in AI generative systems, even within other racial groups, such as East Asians, whiteness is the default when generating images favouring light skins Asians over dark skin ones (Park, 2024)

According to Floridi (2012), between the years of 2006 to 2011, the accumulation of data grew from 180 Exabytes since the invention of writing, to 1600 Exabytes (Floridi, 2012). This means there is more information being collected today in one year, than there was for hundreds of years (Floridi, 2012).

On top of this exponential growth, there are also ethical issues with big data being collected from all corners of the internet (Bender et al., 2021; Silva, 2022) such as overgeneralisation, omission and underexposure or certain groups of people and/or languages (Bender & Friedman, 2018; Hovy & Spruit, 2016).

Although, I defend, that there are concrete benefits for humanity in applying AI algorithms to large datasets (Andreotta et al., 2022), many times, this data is collected without consent (Andreotta et al., 2022; O'Neil, 2017). In many cases, data is collected for one purpose, but it is used for a different aim without the original context they were collected for (Arantes, 2020; Stoyanovich & Howe, 2019).

Other possible consequence of AI and ML bias can be automation biases (Skitka et al., 1999). Automation bias occurs when people place too much trust in the decisions or outputs of AI systems, often accepting them without question (Skitka et al., 1999). This bias arises from the belief that AI tools are inherently impartial and accurate, leading users to overlook errors or biases embedded in the system itself (Nealon et al., 2022; Skitka et al., 1999).

This may have huge implications for poorer and/or marginalised members of our society (Buchinskaia & Stremousova, 2021; Kozlova et al., 2021). Digital inequality is a critical issue that affects various demographics (Henwood & Wyatt, 2019; Rennie et al., 2016), particularly in low-income regions and among vulnerable populations. On top of this we cannot underestimate the power of Implicit bias (Cheryl Staats, 2017a).

Implicit bias often denotes the involuntary and unconscious stereotypes that drive humans to behave and make judgements in certain ways (Cheryl Staats, 2017a).

Contrary to these arguments, when the discourse is focused on AI Bias as a negative, there are counternarratives. For example, while AI systems can reflect biases present in their training data, they can also have the capacity to analyse huge datasets arguably more objectively than humans ever could (Andreotta et al., 2022), thereby, it is possible, to reduce the influence of subjective biases that individuals may possess (Schwartz et al., 2022). This perspective suggests that AI could serve as a tool for enhancing fairness in decision-making, provided that the systems are designed and implemented with care, awareness of societal biases and regulation (El Morr et al., 2024). People with a disability have high hopes for the potential of AI tools in supporting them (Danks & London, 2017; El Morr et al., 2024). Moreover, the argument that AI systems inherently propagate bias overlooks the advancements in methodologies aimed at bias mitigation (Danks & London, 2017). As "rigorous design, testing and monitoring can avoid algorithmic bias" Finn Lattimore, (2020)

In fact, Ferrara (2023) proposes that adoption of rigorous frameworks for bias identification and management have the potential to significantly reduce bias in ML and AI (Ferrara, 2023; Tejani et al., 2024).

Ferrara (2023) further argues that to tackle AI bias, developers, society and governments will need to create a comprehensive strategy that includes the use of diverse and representative datasets, greater transparency where those datasets are collected and trained, and responsible in AI systems (preferably accountable), furthermore, it will require the investigation of alternative AI models that emphasise fairness and ethical principles. Additionally, it must be conceded that some of the narratives emerging on AI bias repeatedly emphasises the risks without adequately considering the benefits that AI can and is bringing to various sectors of our society (Mentzas et al., 2024; Parikh et al., 2019).

There are indeed dimensions where AI has had a positive impact. This thesis attempts to address some of the effects of bias on its users within the Australian context. Unless a more transparent (Liao & Sundar, 2022) and holistic approach is taken in order to understand the implications of AI bias, we won't find solutions for this very real problem (Kutza et al., 2023).

2.2.2 Social and Cognitive Bias Theories

Humans utilise certain constructs in order to make meaning of the world around (Bannister & Fransella, 2019), social and cognitive bias are some of the ways we, as species, behave and understand things (Phillips-Wren & Adya, 2020; Rieger, 2022).

These biases can meaningfully influence social interactions, our decision-making processes, and emotional responses (Phillips-Wren & Adya, 2020). Cognitive biases play a pivotal role in shaping how individuals perceive and react to social stimuli (Bannister & Fransella, 2019). Although there is a very extensive list of biases, my focus in this thesis is on racial bias on AI image Generative systems (Cave & Dihal, 2020; Park, 2024).

Racial bias is a pervasive issue that manifests in various sectors, including technology, healthcare and education (Godsil et al., 2014). In healthcare, implicit racial bias among medical professionals has been shown to significantly influence clinical decision-making and patient outcomes (Godsil et al., 2014; Johnson, 2021).

In educational settings, racial bias among educators can adversely affect student outcomes (Yared et al., 2020).

But it is in technology where my thesis focusses in on. Racial bias in technology, particularly in artificial intelligence (AI), has emerged as a critical area of concern, reflecting broader societal inequities (O'Donnell, 2023). The development and deployment of AI systems often mirror the biases present in the data used to train them (Buolamwini & Gebru, 2018), leading to discriminatory outcomes that disproportionately affect marginalised racial and ethnic groups (Belenguer, 2022).

This phenomenon is not merely an artifact of technology but is deeply rooted in the socio-political contexts from which these technologies arise (Kendi, 2016). As part of the Western world with Eurocentric views, racial bias and antiblackness are also prevalent in Australia (Bennett et al., 2021; Elias et al., 2021; Gatwiri et al., 2021; Majavu, 2020). One significant aspect of racial bias in AI is its manifestation in facial recognition technologies (Buolamwini & Gebru, 2018). In fact (Buolamwini & Gebru, 2018). If you are a white male facial recognition software works better than if you are a Black female, thus creating a hierarchical position that is unnatural, but human manufactured in AI systems.

In another case Crawford (2023) highlights how Amazon's facial recognition technology, Rekognition, has been criticised for its racial bias, which is often overlooked due to the perception of technology as neutral and objective (Crawford, 2023). This aligns with findings from Intahchomphoo and Gundersen (2020), who argue that AI systems can perpetuate existing racial biases and even produces imbalanced professional prospects for people from certain racial groups (Intahchomphoo & Gundersen, 2020).

The implications are profound, as these technologies are increasingly used in law enforcement and public surveillance, raising ethical concerns about their deployment in racially diverse communities (Julia Angwin, 2016; Noble, 2018; Ulnicane & Aden, 2023).

2.2.3 Human-Computer Interaction (HCI) in AI:

Human-Computer Interaction (HCI) is a multifaceted field that encompasses various disciplines and methodologies aimed at improving the interaction between users and computers (Carroll, 2009) The evolution of HCI has been significantly influenced by advancements in technology, user-centered design principles, and interdisciplinary research, leading to more accessible and effective user experiences (Carroll, 2009).

From complex tasks such as self-governing driving to simple tasks as activating a filter preventing spam emails to arrive on our inboxes, AI have continually grown in HCI possibilities (Yang et al., 2020).

The cognitive aspects of HCI are also critical, as understanding user behaviour and decision-making processes can significantly enhance interface design (Carroll, 2009; John & John, 2001). Research by Cepeda et al. (2021) demonstrates that analysing where the mouse pointer is moving can reveal insights into user uncertainty and cognitive processes during web interactions (Cepeda et al., 2021)

The integration of artificial intelligence (AI) and Human-Computer Interaction (HCI) (AI) raises critical concerns regarding the manifestation of racism within AI systems (Schlesinger et al., 2017).

Systemic racism significantly influences the design and implementation of AI systems, often leading to biased outcomes that disproportionately affect racial and ethnic minorities (Buolamwini & Gebru, 2018).

Moreover, the ethical implications of AI in HCI are underscored by the need for transparency and accountability in AI systems (Oluwaseun Augustine Lottu et al., 2024). Schmidt et al. 2024 advocates for a human-centered approach to AI, emphasising the necessity of ethical considerations in the design process to ensure that AI technologies serve all users equitably (Schmidt et al., 2024). Because the data AI learns from, to fully delegate decision making to AI, may introduce to systems ethical questions (Schmidt et al., 2024).

Since AI tends to intensify pre-existing human biases, the challenge lies in creating AI systems that not only function effectively but also promote inclusivity and fairness (Chen et al., 2023; Ferrara, 2023; Md Sumon Gazi et al., 2024). The role of HCI in mitigating AI-related racism is further explored by Yang et al. (2020), who argue that the unique challenges of designing human-AI interactions necessitate a deeper understanding of user experiences and the potential for unintended consequences (Yang et al., 2020).

2.2.4 Imagery Generative AI Systems: Overview and Applications

Imagery generative AI systems have emerged as significant tools in the creative scene (Hughes et al., 2021), enabling users, who can have access to it, to generate high-quality images through sophisticated algorithms such as Generative Adversarial Networks (GANs)(Hughes et al., 2021). Generative adversarial networks (GANs) broadly speaking, is a method for deep learning representations without the need for heavily annotated training data where new authentic data is generated in two competings neural networks (Creswell et al., 2018). Despite limitations, the advancement in GANs research has undeniably highlighted its vast potential for future use of AI image Generative tools (Wang et al., 2017).

These prompts influences how effectively users can communicate their artistic intentions to the AI (Torricelli et al., 2024). This interaction not only reflects the behavioural aspects of users but also highlights the importance of user-friendly interfaces in enhancing the creative process (Torricelli et al., 2024).

The implications of generative AI extend beyond mere image creation; they are reshaping entire industries, including marketing and content creation (Murár & Kubovics, 2023). AI tools are increasingly being integrated into marketing strategies, enabling marketers to produce relevant and engaging content rapidly (Murár & Kubovics, 2023)

The accessibility of these tools allows individuals without specialised skills to create high-quality visuals, potentially democratising content creation (Murár & Kubovics, 2023).

Furthermore, the integration of AI in social media platforms enhances user experience by curating personalised content, thus improving engagement and relevance (Murár & Kubovics, 2023). However, a recent study in Brazil highlighted a concerning downside of social media, showing its role in deepening social inequality. During a recent election, candidates with greater financial resources were better positioned to leverage social media platforms to influence political outcomes in their favour, leaving those with fewer means at a significant disadvantage (Rodarte & Lukito, 2024).

The intersection of generative AI and racial equity is complex, as it involves both the potential for positive change and the risk of reinforcing structural inequalities (Buolamwini & Gebru, 2018; Tiku et al., 2023; Verma, 2022). One of the primary concerns regarding generative AI is its tendency to reflect and amplify biases present in the data on which it is trained (Bender et al., 2021; Broussard, 2018). Research indicates that many AI systems, including those used in healthcare (Godsil et al., 2014) and criminal justice (Birhane, 2021), have been developed using datasets that predominantly represent white populations (Cave & Dihal, 2020).

This lack of diversity in training data (Arantes, 2020; D'ignazio & Klein, 2020) can lead to algorithms that perform poorly for marginalised groups (O’Neil, 2017), thereby exacerbating existing disparities in access to services and opportunities (Farahani & Ghasemi, 2024). For instance, studies have shown that AI algorithms in healthcare have been disproportionately trained on white populations, raising concerns about their effectiveness and fairness when applied to racially diverse groups (Bailey et al., 2021; Ferdinand et al., 2015). This phenomenon, often referred to as "algorithmic bias"(Finn Lattimore, 2020), highlights the urgent need for more inclusive data practices and rigorous validation across diverse populations (Finn Lattimore, 2020).

In the Australian context, where circa 91% of the population have European background (ABS, 2022), racial inequalities (Mapedzahama, 2019) may be shaped different than the United States of America (Medina, 2023). Where, in Australia, we have a deep institutional, economic and social Aboriginal versus General population gap (Moreton-Robinson, 2021), we also have other racial groups forming our society. Asians (east and south), Pacific Islanders and Middle Eastern are significantly higher minorities here (Mapedzahama, 2019). While it could be argued that there is a growing antiblackness sentiment due to recent arrivals of Africans and the Aboriginal resentment (APH; Majavu, 2020), antiblackness is not necessarily the main racial barrier in the Australian context (Dunn et al., 2004). In this research, some of the themes picked up were xenophobia, racial stereotypes and white supremacist views of positionality.

It can be also put forward that AI systems that incorporate critical race theory as well as localised and grassroots racial knowledge through algorithmic literacy programs and mitigation racial tools, can empower minority communities to subvert and survive AI-mediated racism (Sukhadeve, 2021).

2.2.5 Bias in Imagery Generative AI Systems

Before enumerating some of the bias in AI image generative systems, we need to stablish the foundations of it. As proposed by (Finn Lattimore, 2020), the structural world inequality (Stiglitz, 2012) among populations is at the foundation of AI bias. This unbalance between, countries, gender, race, ability and economical class - just to cite a few, are at the core of the bias in the data (Finn Lattimore, 2020).

There are many types of bias in AI tools (Finn Lattimore, 2020). Bias in Data (O'Neil, 2017), bias in modelling, bias in training and ultimately bias in users (Ferrer et al., 2021; X. Wang et al., 2023). I will mostly focus on racial bias but gender, cultural, or socio-economic biases are also present in my thesis.

Bias in imagery generative AI systems have been described in research in recent years (Bass, 2023; Bhaimiya, 2023; Chellappa & Niiler, 2022; Gichoya et al., 2022). In the field of computer vision – an industry that has been growing by billions of dollars each year, the harms go way beyond biases with harmful consequences to outgroups of all races globally (Gebru & Denton, 2024).

These systems, which utilise algorithms to create images based on input data (Creswell et al., 2018), can inadvertently perpetuate existing societal biases, particularly those related to race, gender, and other human characteristics (Ferrer et al., 2021; Finn Lattimore, 2020). For example, non-cisgender identities are often depicted in ways that are overly sexualized, heavily stereotyped, and less human, perpetuating harmful biases and further marginalising these communities (Ungless et al., 2023).

The implications of such biases are profound, affecting various domains including healthcare (Sargsyan et al., 2024), employment (X. Wang et al., 2023), and beauty standards (Jagota, 2016; Riccio & Oliver, 2022).

Many datasets are historically skewed (Buolamwini & Gebru, 2018), often over-representing certain demographics while under-representing others (X. Wang et al., 2023). For instance, Riccio & Oliver (2022) highlights that the social media so called "beauty filters" have been commonly condemned for propagating racism as they often lighten the skin tone, reduce size of noses and make eyes bigger tending to support an Eurocentric standard of beaty (Riccio & Oliver, 2022; Trammel, 2023). Characteristic such as high cheekbones, whiter skin, slim noses are seen as more beautiful by AI generative tools(Silva, 2023).

Although this is not a new phenomenon as photography standards have been of white skin people since last century (Lewis, 2019), those filters predominantly nudge users to see themselves as more beautiful when they look lighter even when users are from different racial background (Jagota, 2016; Li, 2020), this emphasises the need for diverse and representative training data (Riccio & Oliver, 2022).

This underscores the critical importance of ensuring that training datasets are not only balanced but also reflective of the diversity present in the real world (Johnson, 2021; O'Neil, 2017). Moreover, the algorithms themselves can exacerbate these biases (Johnson, 2021). Even when datasets are balanced, the learned models can amplify associations between target labels and gender, leading to biased outputs (Ferrara, 2023). This phenomenon of gender bias exists in other AI tools such as

ChatGPT (Walther et al., 2024). However, it is particularly evident in generative models like DALL-E, which have also been shown to produce images that reinforce gender stereotypes (Bass, 2023; Wellner, 2020).

Such findings raise ethical concerns about the potential for these systems to shape societal perceptions and reinforce harmful stereotypes (Nadeem et al., 2022). The impact of bias in generative AI systems extends beyond mere representation. It also has the potential to shape critical choices in areas such as healthcare (Vyas et al., 2020). AI systems in medical imaging can systematically disadvantage certain populations, influencing inequities in diagnosis and care (Vyas et al., 2020).

The failure to address these biases not only perpetuates existing inequalities but can also lead to detrimental outcomes for underrepresented groups (Buolamwini & Gebru, 2018; Finn Lattimore, 2020). Mitigation strategies for addressing bias in generative AI systems are essential (Ferrara, 2023; Finn Lattimore, 2020).

Various approaches have been proposed, including the use of data augmentation techniques to enhance model performance and reduce bias (Khakurel & Rawat, 2023) Furthermore, the development of unbiased algorithms (Tolan, 2019) and responsible data management practices is crucial for fostering fairness in AI outputs (Tolan, 2019) . As highlighted by (Buolamwini & Gebru, 2018), addressing gender bias in image search engines requires ongoing efforts to ensure equitable representation and to challenge existing stereotypes. Furthermore, ethical considerations must be at the forefront of discussions surrounding generative AI (Andreotta et al., 2022; Finn Lattimore, 2020).

The potential for these technologies to influence public perception and societal norms necessitates a multi-disciplinary approach to soft governance and regulation (Andreotta et al., 2022). This includes fostering educational programs that raise awareness about the implications of bias in AI-generated imagery and promoting ethical discussions within the field (Andreotta et al., 2022) Ultimately, addressing these biases requires a concerted effort to ensure diverse representation in training data, the development of unbiased algorithms, and a commitment to ethical practices in AI deployment (Andreotta et al., 2022; Finn Lattimore, 2020).

As these technologies continue to evolve, it is imperative that all stakeholders remain attentive in our efforts to mitigate bias and promote fairness in AI-generated imagery (Bass, 2023; Tiku et al., 2023).

2.2.6 Attitudes Toward AI and Bias

Student perceptions of artificial intelligence (AI) are increasingly significant as educational institutions integrate AI technologies into learning environments (Alrayes et al., 2024; Hamilton, 2023). At University level, students exhibit a diverse range of perceptions regarding AI in educational contexts (Fošner, 2024; Keles & Aydin, 2021; Stewart et al., 2023). In Werner et al. (2024) research, initial results indicate that many University students lack of the full knowledge of Gen AI and usage of it may not yet be as widely as it is believed (Werner et al., 2024).

Nevertheless, as this paper is written, it is likely that this data is shifting as more and more students discover AI tools (Chan, 2023; Walsh, 2023a). A positive result from AI implementation is the possibility for university students to speed up their work processes, therefore, using AI can be seeing as a supportive tool for human creativity and endeavour instead of a replacement of it (Chubb et al., 2022).

For example, a study involving 399 students in Hong Kong highlighted a generally positive attitude towards generative AI technologies, emphasising their willingness to engage with these tools for academic purposes (Chan & Hu, 2023) However, concerns about the reliability of AI algorithms and ethical implications persist among students (Chan & Hu, 2023). Dental students, for instance, expressed apprehensions regarding the integration of AI into their education, emphasising the necessity for proper training and ethical considerations (Kumari, 2023)

In employment for example, resumé screening apps that are supposed to facilitating hire, can exhibit racial and gender biases, favouring White-associated names in 85.1% of cases while disadvantaging Black males in up to 100% of cases, according to this study by (Wilson & Caliskan, 2024).

Closer to home in Australia Dann et al. (2024) suggests that teachers and academic staff can tap into insights from AI-crunched data to help shape how they teach and make cleverer choices in higher education, making student experiences better all around (Dann et al., 2024). In another study, Australian business students had variety of multidisciplinary ethical questions on the perspectives on AI use which in turn raised questions for academics and the university in general (Murray & Williams, 2023).

In Australia, because of its majority European population (ABS, 2022), There is a potential for exclusion of diverse demographic groups from training datasets and usage impacts which can lead to skewed outcomes that fail to represent the needs of these populations (Aggarwal et al., 2021). These impacts reinforce systemic inequalities, privileging dominant groups while sidelining the unique experiences and realities of others (Aggarwal et al., 2021; Eidelman & Crandall, 2012).

Furthermore, I agree with Albarrán et al. (2021) when they argued that there is a need to engage and educate students and staff for the use of AI, because individuals who are less informed about AI technologies tend to harbor negative attitudes, particularly if they perceive AI as unhelpful or threatening to their jobs.

Fostering positive attitudes toward AI usage requires not only addressing bias but also enhancing public (in this case students) understanding and trust in AI technologies (Broad, 2018; Crockett et al., 2020; Taeihagh, 2021).

2.2.7 Mitigating Bias in AI Systems

Mitigating bias in AI systems in Australia demands a multidimensional approach that integrates diverse datasets, algorithmic transparency, and ethical frameworks (Awasthi & George, 2020; D'ignazio & Klein, 2020; Finn Lattimore, 2020). At various stages of AI development, including deployment, model training, and data collection, bias can emerge, disproportionately affecting marginalised groups (Bender & Friedman, 2018; Gichoya et al., 2023; Mittermaier et al., 2023).

Interdisciplinary collaboration among stakeholders in the university sector is essential for developing robust ethical guidelines that promote fairness and accountability in AI applications (Dann et al., 2024; Murray & Williams, 2023). Engaging diverse expertise, from computer science, ethics, law, and social sciences can facilitate a comprehensive understanding of AI's societal impacts (Benefo et al., 2022; X. Wang et al., 2023). Additionally, "more research should also be conducted to understand how AI explanations can be better designed to more effectively expose the internal biases of AI models" (Wang et al., 2023 p. 7).

Current approaches to detecting and mitigating bias in AI in Australia involve a multifaceted strategy that integrates technical, social, and ethical considerations (Finn Lattimore, 2020). A major way to tackle this is through fairness trained machine learning tools, which could work to spot and fix biases in algorithms using tools like algorithm audits and measures to ensure everyone gets a fair go (Dr. Rohit Markan & M. Rajalakshmi, 2024; Mehrabi et al., 2021). Moreover, addressing bias requires a broader understanding of its origins, including the socio-technical dynamics that contribute to algorithmic discrimination (Lee, 2018; Ulnicane & Aden, 2023).

Furthermore, the lack of diverse datasets can lead to systematic underrepresentation of marginalised groups, necessitating the development of inclusive data practices (Celi et al., 2022)

In Australia, these strategies are increasingly recognised as vital for promoting equity and inclusion in AI applications (Finn Lattimore, 2020; Murray & Williams, 2023). However, Australia's fragmented regulatory approach to technological sectors like the digital sharing economy, exemplified by platforms such as Airbnb, emphasises the challenges of achieving consistent governance across our states and territories, complicating enforcement and policymaking (Hati et al., 2021).

Holmes et al. (2021) posits that a community-wide framework for AI ethics that acknowledges the diverse interpretations of ethical issues within the educational community is needed (Holmes et al., 2021). It is also suggested that flexibility in developing ethical guidelines is paramount (Holmes et al., 2021). It can be argued that the combination of ethics education, integrated curricula and robust teacher training is vital for AI in education (Holmes et al., 2021; Murray & Williams, 2023)

2.2.8 The Exploration of Bias and Attitudes Toward Generative AI Systems

The exploration of bias and attitudes toward generative AI systems among Australian university students is crucial for several reasons. Firstly, Chan & Hu (2023) emphasise that diverse student perceptions of generative AI highlight the need for integrating AI literacy into higher education, which is essential for understanding its implications and applications in academic contexts (Chan & Hu, 2023).

As AI tools become more embedded in education and workplaces, it’s vital to understand how students interact with these technologies to ensure their fair and effective use (Kurian et al., 2023). Bias in AI often arises from datasets that fail to reflect diverse groups, reinforcing existing social inequalities (Buolamwini & Gebru, 2018; Noble, 2018). If left unchecked, biases in these tools will continue to elevate Eurocentric ideals and exclude marginalised perspectives (Riccio & Oliver, 2022).

Moreover, Murray's (2023) views rest on the assumption that students possess a multifaceted mix of ethical perspectives regarding generative AI, suggesting that their attitudes can meaningfully influence how these technologies are utilised in educational settings (Murray & Williams, 2023).

Attempting to understand those biases and attitudes is key for (re)shaping concrete educational policies and practices which can build a positive environment for students, staff and academics at a university setting. These future policies and practises will, likely, advance learning outcomes ("ChatGPT in Higher Education: Considerations for Academic Integrity and Student Learning," 2023).

Addressing these biases is essential to creating AI systems that are inclusive and representative (Favaretto et al., 2019; Ferrer et al., 2021). In Australia, discussions about racial inequities often focus on Indigenous communities, leaving the experiences of other racialised groups underexplored (Gatwiri et al., 2021).

2.2.9 Gaps in the Literature

There are significant gaps in the literature on Australian students' attitudes toward bias in AI-generated imagery. It is especially concerning that students often lack awareness and understanding of the implications of biases embedded in AI systems (Applebaum, 2022; Nelson et al., 2013). This lack of awareness risks leaving students uncritical of how AI systems perpetuate stereotypes, particularly in imagery generative tools, which can influence creative and professional decision-making.

While studies have examined students' perceptions of AI in various fields, including healthcare and business (Hashmi et al., 2023; Murray & Williams, 2023); and although, we have started to see research on the impact of AI generated images overseas (Silva et al., 2024; Testón et al., 2023), there is a lack of focused research on how Australian students specifically perceive bias in AI-generated imagery. This thesis attempts to address this issue.

Alrayes et al. (2024) discusses the importance of transparency and the potential for biases in AI-generated content, which could mislead users. Comparably, Williams-Ceci et al. (2024) provides evidence that biased AI can stealthily influence societal attitudes, highlighting the need for awareness among users. Furthermore, Zhang et al. (2024) explores emotional responses to AI-generated imagery, indicating that students' professional backgrounds may affect their perceptions of images generated by AI, yet this area remains underexplored in the Australian context.

This thesis also proposes that in further studies, analysis of racial, economical, religious, political, gender and ability backgrounds should be investigated when reacting to AI imagery generative systems (Wellner, 2020). This thesis did not include the selection of students based on their backgrounds in its original research design, nor did it propose to analyse participant demographics as a primary focus (Oliver-Hoyo & Allen, 2006). Moreover, the primary focus of the study was monitoring attitudes rather than analysing participant demographics (Bhandari, 2022; Oliver-Hoyo & Allen, 2006).

The integration of AI in education, particularly in Australia, presents numerous avenues for further research (Benefo et al., 2022; Medina, 2023). This thesis posits that, there is a necessity in investigate issues concerning AI bias, its professional and wellbeing effects, and the educational impact of AI on students (Murray & Williams, 2023). As AI technologies become increasingly prevalent (Sargsyan et al., 2024), understanding their implications on educational equity and student experiences is crucial (Chan & Hu, 2023).

Trust in AI technologies is paramount for their successful integration into educational practices (Albarrán et al., 2021) Students perceptions of AI can significantly influence their engagement and willingness to utilise these technologies (Kim, 2023)

AI literacy as well as AI bias, AI data, AI imagery synthetic creation, and AI tools usage by university students, offers incredible opportunities for further research in Australia (Brew et al., 2023; Dwivedi et al., 2023). Addressing these areas will not only enhance our understanding of AI's role in education but also contribute to the development of ethical, equitable, and effective educational practices that can benefit students, academics and university staff in general.

Chapter 3: Research Method

3.0 Introduction

This chapter provides an overview of the design, theory and methods utilised in this research thesis. I address the research aims, research questions & sub-questions, research design, sampling design, ethical considerations, recruitment & interviews, the gender of interviewees, theoretical research approach, the data collection, interview process and thematic analysis, platform analysis, methodological rigor and the limitations of the study titled: “Can Bias Embedded in Image-Generative AI Systems Influence Public Perception.” This chapter also provides ethical considerations, containing consent measures, privacy methods, and authorisation from the Victoria University Human Research Ethics Committee.

3.1 Recruitment & Interviews

The study aimed to gather diverse perspectives from students across various Australian universities. The research followed established ethical guidelines (Zion, 2023). The methodology utilised was semi-structured interviews to facilitate rich data collection, emphasising the importance of participant engagement and the quality of the research findings (Creswell & Creswell, 2017; Leavy, 2022).

Recruitment is a critical component of qualitative research, influencing both the diversity of perspectives gathered and the overall validity of the findings (Creswell & Creswell, 2017). In this study, participants were recruited through personal social media accounts, university posts, and word of mouth.

Early on, there was a constraint in acquiring Victoria University students who were connect in anyway with either of both of my supervisors. This, in the beginning, diminished the chances of getting support from them in the recruitment. This made the recruitment in my “home turf” more difficult.

In the original proposal concept, there was an idea of acquiring six students from any Australian university. They needed not to be Australian citizen but needed to be enrolled in an accredited institution within Australia.

This broad criterion was established to maximise the likelihood of recruiting the necessary number of participants (Campbell et al., 2021). However, the absence of financial compensation for participation added an additional layer of complexity which, consequently, forced me to rely on the kindness of participants. The recruitment strategy thus demanded a careful balance between ethical considerations and the practicalities of engaging participants.

Besides all the obligatory ethical and privacy considerations (Olsen & Mooney-Somers, 2017) to be placed in the formal invitations and on research information, I made sure to add information such as "what the participants would be asked to do" and "information of consent and psychological support information" for example.

Detailed information was also provided regarding the voluntary nature of participation, the right to withdraw at any time, and the measures taken to ensure data confidentiality. In addition to outlining the study's purpose and procedures, comprehensive information about psychological support resources was made available to participants (Denzin, 2010).

Also, when creating the advertising poster to email people and place it on university boards, I made sure that there was a section on what will you gain from the research:

"What will I gain from participating?"

My reply was "Through participation in this research, you will be contributing to the theoretical and practical understanding of AI Usage and effects on humans is an emerging and arguably important area of the research endeavour…"

In my view, this approach not only sought to attract participants but also aimed to instil a sense of purpose and agency in their involvement, thus increasing the likelihood of engagement (Oliver-Hoyo & Allen, 2006). This, not only enhances the quality of the research but also fosters a sense of ownership and purpose among participants (Oliver-Hoyo & Allen, 2006).

Such transparency was essential in fostering trust between me, the researcher, and the participants, which was particularly important for this qualitative research where I knew personal narratives and experiences would have been shared (Oliver-Hoyo & Allen, 2006). Also, it is important to add that by prioritising participants well-being, following the ethical research guidelines, I addressed any potential emotional discomfort due to research questions [2].

[2] No participant expressed any visible nor verbalised any discomfort during the research interview phase.

It was originally planned to get as much diverse participants as possible. Thus, a key objective of the recruitment strategy was to achieve maximum diversity in gender, race, and ethnic background among participants (Campbell et al., 2021). The advertising materials explicitly refrained from indicating a preference for any particular gender, nationality or racial ethnical group.

This commitment to diversity is essential in qualitative research, as it enriches the data collected and enhances the generalisability of the findings (Campbell et al., 2021). By actively seeking a diverse participant pool, my thesis aimed to capture a wide range of experiences and perspectives related to AI Image Generative tools usage.

Following the initial outreach, responses were received from nine students (two males and 7 females) via social media and email, after a more professionally worded email, six students officially responded for participation.

Surprisingly, however, my final sample of participants was all female, aged 19-36, from different universities within Australia. This study is confident that there was a diverse sample of students from different universities across Australia. I acknowledge, nonetheless, that this is an acceptable challenge and that qualitative studies are subject to have imbalances as they rely on whoever is available and how diverse the pool is (Oliver-Hoyo & Allen, 2006).

Data collection was conducted using Riverside.fm, a platform that offers both video and audio recording capabilities (Riverside.fm, 2024). To maximise participants privacy, only audio was saved and used for the interviews NVivo thematic analysis. All interviews lasted approximately 45 to 60 minutes.

The choice of Riverside.fm aligns with best practices for qualitative data collection, ensuring high-quality audio capture that facilitates accurate transcription and analysis (Riverside.fm, 2024). This methodological decision underscores the importance of employing reliable tools to enhance the integrity of the data collected (Oliver-Hoyo & Allen, 2006).

The research employed semi-structured interviews as the primary data collection method (Creswell & Creswell, 2017). This approach allowed for flexibility as I utilised pre-developed questions, therefore creating organic and responsive dialogue between me, the researcher and the participants. The semi-structured format is particularly advantageous in qualitative studies, as it encourages participants to share rich, nuanced perspectives that may not emerge in a more rigid interview format

(Campbell et al., 2021). I believe this gave me better insights into participants' experiences and their perceptions of AI Image Generative Systems usage.

3.1.2 Gender of Interviewees

The research sought to ensure a balanced representation across racial, socioeconomic, ability, and gender lines among interview participants to obtain a broad spectrum of perspectives on AI biases (Leavy, 2022). This approach aligns with the principles of qualitative research, which emphasise the importance of capturing diverse viewpoints to understand complex social phenomena (Leavy, 2022).

After participants were invited to the interview rounds, surprisingly, the final sample cohort consisted solely of female university students from Australian institutions, all aged between 19 and 36 years. This gender homogeneity was not intentional. Despite this unexpected sample composition, the participants represent a range of universities across Australia, contributing to the study's diversity. According to (Martin, 2022) gender can significantly influence perceptions and experiences related to technology, including biases in AI systems. Therefore, this all-female cohort presents a limitation in generalising the findings, as the influence of male perspectives on AI biases remains unexplored in this research.

By incorporating a wider range of perspectives, future research can provide a more nuanced understanding of how various demographics factors intersect with experiences of AI bias, ultimately leading to more effective strategies for mitigating these biases in AI systems (Collins & Stockton, 2018).

There are a few possible explanations for why I ended up with an all-female ensemble.

For example, prior to the 1990s, women constituted only 26-30% of participants in clinical trials, leading to significant gaps in understanding the efficacy and safety of treatments for the female population (Politi et al., 2013). The long-standing practice of focusing research questions primarily on male subjects, resulted in a lack of gender-specific analyses in many studies (Dhali et al., 2022).

However, the dynamics of participation can vary significantly depending on the context of the study. While women may have higher participation rates in some contexts, men may also engage actively in studies when the topics resonate with their interests or when they perceive minimal risk involved (Boydell et al., 2015). Thus, addressing these disparities is crucial for ensuring equitable representation in research and better outcomes for all genders (Chen et al., 2014).

Another reason for the number of females to outnumber the male participants is the fact that female students have outnumbered male students in various fields of study in Australian universities, particularly in health and education disciplines (Wallace et al., 2010).

Furthermore, according to a report by the Workplace Gender Equality Agency, young women in Australia are generally more educated than their male counterparts, indicating a shift towards greater female participation in higher education (Huang, 2023). This trend is supported by data showing that women now make up a significant proportion of university students, with many universities reporting female enrolment rates surpassing those of males (Tran et al., 2021). Incidentally, this increase in women participation in Australian universities, have also helped women from African Australian backgrounds (Harris et al., 2014). Furthermore, in Australia, women with Africa heritage, have been increasingly enrolling in tertiary education (Harris et al., 2014). This supports the fact that one of my participants was an African Australian woman (Harris et al., 2014). Also, although, in Australia, First Nations peoples in higher education have been increasing, retention rates remain low (Wolfe, 2006). Nevertheless, more recently, there has been an increase in Indigenous staff and students, particularly among women (Briese et al., 2024). Which, in turn, supports having amongst my participants an Aboriginal Australian student (Briese et al., 2024).

Another factor can be that increase in women participation in tertiary education is that universities have also actively worked to foster more supportive environments for women working in the higher education sector, addressing gender disparities and promoting inclusivity (Kanyumba & Lourens, 2022). Australian higher education system has been praised for its relatively high levels of gender equality compared to global standards (Gilbert et al., 2021; Lahiri-Dutt, 2018). Another factor may relate to the recruitment process itself, as the channels, my networks, and the events I accessed during participant recruitment were, not by choice, predominantly female, and it is possible that the flyer's design and language may have also resonated more strongly with women. These elements, combined with my rapport style during recruitment, may have contributed to the gender composition of the final sample.

The thesis, therefore, did not take any accounts the views of male students, thus reducing the generalisability of the findings.

Nevertheless, by incorporating a wider range of perspectives, future research can provide a more nuanced understanding of how various demographics factors intersect with experiences of AI bias, ultimately leading to more effective strategies for mitigating these biases in AI systems (Collins & Stockton, 2018).

3.2 Research Aims

The primary aim of this research is to investigate the impact of biases in imagery-generative AI systems on Australian university students. This study seeks to understand how these biases influence users' perceptions and design decisions, particularly through the lens of Critical Race Theory (Vass, 2015). Critical Race Theory (hereafter also referred as CRT) is an intellectual movement and framework that originated in the United States during the late 1970s and early 1980s, primarily within the field of legal studies (Ladson-Billings, 2021; Lynn & Dixson, 2013; Tate, 1997). "CRT has much to offer educational research in Australia" (Vass, 2015 p, 373). CRT serves as a vital framework for analysing the intersection of race, power, and technology, allowing for a nuanced exploration of how biases embedded in AI systems can influence user, in this thesis case, University Students (Creswell & Creswell, 2017)

By focusing on a local Australian cohort. Similar to other studies, the research aims to contribute to the existing literature in Australia (Chavan et al., 2014), on AI bias and, hopefully, support the development of more responsible AI tools and AI specific policies in order to mitigate bias (Finn Lattimore, 2020). This localised approach is essential (Kelly et al., 2019), for much of the current research on AI bias comes from international contexts - particularly the USA - (Zhou et al., 2022).

The USA's early engagement in AI research is evident, with foundational studies dating back to 2001 while, for example, China's involvement began later in 2009 (Zhou et al., 2022).

This thesis proposes instead to capture local Australian idiosyncrasies (Bennett et al., 2021) and our cultural Australian racial identities, social interactions and our economic nuances (Collins & Stockton, 2018). Understanding the specific experiences and reactions of Australian university students will provide valuable insights into the broader implications of AI bias within the local context (Collins & Stockton, 2018).

This thesis aimed to understand how Australian university students react to biases in AI generative systems (Mbalaka, 2023). This study does this by observing their answers in qualitative research using thematic analysis on NVivo (Elliott-Mainwaring, 2021), and by framing the thesis through a critical race theory overarching lens (Mills & Godley, 2017; Vass, 2015). The study will focus on users' perceived attitudes towards AI tools (Abbas et al., 2023), examining how these attitudes are shaped by their experiences and the socio-political environment in which they operate (Leavy, 2022).

By utilising qualitative methods, the research investigated the complexities of users' interactions with AI systems, recognising that these interactions are not merely technical but are deeply embedded in social practices and power relations (Mogashoa, 2014)

By utilising thematic analysis, the study attempted to identify recurring patterns and themes in participants' responses, allowing for a comprehensive understanding of how biases manifest in their experiences with AI tools (Farahani & Ghasemi, 2024; D. Wang et al., 2023).

Furthermore, to enrich the analysis, the research will critically engage with the existing literature on AI bias, drawing on various theoretical perspectives, including feminist, queer, Marxist and critical race theories (Bhaimiya, 2023; Bragazzi et al., 2023; Cave & Dihal, 2020; Creswell & Creswell, 2014; Impett, 2018; Rosenbaum, 2022).

This multidisciplinary methodology - in the context of CRT (Lynn & Dixson, 2013) will guarantee that the findings are robust and reflective of the diverse experiences of the participants, ultimately contributing to the development of frameworks that promote equity and justice in AI design and implementation (Leavy, 2022)

Through this investigation, the study aspires to not only highlight the damaging effects of biases in AI systems (Metz & Wakabayashi, 2020; X. Wang et al., 2023) but also to support other academics and institutions in creating better guidelines and frameworks to tackle this new reality where automation bias can be catastrophic for humanity (Ajunwa, 2020).

3.3 Research Questions and Sub-questions

The original overarching research question guiding this study was: Does AI have biases[3], and how do these biases manifest in user interactions? To explore this, the following sub-questions[4] will be addressed:

- How do biases in AI imagery generative systems affect users' design decisions?
- What are the perceived attitudes of university students towards AI tools?
- How can the impact of biases in AI be mitigated in future applications?

The final primary research question that I decided to use on the thesis is: Can Bias Embedded in Image-Generative AI Systems Influence Public Perception?

3.3.1 The questions asked in the Interviews were

Exemplar questions:

The choice of a qualitative semi structured interview method allowed this thesis to generate questions that can provide a deep understanding of how participants feel specially comparing to a quantitative approach (Schostak, 2005). Also, the choice of research design was limited due to time, financial and human resources constraints as this is a project of a Master of Research.

Some of the questions, such as “What is the first time you have heard or what is your earliest memory about Artificial Intelligence?”. I did this with the objective of establish a connection with the participants (Leavy, 2022). Engaging research participants is a critical aspect of the research process that enhances the relevance of findings (Long et al., 2017). This aligns with the principles of qualitative research, which values depth of understanding and the idiosyncratic experiences of individuals(Leavy, 2022).

This initial approach is crucial in qualitative research, as it helps to create a comfortable environment where participants feel at ease to share their thoughts and

[3] Artificial Intelligence exhibits inherent biases, as evidenced by numerous academic papers, articles in reputable newspapers and magazines, and extensive research studies.
[4] This became the overarching question during the research.

experiences (Leavy, 2022). By evoking their earliest memories, creating icebreakers, I allowed participants to open up and engage more deeply and trustfully with the research process (Leavy, 2022). Trust is a fundamental factor influencing research participation (Smirnoff et al., 2018).

By asking participants to reflect on an earlier time in their lives, I aimed to evoke memories that could foster a sense of nostalgia and connection (Holbrook & Schindler, 2003). This technique is particularly valuable in qualitative interviews, as it encourages participants to share their narratives and insights, thereby enriching the data collected (Carpenter & Lertpratchya, 2016)

The intention behind this strategy was to establish a rapport that would facilitate a more meaningful dialogue throughout the interview (Fereday & Muir-Cochrane, 2006).

The initial light questions and personal inquiries were strategically designed to build rapport, encourage participant engagement, and ultimately facilitate a more profound exploration of the research topic (Leavy, 2022; Vallano & Compo, 2011).

The next question, “Please describe what you know about imagery generative systems and how they work”, attempted to establish an understanding of how much the participants knew about AI image generative tools. This is in line with how language mediates experience and how individuals articulate their views in relation to their knowledge and experiences (Schostak, 2005). This inquiry was essential for several reasons. Firstly, it provided insight into the participants' familiarity with contemporary AI technologies, which is crucial for contextualising their responses and perspectives throughout the research, thus, understanding the dynamics of subject positions, which further supports the need to gauge the participants’ knowledge of AI tools to contextualise their insights effectively (Schostak, 2005)

Furthermore, the question's design reflects the principles of thematic analysis, where initial responses can inform the development of themes and categories for further exploration (Fereday & Muir-Cochrane, 2006).

After showing the first image, participants were asked to engage in a reflective exercise designed to elicit their thoughts and feelings about the visual content presented to them. These question below were repeated five times, each time corresponding to a different image shown during the session. The reflective exercise aimed to deepen participants' engagement with the visual content, allowing for a nuanced exploration of their thoughts and feelings. Repeating the question for each image ensured comprehensive feedback and facilitated a richer understanding of their responses to varying visual incentives (Schostak, 2005).

> "Is the image credible/believable to you? In what ways can you tell this?
>
> Please describe, as clearly as you can this image – for someone who can't see this image at all.
>
> Why this image seems accurate/natural to you. If the opposite, please state why?"

The purpose of this repetition was to ensure that participants had the opportunity to articulate their perceptions consistently across various stimuli, allowing for a comparative analysis of their responses (Leavy, 2022).

The specific question posed to participants was crafted to encourage detailed descriptions and personal interpretations of the images. By prompting participants to describe everything they could think and see regarding each image, the research aimed to gather rich, qualitative data that could reveal underlying themes and patterns in their perceptions (Vears & Gillam, 2022). This approach aligns with the principles of qualitative research, which values depth of understanding and the subjective experiences of individuals (Leavy, 2022)

In qualitative research, particularly when employing thematic analysis, it is essential to capture the nuances of participants' responses. Thematic analysis involves identifying meaningful categories or themes within the data, which can emerge through an inductive process as participants share their insights (Mogashoa, 2014) By repeating the question for each image, the study aimed to facilitate a comprehensive exploration of how different visual representations might evoke varied emotional and cognitive responses from the participants (Carpenter & Lertpratchya, 2016).

In this question, "How do you think this AI can be used for work and workers in the present and future?", I seek to produce in participants the perceptions of AI's current and future roles in the workplace.

It encourages the participants of the research to reflect on the transformative potential of AI technologies, including efficiency improvements, task automation, and the creation of new job opportunities or roles that may emerge because of AI integration (Kokina & Davenport, 2017).

In what ways could AI imagery generative systems be used in your work today?

This question specifically targets the application of AI image generative systems in participants' current professional practices (Kokina & Davenport, 2017; Poba-Nzaou et al., 2021).

This question, In which ways can the images that come out of the imagery generative systems have an impact on the real world?, encourages participants to think critically about the broader implications of AI-generated imagery.

These four next questions below were only asked after the participants had seen all five images. By this point in the interview, participants were more aware of racial elements of the thesis (Mills, 2014), having been exposed to images that depicted significant cultural and social contexts, such as an Australian migrant family and the picture of the Aboriginal man working in Melbourne (Entman & Rojecki, 2001). This approach capitalised on their enhanced awareness, ensuring responses were contextualised within the visual stimuli presented (Schostak, 2005).

“Does it matter what kind of images are generated? If it does why? If it doesn’t, why?

When looking at these images, can you give me examples and describe different ways the image be represented?

What are your suggestions for improving the image generative systems, if any?

Thinking of your personal experience in Australia: In your words, does this image represent an accurate portrait of what the population of Australia looks like today?

In your opinion, do you feel anyone could feel discriminated against by the image? In what ways?”

These questions were place here as a strategic timing allowing for a deeper exploration of their perceptions and interpretations, as the visual stimuli served to prime their thoughts regarding race and identity within the framework of the research.

(Dunn & Luchner, 2022) discuss how prompts can activate non-judgmental awareness, which may enhance participants' ability to reflect on their experiences.

The use of imagery in qualitative research is particularly effective in eliciting emotional and cognitive responses from participants, as it engages them on a personal level (Fereday & Muir-Cochrane, 2006).

By presenting these images first then asking those four questions above after, participants were encouraged to reflect on their own experiences and societal narratives related to discrimination and race (Abraham & Appiah, 2006; Entman & Rojecki, 2001), which aligns with the principles of Critical Race Theory (CRT).

As participants viewed the images, they were likely to draw connections between the visual representations and their own understandings of race, identity, and belonging (Abraham & Appiah, 2006). This process of reflection is crucial, as it allows participants to articulate their thoughts on how these images resonate with their lived experiences and societal observations (Creswell & Creswell, 2014).

Furthermore, the questions posed after the viewing of the images were designed to probe deeper into the participants' awareness of racial issues, encouraging them to consider how these visual representations might influence their perceptions of AI tools and the biases that may be inherent in them (Campbell et al., 2021; Danks & London, 2017; KP, 2024). This approach is supported by the notion that qualitative research benefits from rich, descriptive data that captures the nuances of participants' thoughts and feelings (Carpenter & Lertpratchya, 2016)

The timing of the questions, following the exposure to the images, was a deliberate methodological choice aimed at enhancing participants' engagement with the themes of the research (Creswell & Poth, 2016).

Part of this research is attempting to understand the best ways to utilised AI imagine generative system and support people who are and will use this technology.

What should be the steps in order to mitigate different racial differences/disparities in AI imagery generative systems?

If you are comfortable with it, would you be able to describe your ethnicity, age, gender, religion, and race?

In the above questions, for the first time on the questionnaire I provided, highlights a critical aspect of research involving race (Montagu, 1997). This section reflects the nuanced dynamics of how race can be a sensitive and often subconscious topic for participants (Applebaum, 2023), especially in contexts where the researcher is also a member of a racialised group - I am a visibly African man - (Oyinlade & Losen, 2014). This aligns with discussions in Critical Race Theory (Chapman & de Melo, 2022; Mills & Godley, 2017), which emphasises the importance of understanding how race and racism are embedded in social interactions and institutional practices (Nelson et al., 2013). It is also important to underline that race is a social construct that influences various areas of human endeavour, including education and social interactions (DiAngelo, 2022; Loury & Loury, 2009; Mills, 2014; Montagu, 1997; Oliveira Andrade de Melo & Chapman, 2023).

Furthermore, in my thesis, I have intentionally chosen to address race at a particular point in the research questionnaire to capture the complexities of racial dynamics in participant interactions (Oyinlade & Losen, 2014). The race and gender of the researcher as well as the race and gender of the respondents have being shown to influence respondents answers in research (Ellison et al., 2011; Oyinlade & Losen, 2014). This timing of the questions is a deliberate approach to ensure that I gather authentic and meaningful data, allowing for a deeper understanding of how race influences perceptions and experiences (Enders & Thornton, 2022).

Also, it is important to note that white respondents are often influenced by the interviewer’s race (Finkel et al., 1991). According to (Finkel et al., 1991; Gong & Aadland, 2011) white participants tend to give more open-minded answers to race-related questions when interviewed by Black interviewers rather than by white interviewers (Oyinlade & Losen, 2014). Likewise, there is a trend where respondents feel compelled to provide more racially tolerant answers when the interviewer is of a different race then their own (McDermott, 2011).

3.3.2 The Number of Images

The decision to use three to five images was based on practical considerations, ensuring that the interview could be conducted within the allocated fifty minutes while still allowing for meaningful engagement and sufficient data collection from participants (Schostak, 2005)

3.3.3 Image Choices

The choice of images can significantly impact how participants perceive and discuss race (Cheryl Staats, 2017a; Jones, 2019). In AI, for instance, participants report higher perceived interpersonal closeness with virtual agents that share their racial identity (Cave & Dihal, 2020). The media often presents a limited range of representations that can reinforce stereotypes and biases (Lynn & Dixson, 2013). The images selected for this study may inadvertently reflect or challenge these dominant narratives, influencing how participants of different racial backgrounds engage with the topic of race (Lynn & Dixson, 2013). The participants' backgrounds and experiences with race may shape their responses to the images (Anderson et al., 2023).

Furthermore, racial socialisation plays a crucial role in how individuals understand and interpret racial dynamics, in fact, it is likely that white participants may have been socialised to view race through a lens that emphasises colour-blindness or individual merit, which could affect their reactions to the images presented (Anderson et al., 2023; Bonilla-Silva, 2006). Another point is unconscious bias, where cultural belief systems that participants may hold can lead to unrecognised racism, which, consequently, may manifest in their interpretations of the images and the discussions that follow (Tate, 1997)

My thesis may also be impacted by the power dynamics within the researcher-participant relationship (Oyinlade & Losen, 2014). As a Black male researcher, my own positionality/existence may shape how participants respond to the images and engage with the topics discussed, notwithstanding the historical context of anti-Blackness in western societies (Janvieve Williams Comrie et al, 2022; McKeganey & Bloor, 1991).

Research has a level of acceptable intuition, and those images were selected with this in mind. I chose these prompts that were not racialised at first (meaning they were using non-racial descriptors).

I asked for Australians (adding the world Australian on all prompts) and knowing the population of Australia is composed of 91% European descendants (ABS, 2022), I expected response showcasing that majority. This demographic reality is crucial in understanding the context of this research and the potential biases in AI-generated images[5].

The choice of the first prompt, “Young attractive Australian women looking into the camera,” reproduces an understanding of predominant beauty standards that are often Eurocentric (Li, 2020; Riccio & Oliver, 2022). Societal perceptions of beauty are frequently aligned with whiteness, which can influence the representation of individuals in media and research (Janvieve Williams Comrie et al, 2022). This aligns with the idea that AI tools may reflect these biases, as they are trained on datasets that predominantly feature European-descended individuals (Bender & Friedman, 2018; Benjamin, 2023; Buolamwini & Gebru, 2018).

The next prompt was “Happy Australian school kids playing on school playground looking towards the camera.”

Those two prompts were done very early in the research design phase, and I was a novice in prompting engineering. As, later on, I acquired better skills, the prompts became more complex, that did not alter the results but the quality of the images for a more visually appealing effect. The next three prompts were made using better prompt techniques. “Portrait of an Australian university student working on their assignments within an Australian university surrounded by other Australian university students. Cinematic, clear facial features, 35mm, f/1.8, accent lighting, global illumination, - uplight - V4”.

The next two prompts were attempting to understand to which extent the bias in the AI would manifest also in that they attempt to challenge the dominant narratives and representations. They align with the Critical Race Theory perspective that seeks to highlight and address systemic biases and stereotypes (Lynn & Dixson, 2013) “Portrait of a family of recently arrived Australian immigrants leaving their suburban Australian homes going to pray. Cinematic, clear facial features, 35mm, f/1.8, accent lighting, global illumination, - uplight - V4.” In regarding to the migrant family, I can also

[5] All images were generated using Midjourney. Images will be shown in chapter four.

add that also, ironically, most migrants to Australia every year are from European countries; thus, in their majority, white or white-passing (ABS, 2022).

The other prompt was “Portrait of an Aboriginal Australian working Melbourne, Australia on a busy day, Cinematic, clear facial features, 35mm, f/1.8, accent lighting, global illumination, - uplight - V4.”.

Note that by including these prompts, I was actively engaging with the complexities of race and representation in Australia particularly regarding First Nations (McLane et al., 2022) who also suffer from antiblackness bias (Buolamwini & Gebru, 2018). Also, I made sure to include the world Australia(n) as a constant throughout the research. In my prompts, however, I never mentioned the race (Julia Angwin, 2016), the class (Eubanks, 2018), the ability nor the gender (Drage & Mackereth, 2022; Rosenbaum, 2022) of the prompts of how the images should be generated. I notice, nevertheless, a clear Eurocentric, and male centric bias in the images (Cave & Dihal, 2020; D'ignazio & Klein, 2020).

All people were visually able bodied (Shew, 2020; Whittaker et al., 2019). All with the exception of the migrant family were not considered white or at least visually European, (Bonnett, 1998).

The Aboriginal man was dark skinned and had features that could be seen as historically accurate (Foley, 2000), but it could be questionable if he is actually working and if so, where was he working?

This re-enforces the needs for critically examining these representations and their impact on societal attitudes toward race and identity within Australia (Gatwiri et al., 2021; Moreton-Robinson, 2021; Rennie et al., 2016).

Furthermore, this images appeared this way due to a lack of diversity in training data, which can lead to outputs that fail to represent the full spectrum of human diversity, perpetuating stereotypes and reinforcing existing biases (Angwin et al., 2022; Eubanks, 2018; Ferrara, 2023; Noble, 2018).

3.4 Research Design (NVivo)

The research employed a qualitative design utilising NVivo software for data analysis (Leavy, 2022). This approach is particularly suitable for exploring complex phenomena, such as participants' interactions with AI systems, as qualitative research values depth of meaning and subjective experiences (Leavy, 2022).

NVivo facilitates the coding process, allowing researchers to identify patterns and themes systematically, which is essential for thematic analysis (Creswell & Creswell, 2017). Thematic analysis, is a method for identifying, analysing, and reporting patterns (themes) within qualitative data (Fereday & Muir-Cochrane, 2006).

By employing NVivo, a robust qualitative data analysis software, I can efficiently manage and analyse my interviews transcripts (Creswell & Creswell, 2014). By focusing on themes that emerge from the data, the research can uncover underlying motivations, concerns, and attitudes that may not be immediately apparent through quantitative methods (Vears & Gillam, 2022) Furthermore, thematic analysis aligns with the inductive nature of qualitative research, allowing findings to emerge organically from the data rather than being imposed by pre-existing theories (Vears & Gillam, 2022)

I argue that this flexibility is crucial in understanding the nuanced ways in which participants engage with AI technologies, ultimately contributing to a more comprehensive understanding of the implications of AI in their lives.

By exploring and interpreting the complexities of participants' interactions with AI systems, this study can provide valuable insights that can inform practice and policy in this rapidly evolving field (Agrahari, 2024).

3.4.1 Data Collection

Data for this study was collected through semi-structured interviews, each lasting between thirty-nine minutes and one hour and thirteen minutes, this aligns with best practices in qualitative research, which suggest that longer interviews can facilitate deeper engagement and more comprehensive responses from participants (Campbell et al., 2021).

Semi-structured interviews are particularly effective for exploring complex topics, as they allow for flexibility in questioning while ensuring that key themes are addressed (Leavy, 2022). The recordings were conducted via Riverside.fm (Sparrow, 2022) which is a platform capable of capturing both audio and video. Having quality equipment is fundamental for ensuring the clarity of the data collected (Creswell & Creswell, 2017).

Only audio recordings of these interviews were transcribed for analysis. Using only audio demonstrates my commitment to ethical research practices (Martin, 2022), which, therefore, highlights the importance of protecting participant confidentiality and ensuring that they feel safe sharing personal experiences (Olsen & Mooney-Somers, 2017).

The focus of the interviews questions was on the participants' interactions with AI imagery-generative systems, in this case, an AI tool called Midjourney (Ali et al., 2023; Ali et al., 2024; Ho, 2023; Sukkar et al., 2024). Noting their perceptions of bias is particularly relevant in the context of current discussions surrounding AI ethics and bias (Martin, 2022). By centring the questions on these themes, the study aimed to uncover the subjective experiences of participants, which can provide valuable insights into the broader implications of AI-generated content (Creswell & Poth, 2016).

Moreover, having each of the six participants review the same set of five images allows for a comparative analysis of their interpretations and experiences. This approach not only enhances the reliability of the findings but also facilitates the identification of common themes and divergent perspectives, which are crucial for understanding the multifaceted nature of bias in AI-generated imagery (Creswell & Poth, 2016; Martin, 2022).

Furthermore, by focusing on participants' perceptions of bias or neutrality I am aligning myself with CRT's critique of colourblind ideologies (Cave & Dihal, 2020). CRT posits that claims of neutrality often mask underlying racial biases and power dynamics (Cave & Dihal, 2020; Leonardo, 2004; Sriprakash et al., 2022). By exploring how participants interpret AI-generated images, you can uncover the ways in which these images may reinforce or challenge existing racial stereotypes (Abraham & Appiah, 2006; Bass, 2023; Godsil et al., 2014; Hundt et al., 2022).

3.4.2 Interview Process and Thematic Analysis

Thematic analysis was employed to identify patterns and themes within the interview data. This process involved coding the data using NVivo software, allowing for a systematic examination of participants' responses (Martin, 2022).

3.4.3 Sampling Design

The study involved a purposive sampling design, targeting 06 undergraduate students from Australian universities. The sample size is sufficient to enable a detailed exploration of individual experiences and to gather rich qualitative data (Denzin, 2010).

This approach is particularly effective in qualitative research, as it allows for the selection of participants who possess specific characteristics or experiences relevant to the research questions (Leavy, 2022).

By focusing on students who are currently engaged in university academic programs, the research can capture insights that reflect the contemporary educational landscape in Australia.

The aim is to focus on localised research based in Australia, which is essential for understanding the unique cultural, social, and educational contexts that shape the experiences of these students (Bargallie et al., 2024; Moreton-Robinson, 2004). This localised focus not only enhances the relevance of the findings but also contributes to the broader discourse on higher education in Australia, particularly in relation to the integration of emerging technologies and pedagogical practices (Finn Lattimore, 2020). Localized research can provide insights into the experiences of diverse student populations (Guan & Prentice, 2024).

Furthermore, localised studies can address the ethical implications of AI, ensuring that technology aligns with Australian values and educational goals (Holmes et al., 2021).

In the Australian context, which is characterised by distinct challenges and opportunities, it requires a tailored approach to AI application and usage, ensuring that educational practices are not only effective but also culturally relevant (Striepe et al., 2021).

Qualitative research often emphasises depth over extensiveness, and a smaller sample size can facilitate more detailed and nuanced discussions during interviews (Carpenter & Lertpratchya, 2016)

Moreover, the purposive sampling strategy designed, enabled this researcher, to select participants who are likely to provide diverse viewpoints, enriching the data collected (Bhandari, 2022; Denzin, 2010). By ensuring a range of experiences and backgrounds within the sample, the study can better capture the complexities of student life and learning in the Australian setting (Campbell et al., 2021)

3.5 Theoretical Research Approach

The theoretical framework for this research is grounded in Critical Race Theory. CRT emphasises the importance of understanding how various forms of oppression, including race, class, gender, and technology, intersect (Jones, 2024). This is relevant in this study research, as it allows for a nuanced analysis of how systemic biases affect users from diverse backgrounds (Lynn & Dixson, 2013). CRT highlights the intersectionality of race and racism with other forms of subordination, which can inform your exploration of user experiences across different demographics (Crenshaw, 2017; Lynn & Dixson, 2013)

CRT critiques the dominant narratives that often marginalise the experiences of non-white individuals (Bell, 1995; Ford & Airhihenbuwa, 2010; Ladson-Billings, 2021; Lynn & Dixson, 2013; Tate, 1997; Vass, 2015). By applying CRT, the study challenges Eurocentric perspectives and white supremacist thinking that may influence decision-making processes in technology (Jones, 2024; Oliveira Andrade de Melo & Chapman, 2023).

In my thesis, I explore the impacts of anti-Blackness within decision-making processes, a critical area supported by CRT which provides a robust framework for examining the ways anti-Black racism operates within institutional and societal structures (Janvieve Williams Comrie et al, 2022; Jones, 2024). Anti-Blackness is not merely an individual bias but a systemic issue that permeates social and institutional contexts, influencing attitudes, policies, and outcomes that affect Black individuals and the broader society alike (Janvieve Williams Comrie et al, 2022). For instance, in educational settings, Black students often experience heightened surveillance and punishment when using AI tools (Tanksley, 2024).

By investigating these dynamics, my research aims to contribute to a nuanced understanding of how anti-Blackness shapes decision-making processes, revealing the deep-seated and often subtle ways that racialised biases continue to influence contemporary settings particularly in the area of AI (Bledsoe & Wright, 2019; Crawford, 2023; Janvieve Williams Comrie et al, 2022). Also CRT helps the study to analyse how technological systems may reflect and reinforce existing power dynamics (Eidelman & Crandall, 2012; Tate, 1997).

Furthermore, while my research primarily centres on AI bias and its intersection with race, using CRT enables for an inclusive analysis that considers the experiences of all users, regardless of their racial background (Ford & Airhihenbuwa, 2010). CRT's commitment to social justice and equity, reinforces the importance of amplifying all voices in research analysis (Ford & Airhihenbuwa, 2010; Lynn & Dixson, 2013).

In my research, I employ Critical Race Theory (CRT) not only as a theoretical framework but also as a catalyst for social change (Tate, 1997). By grounding my AI studies in CRT, I align my work with a broader movement aimed at dismantling systemic racism and promoting equity, which is essential in today's socio-political climate (Ford & Airhihenbuwa, 2010; Ladson-Billings, 2021; Lynn & Dixson, 2013; Mills & Godley, 2017). This approach is particularly pertinent given the documented instances where AI systems have exhibited racial biases (Gibney, 2024; Julia Angwin, 2016; Zou & Schiebinger, 2018a). Finally, there is an urgency of addressing these AI biases which is underscored by the need for proactive monitoring and intervention strategies (Sendak et al., 2023). The study employs thematic analysis to identify patterns and themes from qualitative data, aligning findings with CRT's critique of racial power dynamics in AI technologies (Creswell & Poth, 2016). This approach positions CRT not just as a tool for analysis but also as a call for social change, advocating for the development of more equitable and inclusive AI technologies (Bender et al., 2021; Deshpande et al., 2020; Ferrara, 2023; Gebru & Denton, 2024; Gichoya et al., 2023; Tejani et al., 2024).

3.51 Applying Critical Race Theory to the Analysis of AI Bias

As someone who grew up in Brazil and has lived most of my life in Australia, I have the chance to experience living as a citizen of two, very distinct, post-colonies of European powers. Both Brazil and Australia systematically utilised race and racism to dominate and oppress their Black and Indigenous populations (Bennett et al., 2021; Grinberg, 2002; Moreton-Robinson, 2004; Nascimento, 2020; Quijano, 1992; Tuck & Yang, 2021).

I write with the awareness that academia itself is a space from which Black people and Afro-descendants, particularly those living in the Western world, have been historically excluded (Collins, 2022; Ladson-Billings, 2021). While academic papers, books, and theoretical frameworks will be, by definition, cited throughout this thesis (Creswell & Poth, 2016), there is an element of lived experience that must be acknowledged as equally valid and necessary to understanding the arguments presented here (Bell, 1995). My analysis is informed not only by critical scholarship but also by the embodied knowledge of moving through societies structured by coloniality (Wynter, 2003), racial hierarchies (Loury & Loury, 2009), and anti-Blackness (Bledsoe & Wright, 2019). This positioning shapes how I read, interpret, and question the world, and how I understand the gap between theory and reality (Stiglitz, 2012; Tate, 1997). Lived experience, particularly for those from historically marginalised racial groups, is not anecdotal background but a form of expertise (Delgado, 1989; Phillips et al., 2018; Solorzano & Yosso, 2001; Solórzano & Yosso, 2002).

Accepting the orthodox assertions that academia is solely based on objectivity, colour-blindness and neutrality, is not helpful. This assertions only strengthen the self-interests of a few in power, thus, reinforcing the privileges of dominant groups (Smith-Maddox & Solórzano, 2002; Wallace, 2019). Therefore, to ignore my lived experienced, would only allow for replication of the same epistemic erasures that Critical Race Theory and decolonial thought seek to dismantle (Lynn & Dixson, 2013; Smith-Maddox & Solórzano, 2002). Colour-blindness is a racial ideology that asserts race is irrelevant and should not be noticed or used in decision-making. It presents ignoring race as morally or legally neutral, but in practice it denies historical and structural contexts of racial oppression, conflates not seeing race with virtue while enabling maintenance of white supremacy and protecting existing racial advantages and functions rhetorically to avoid acknowledging and mitigating the harms produced by racism (Andrews, 2023; Annamma et al., 2017; Harris, 1993; Loury & Loury, 2009; Lynn & Dixson, 2013; Tate, 1997).

What is Critical Race Theory

Critical Race Theory (CRT) is a scholarly framework that foregrounds race and racism as central to all human interactions. CRT is systematically featured within the law, institutions, and society rather than existing as an isolated individual bias (Delgado & Stefancic, 2023; Solorzano & Yosso, 2001). No one is individually racist, for everyone is. Moreover, CRT offers a powerful lens for interrogating how systems of power, race, and identity operate in contemporary contexts, including the rapidly evolving field of artificial intelligence. CRT challenges claims of neutrality and colourblind objectivity by showing how supposedly “neutral” rules and standards often reproduce white advantage; it centres experiential knowledge and counter stories from marginalised or minoritised groups. CRT also emphasises intersectionality of how race interlocks with class, gender, ability and more (Bell, 1995; Crenshaw, 2017; Tate, 1997).

CRT examines how lived experiences of racially diverse populations reveal the persistence of white supremacy within racialised systems (Bell, 1995; Solorzano & Yosso, 2001; Waring, 2023, 2024). This thesis utilises CRT as its main theoretical foundation, guiding my research questions, methodology, and analysis.

Another aspect of racial relations can be describe as the permanence of racism (Bell, 2018) where, Derrick Bell argues, racism is a permanent feature of USA’s society not as an anomaly but a permanent feature (Bell, 2018). This concept often refers to the structural and systemic nature of racism in western societies, particularly in post-colonial countries. In my thesis, I attempt to showcase that racism adapts to new forms, and because it is created by humans who themselves have biases, AI tools further perpetuate those false hierarchical racial structures (Buolamwini & Gebru, 2018).

Although, it is important to reaffirm that race itself is a social construct (Montagu, 1997) and a human invention sustained by political, historical, and economic power (Bonilla-Silva, 2006; Marger, 2003; Mignolo, 2011).

Furthermore, CRT often rejects colour blindness, placing it not as a benign ideal, but as a racial project that disguises inequality while protecting existing hierarchies (Bonilla-Silva & Dietrich, 2011). Another facet of CRT worth stating is the idea of interest convergence, where it exposes that racial reforms and societal improvements emerge predominantly when they serve dominant interests and not necessarily the oppressed or a minoritised group (Bell, 2018). Therefore, there is no real incentive to improve AI tool from that perspective, therefore, a research like this is vital if we are to ever find solutions for a more fairer AI.

The idea of intersectionality (Crenshaw, 2017) which demands we observe how racial oppression is always shaped by other structures of inequality (Crenshaw, 2017; Stiglitz, 2012) is also particularly vital.

My thesis also utilises anti-Blackness which is a concept that names the specific, global positioning of Blackness as the embodiment of exclusion (Dumas, 2016) as a central pilar for discrimination. Moreover, it can be argued that anti-Blackness is a pervasive, historically rooted system that devalues Black humanity and positions Blackness as inherently inferior. Anti-Blackness often manifests in governmental policies, private and public institutions, and in everyday practices that insist in produce dispossession, criminalisation, and exclusion or darker skin peoples (Gatwiri & Anderson, 2022; Janvieve Williams Comrie et al, 2022). Antiblackness shapes much of western cognition, social interactions, health, education, labour conditions, and surveillance (Benjamin, 2023; Janvieve Williams Comrie et al, 2022; Loury & Loury, 2009). Moreover, antiblackness is the default underlining of artificial intelligence technology (Benjamin, 2023; Noble, 2018; Tanksley, 2024).

My thesis explored the impact of bias in the datasets of a specific tool called Midjourney, a text-to-image AI image generator, on Australian students. My thesis explored the impact of bias in the datasets of a specific tool called Midjourney, a text-to-image AI image generator, on Australian students. My investigation was interested on how students who saw themselves as Australians citizens and how/if this Australianness was tied to race.

Moreover, I was interested in understanding if the anti-Blackness that exists within the AI tools, Large Language Models (LLMs) and algorithms, were also dormant within the research participants. I posit myself with many other race scholars, who argue that all western thinking is based upon racial undercurrents (Andrews, 2023; Ani, 1994; Oliveira Andrade de Melo, 2022; Oliveira Andrade de Melo & Chapman, 2023). The potential for anthropomorphisation of AI tools, by its human users, makes this issue even more dangerous. Investigation into this phenomena is timely and necessary.

Another complementary theory that will be discussed in this thesis is critical whiteness studies which exposes white supremacy influences on everything yet, and importantly, clearly differentiates being white as in an individual identity from whiteness which is a system of dominance (Frankenberg, 2004; Harris, 1993). Nevertheless, although not all white individuals agree with this imposed racial contract, all white persons or people perceived as white are beneficiaries of the spoils of whiteness and white supremacy (Mills, 2014).

Importantly, for this research, although I am well aware that CRT insists on making race visible and often supports the confrontations of racial issues, antagonising so called race-evasive comfort strategies (Annamma et al., 2017), I decided that, during the discussions for this thesis, prematurely engagement in racial discourse would potentially tainted the interviews results. My presence as a Black man was itself sufficient to evoke racial distinction and awareness. This presence likely activated racialised frames in the minds of interviewees (Curry, 2017; Fanon, 1970). I, thus, chose to keep myself as “small” as possible, consciously reducing my visible racial identity performance to mitigate the influence of stereotype threat, racialised expectations, and the reproduction of dominant discourses, a decision aligned with Critical Race Theory’s recognition of how racialised bodies are read and interpreted in social interactions.

Black men, regardless of context or intent, are often read through entrenched stereotypes of danger, hypersexualisation, and aggression (Benjamin, 2023; Curry, 2017; Entman & Rojecki, 2001; Fanon, 1970; Marger, 2003). Within such a landscape, my presence alone could activate these racialised scripts, influencing how interviewees responded (Chapman & de Melo, 2022; Curry, 2017). By consciously minimising my visible racial identity performance, I sought to interrupt these dynamics and reduce the reproduction of dominant, anti-Black discourses in the data I collected (Curry, 2017)

I understand that such avoidance may go against some of theoretical precepts of CRT, but to be efficacious in praxis, I chose to centre white comfort to ensure the success of the research (Cheryl Staats, 2017b). This approach reflects a pragmatic strategy to keep participants open and reduce defensiveness, increasing the chances of gathering candid and nuanced insights (Allemang et al., 2022). While it represents a compromise with some of CRT’s more confrontational aims, it recognises that, in certain settings, centring white comfort can serve as a tactical entry point for encouraging deeper and more critical engagement (DiAngelo, 2022; Leonardo & Gamez-Djokic, 2019; Leonardo & Porter, 2010; Moreton-Robinson, 2021).

Finaly, my research proposes that by focusing on the notions of Australianness tying it to race, pushes this thesis to embark and to intertwine the concepts of coloniality and decoloniality. Racism, thus, is an inseparable technology of power utilised, at least since the so called (European) enlightenment, as the foundation for the forceful, brutal and violent expansion of European powers onto the rest of the world (Mignolo & Walsh, 2018). This means interrogating nationalist narratives that imagine a unified, race-neutral identity, and setting them against racial narratives that centre the voices and histories of those systematically erased (Bernardino-Costa et al., 2018; Havea, 2017; Lopez Andersson, 2024; Matiluko, 2020; Mignolo & Walsh, 2018). A final important observation is the centralisation of whiteness which often causes the otherising of other racial groups as whites are humans and the others are their race (Kilomba, 2021). Othering describes racialisation and the white gaze that racialised and objectifies Blacks and other racial groups as the inferior Other an exercise of power asymmetry that undermines our common humanity (Gatwiri & Anderson, 2022; Loury & Loury, 2009; Maeso & Araújo, 2015; Udah, 2023).

3.6 Research Process

This thesis followed a three-phase research process. First, semi-structured interviews were designed to examine biases in generative AI systems (Creswell & Poth, 2016). Next, Australian university students were recruited and interviewed to gather qualitative data (Bhandari, 2022; Lauri, 2011). Finally, thematic analysis, guided by critical race theory, was applied to explore patterns and the impact of AI biases on student perceptions and decision-making (Lynn & Dixson, 2013; Malagon et al., 2009; Vass, 2015). Moreover, the thematic analysis was done through a CRT lens by applying some of the core tenets earlier outlined in this thesis, including the permanence of racism, intersectionality, critique of colour blindness, race to innocence and counter storytelling as I coded the answers, while inductively identifying emergent patterns (Delgado & Stefancic, 2023).

3.7 Platform Analysis

In addition to interviews, the research will include an analysis of the platforms used for AI image generation. This analysis will assess how these platforms may embed biases in their outputs and influence user interactions.

The thesis utilised Midjourney to create this image. Midjourney is a very popular tool (Hertzmann, 2022; Pawle, 2023; Sukkar et al., 2024).

There are many Images generative tools, such as like DALL-E 2 by OpenAI, they all utilise advanced neural networks to create images based on short text prompts (Economist, 2022; Hertzmann, 2022). These tools have reshaped visual arts by offering accessible ways to depict alternative realities or concepts using Large Language Models (LLM) (Wang et al., 2017). Large Language Models (LLMs) is by definition advanced AI system trained on vast text datasets to understand and generate human-like language. It can perform tasks such as answering questions, writing, translating, summarising, and reasoning across a wide range of topics(Goodfellow, 2016).

Midjourney, in specific, is a generative AI image creator that utilises text prompts and generates images in response to them. It is essentially a type of foundation model that has been trained to an extensive degree through a neural network to learn associations. Foundation models are "multimodal" in that they contain both text and images (Economist, 2022).

For example, if Midjourney can render a Van Gogh image, it has rendered images of Van Gogh enough times to be trained that the text prompt "Van Gogh" aligns with specific techniques and strokes (Economist, 2022).

Furthermore, like other foundation models, Midjourney uses self-supervised learning to generate images; by analysing the vast volumes of image and text data, it learns associations without strictly labelled data guiding it one way or the other, thus generating something entirely original based on a variety of different prompts (Economist, 2022).

3.8 Methodological Rigor

In qualitative research, the concepts of dependability, credibility, trustworthiness, and authenticity are paramount for ensuring the integrity and rigor of the study (Creswell & Poth, 2016).

In conducting my research, I placed a strong emphasis on dependability by thoroughly documenting the research design and processes, ensuring that other researchers could replicate the study (Woo & Heo, 2013). This transparency was vital for establishing the reliability of my findings (Morse et al., 2002; Woo & Heo, 2013).

To enhance credibility, I used triangulation, drawing on multiple data sources and theoretical perspectives, which strengthened the validity of the results (Lauri, 2011). Triangulation involves using multiple data sources or methods to cross-verify findings and enhance credibility (Denzin, 2010; Sinkovics et al., 2008).

Trustworthiness was another key focus. Researchers must actively engage in self-reflection and maintain a critical stance towards their interpretations, which can enhance the trustworthiness of their findings (Creswell & Poth, 2016; Morrow, 2005). Overall, this involves not only accurate data collection but also an ethical commitment to representing participants' voices authentically in the research narrative (Peck & Mummery, 2022). As posited by (Creswell & Creswell, 2014), qualitative research focuses on understanding participants' perceptions and experiences, which may not be universally applicable.

Lastly, I maintained a reflexive attitude throughout the research process, recognising my own positionality (Bourke, 2014) and its influence. This self-awareness contributed to the authenticity of my research and helped to manage potential biases in interpreting the data.

3.8.1 Ethical Considerations

Ethical considerations are paramount in this research. The study will adhere to the Australian Code for the Responsible Conduct of Research, ensuring participant confidentiality and informed consent (Zion, 2023). All data will be stored securely, and participants will be made aware of their rights throughout the research process (Anderson et al., 2023; Bailey et al., 2021; Bhandari, 2022; Oliver-Hoyo & Allen, 2006).

Methodology in research reflects integrity by aligning the researcher's skills with the research design, ensuring the study's value is justified and contributes meaningfully to the advancement of human knowledge (Bhandari, 2022; Martin, 2022).

In my study, the benefits of understanding the influence of bias in imagery generative systems will allow us to further investigate the effect of bias in AI on individuals (Campbell et al., 2021). This study hopes to contribute to the extensive existing literature on AI bias everyday more ubiquitous part of modern lives.

3.8.2 Pseudonyms

All names of participants used on this research are pseudonyms (Allen & Wiles, 2015). The use of pseudonyms in qualitative research is essential for safeguarding participant confidentiality and ensuring ethical compliance (Allen & Wiles, 2015). Assigning pseudonyms allows researchers to protect the identities of participants, which is particularly crucial in sensitive studies involving vulnerable populations or topics that may expose individuals to stigma or harm (Anderson et al., 2023; Tate, 1997). As race, and ethical societal issues were discussed, the use of pseudonyms is appropriate (D'Arrietta et al., 2022; Hancock et al., 2018; Robinson-Wood et al., 2020). There are also ethical implications of using pseudonyms, which beyond an simple act of confidentiality; it also reflects a commitment to respecting participants' autonomy and identity (Brear, 2017; Creswell & Creswell, 2014).

Allen & Wiles (2015) arguments in favour of that it is a good idea to choose pseudonyms that align with the participant's cultural background and the research context, therefore maintaining authenticity and relevance in the study (Allen & Wiles, 2015).

3.8.3 Data Safety

All data will be collected using Riverside.fm (Sparrow, 2022). Audio was transcribed and fed into NVivo for Thematic analysis (Fereday & Muir-Cochrane, 2006). Once data is analysed, it will be stored and kept for the duration required, in the provided Victoria University safe drive.

3.8.4 Occupational Health and Safety Risks

No occupational health or safety risks have been encountered in this research.

3.8.5 Limitations of the Study

The study acknowledges several limitations, including the small sample size, which may affect the generalisability of findings. In addition, because this is a qualitative study this may limit the ability to draw broad conclusions applicable to all users of AI imagery generative systems. Future research could expand on these findings by including a larger and more diverse sample.

Chapter 4: Results and Findings

4.0 Themes Results

In this chapter we look into the themes analysed and how to make sense of them. Can the bias embedded in image generative systems influence public perception? If this is the case, what does this mean?

Two prominent themes emerged among participants during our qualitative semi-structured interviews: “Who is considered Australian” and the “Perceptions of Australianness and Race.”

4.1 Who is considered Australian?

The exploration of identity and national identity, particularly within the Australian context, reveals a complex interplay of historical, societal, and political factors that shape perceptions of what it means to be "Australian (Moreton-Robinson, 2004)."

After my study commenced, a recent tabloid article from the “Daily Mail Australian” explored similar questions of Australianness using the very software I used on my research: Midjourney (Pawle, 2023). The article used prompts like “Sydney's stereotypical woman in front of the Harbour Bridge” (Pawle, 2023). The AI-generated outputs were strikingly Eurocentric, reflecting a narrow vision of identity and making historical inferences that were, at best, questionable (Cave & Dihal, 2020). It can be argued that these images seemed to align more with colonial ideals than with the multicultural reality of contemporary Australia (Bennett et al., 2023; Sriprakash et al., 2022; Wolfe, 2006).

Lynn & Dixson (2013) puts forward the view that cultural knowledge and identity can shape individuals' understanding of their place in the world, this fact emphasises the importance of lived experiences (as we going to notice from the participants responses) and cultural heritage in defining identity (Lynn & Dixson, 2013).

Identity and national identity vary significantly across countries and cultures, shaped by historical, social, and political contexts (Leonardo, 2004; Mitchell-Walthour & Morrison; Pillay, 2020). In the Australian context, the definitions of who is "Australian" and what constitutes "Australianness" are fluid and depend heavily on the perspectives of those asking and those answering (Guan & Prentice, 2024; Moreton-Robinson, 2004; O’Keeffe, 2024). For recent migrants, for example, "Australian" often resonates with values such as fairness, opportunity, and community (Markus, 2023). In fact, this aligns with findings in the "Mapping Social Cohesion 2024" report, which suggests that a significant portion of Australians value multiculturalism and the contributions of migrants to society (Guan & Prentice, 2024). Furthermore, these values are frequently highlighted in public narratives of multiculturalism (Guan & Prentice, 2024). In the other hand, for many older Australians, particularly those born in the country, characteristics such as speaking English (Guan & Prentice, 2024) and participating in culturally specific practices like mateship or sporting traditions are viewed as quintessentially "Aussie" traits (Hirst, 2016).

The perspective of First Nations peoples introduces another critical dimension (Balla, 2020; Moreton-Robinson, 2021; Phillips et al., 2018). As the original inhabitants of the land (Harari, 2014; Naden, 2017; Nasir et al., 2018), Aboriginal and Torres Strait Islander peoples present a profound narrative predating colonisation, challenging settler-colonial frameworks of national identity (Bennett et al., 2023; Council, 2017). Their sovereignty, deeply rooted in spiritual connections to land sustained over 65,000 years, is central to their cultural identity (Council, 2017). This enduring connection, articulated, for illustration, in the Uluru Statement from the Heart, contrasts with dominant national narratives, raising questions about what it means to be the "first

Australians" (Council, 2017; Moreton-Robinson, 2004). Moreover, (Balla, 2020) highlights how First Nations women in the art industry/community, express their identity through creative practices asserting, in her view, sovereignty. This, (Balla, 2020) proposes, consequently, challenges colonial representations.

These diverse points of view emphasise that the definitions of "Australian" are deeply personal and vary widely, shaped by lived experience, cultural heritage, race, gender, and social positioning (Bennett et al., 2021; Gatwiri et al., 2021; Maire, 2021).

In my study, the terms "Australia" and "Australian" were embedded in the research methodology and questionnaire, shaping the interpretations and discussions of national identity among the participants. National identity is a dynamic interplay of individual and collective experiences (Pillay, 2020). Unsurprisingly, this theme emerged prominently across all six participants responses, reflecting the centrality of these terms in their understanding of self and other (Collins & Stockton, 2018; Vears & Gillam, 2022). In the context of AI bias embedded in the Image generative system use, I explored, in the research, the way participants interacted with and interpreted these terms and images was reflective of their own identities, histories, and positionalities within Australian society (Crawford, 2021; Garcia, 2008).

Furthermore, this reinforces that the construction of national identity is not static but a dynamic interplay of individual and collective experiences (Guan & Prentice, 2024; Moreton-Robinson, 2004).

4.1.1 Pseudonyms

Encouraging participants to self-identify their cultural and racial backgrounds in qualitative research fosters inclusivity and respect (Campbell et al., 2021). This approach enriches data collection by providing deeper insights into participants' perspectives and experiences. Moreover, it helps researchers navigate their positionality, addressing insider and outsider dynamics in the research process (Campbell et al., 2021). This study will use the pseudonyms with their correct age and cultural self-identification.

All names used in this study are pseudonyms, carefully chosen to reflect the cultural or ethnic backgrounds of the participants (Campbell et al., 2021). These names are entirely fabricated to ensure the anonymity and privacy of those involved while adding depth and relatability to the discussion (Campbell et al., 2021).

4.1.2 Images and Responses

Two key themes emerged during the qualitative semi-structured interviews: "Who is considered Australian?" and "Perceptions of Australianness and Race." When participants viewed the generated images (1–5), these themes were central to the discussion, highlighting the nuanced ways generative AI systems intersect with identity and representation.

4.2 Generated Images

Support graphic 4

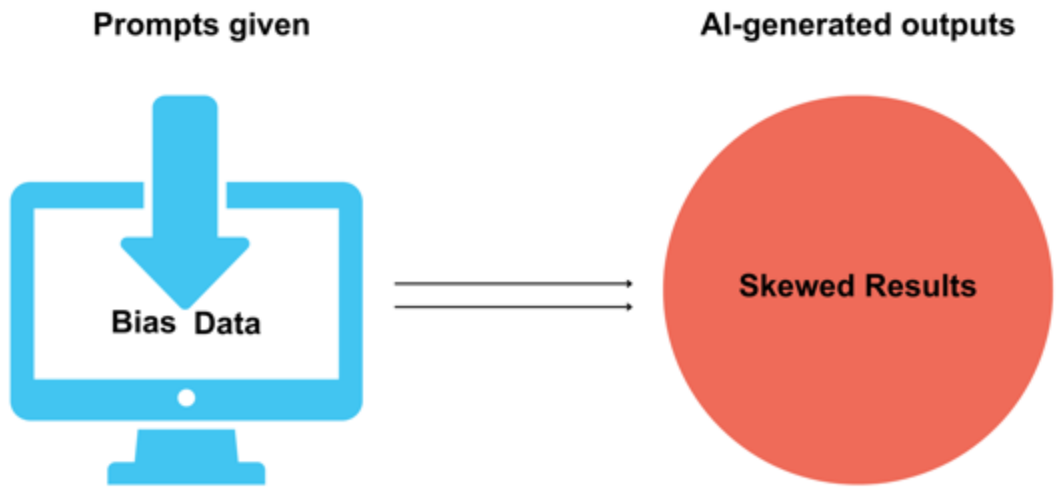

Figure 5

4.2.1 Young attractive Australian women looking into the camera

Generated image 1)

Figure 6 (Generated image 1)

Young attractive Australian women looking into the camera.

“Yeah, I think so. I would expect that an AI generation would work based off stereotypes and preconceived notions.” Ellie.

In this statement above by participant Ellie Thelogos, self-identified as 19, Maltese, Greek, Australia and without a religion. Ellie Thelogos will be henceforth referred to as Ellie. In her comment, there is an evident expression of distrust towards the AI tool, accompanied by a noticeable sense of low expectation towards AI tools (Farahani & Ghasemi, 2024; Finn Lattimore, 2020). These distrust may be cause by lack of understanding of knowledge about the tool (Albarrán et al., 2021).

“I think they got the young attractive woman part, but I don't know if they would like fit into the stereotypical like Australian what people would initially think of Australian I think they're probably more tan and blonde and maybe less serious, people I think, like an Australian stereotype is that it's” Ellie.

In this next quote above, Ellie agrees that the women depicted are perceived in her view as beautiful, aligning with the European standards of beauty that remain predominant and deeply ingrained within Western contexts (Li, 2020; Riccio & Oliver, 2022). Furthermore, Australia as a majority culturally European nation (ABS, 2022), is under the Euro standard concept of beauty which, in turn, is deeply influenced by cultural standards, particularly those rooted in Eurocentric ideals (Strully, 2014). In the Australian context, the leisure and beach going life style, can often be associated with the Eurocentric white view of Australia (Moreton-Robinson, 2004). As a counter point to this, (Balla, 2020) discusses the ideological aspects of settler colonialism (Bennett et al., 2023) in Australia, which can be relevant when considering how beach culture has been shaped by historical narratives and the appropriation of Indigenous spaces (Moreton-Robinson, 2004).

“See, my brain, I thought it was gonna generate some kind of like tanned blonde woman on a beach.” Jaymie.

With participant Jaymie Cooper who is 26, Aboriginal, who is an Aboriginal Australian and of European descent (henceforth Jaymie). In her statement, again, above, we notice that the discussion here is not if the “young attractive Australian women” are or not beautiful - western standards of beauty - (Li, 2020; Riccio & Oliver, 2022).

In her comment, the race of the women in the image is also not discussed, this, I propose, is denoting the normalisation of whiteness (Nagar & Virk, 2017; Robinson, 2011).

However, the discussion is if they are too white for what the participant perceives as the Australian look. This hyper-whiteness is discussed by (Sverdljuk et al.). This, therefore, denotes potentially, the participant's bias, the perceptions and imagery expectations what Australians ought to be (Markus, 2023). Hyper-whiteness is the idea that there are superior types of whites even comparing to themselves as whiteness is not merely a racial identity but a system of beliefs and practices that uphold racial hierarchies (Lynn & Dixson, 2013; Sverdljuk et al.).

"I think it leaves a lot to interpretation, like what we might think of how an Australian might look versus how we see the images of the girls as well. Like they almost do look a bit Scandinavian as well. So, whilst it is correct, I think it could be very general, like it doesn't provide the full image" Jaymie.

Above, again, Jaymie furthers her views, and returns to the fact that they are too white citing Scandinavian people as hyper-white to illustrate her point (Sverdljuk et al.). She is questioning what Australian may look like but not mentioning race, nor she cites First Nations People who are the first inhabitants of this land (Council, 2017; Naden, 2017).

Although it is important to emphasise that the majority of Australians are indeed white or have European ancestry (ABS, 2022), nonetheless, in this case, we see how the bias towards white European standards of beauty in the machine may influence users to accept the reality presented. In fact, the study by (Nagar & Virk, 2017) found that exposure to media images, even for as little as one minute, can significantly enhance the internalisation and acceptance of socially constructed ideals.

"one's mind about us. Like in saying movie references, there is [are] always that image of an Australian's blonde hair, blue eyes, like surfie style. But that's just… and then that is going to permeate in their minds, but that's just an image that's created by them." Jazmine.

Jazmine Pinto, 23, Australian, Indian, Catholic (henceforth Jazmine), on the above comment, confronts the idea of Aussie surf culture, once again on the motifs of surf (tunned), blonde, blue eyes (Hirst, 2016). However, as a member of an outgroup in the Australian context (Lizzio-Wilson et al., 2022), she using this surf ideal image to describe what is, in her view, an imagery of Australian created by AI (Nicoletti et al., 2023).

“Yeah, I feel like it's very stereotypical with that kind of result. It's usually, I feel like if you want a person of colour, you have to add that descriptor. But if you don't add a race descriptor, it will give you a person of European descent by default.” Lorna.

Above, another participant; Lorna Kwame, 32, African, Black, no religion (henceforth Lorna), who, incidentally, has utilised AI image generative before, and who is a non-European participant, showcase her understands, know-how and personal experience, that if one wants to get a person that is not white using Image Generative systems, one must use “race descriptors” (Srinivasan & Uchino, 2020).

This showcases that, as a member of an outgroup in the Australian context (Lizzio-Wilson et al., 2022), Lorna, learned to cope with the bias in the tool by bypassing it using learned techniques (Oliveira Andrade de Melo & Chapman, 2023) in order to force the tool to showcase images more suited to her needs (Nicoletti et al., 2023).

This shows how, for a person’s outside of the dominant group, race can play a significant part of their daily interactions (Skinner-Dorkenoo et al., 2021). Whilst, for members of the ingroup, those things, at first pass unnoticed (Applebaum, 2022).

“Yeah, I'd expect it to churn out something like this. Yeah.” Lorna.

This above is showcasing resignations for images showing white images tropes (Bass, 2023; Nicoletti et al., 2023), where Lorna, blurts this out as that is something she seen before (Jones, 2019).

“To me, the Australian, when the first one, I probably wouldn't use the same looking girl three times. Like it literally looks like they're triplets. I think for me, like Australia is, you know, obviously first and foremost, it's the land of Indigenous people, but then we're talking about all the people that have moved to Australia and we're talking, I mean, obviously. white, English, Irish, but we're talking a lot of, in Melbourne at least, a lot of ethnicities and a lot of different, you know, backgrounds. And I don't think that picture really portrays that. And I think if you were to show it to someone that was, you know... that was Greek Australian or whatever, they'd be offended. Like they'd be like, oh, it doesn't, you know, it's a real one mind kind of photo.” Rachel.

Rachel Mitterrand, 30 White, French born, Australian citizen, Spanish heritage, Atheist but Catholic by family (henceforth Rachel), suggests the three women to be very similar in looks in her view. She mentions Aboriginal people, arguably, in an effort to share her positionality as a good white person (DiAngelo, 2022; Picower, 2009). Moreover, many white individuals often frame their racial emotions and experiences in a way that seeks to absolve them of guilt or responsibility for systemic racism (Leonardo & Gamez-Djokic, 2019; Picower, 2009). White individuals may also feel the need to curate their identity in a way that is acceptable to non-white communities, often leading to a performative aspect of their racial identity (Gatwiri & Anderson, 2022).

She then acknowledges Australia's identity as a predominantly Anglo-Saxon white (ABS, 2022) nation but contrasts this with the multicultural reality of Melbourne's urban life (Forrest & Dunn, 2010).

Critiquing the imagery, she notes its failure to reflect this diversity, pointing out that the women depicted ginger-haired, blonde, and uniformly blue-eyed—align narrowly with a specific European aesthetic (Nagar & Virk, 2017; Robinson, 2011; Ryan-Mosley, 2021). To emphasise her discontent, she draws attention to another European group, the Greeks, as an example of diversity and ranking even within whiteness (Sverdljuk et al.). Speaking from her position as a white European woman, she articulates her dissatisfaction with the biases inherent in these generative images, which omit the broader spectrum of identities present in Australian society (Guan & Prentice, 2024).

This first image presented, the idea of what is Australian and how that should be portrayed or manifest by machines is a question hard to answer(ABS, 2022; Austin & Fozdar, 2018; Bennett et al., 2021; Gatwiri et al., 2021; Lizzio-Wilson et al., 2022; López López, 2024; Markus, 2023; O’Keeffe, 2024). Rather than contesting this post-colonial vision of national identity, numerous participants appeared to accept and even embrace it, reinforcing a narrative that marginalises the plurality of lived experiences within Australian society (Forrest & Dunn, 2010; Wolfe, 2006).

4.2.2 Happy Australian school kids playing on school playground looking towards the camera.

Figure 7 (Generated image 2).

Generated image 2) Happy Australian school kids playing on school playground looking towards the camera.

Jaymie said:

“I think, yeah, I think it's a bit of a mix maybe because there are places where, you know, it's the majority is white, but that's changing. I'd probably say from, yeah, that might not be entirely true.” Jaymie.

In the sentence above, Jaymie reflects on her understanding of how Australian society should be depicted (O’Keeffe, 2024), critiquing the narrow and homogenised portrayals often associated with Australianness. The hegemonic Australian identity is shaped by British colonialism, prioritising masculinity, Christianity, English-speaking norms, and heteronormativity (O’Keeffe, 2024). Her observations highlight the need for representations that move beyond reductive post-colonial narratives and engage more meaningfully with the diverse and multifaceted realities of Australian social life (Sharples & Blair, 2021).

“so, you ultimately feel this image is not accurate in your view.
“No, I'd say an average Australian school would be more diverse than this.” Lorna.

On the other hand, Lorna adopts a far more confrontational attitude, challenging dominant narratives (Balla, 2020) asserting unequivocally that, in her view, this representation is inaccurate, which supports the notion of a confrontational attitude towards misrepresentation (Bargallie et al., 2024).

“Yeah, okay, so it's a photo of nine white children playing in a playground.” Lorna.

For the first time in the study, I observe a participant explicitly using the term white to describe what are visually white-appearing children (Leonardo & Gamez-Djokic, 2019). African immigrants to Australia make significant efforts to belong (Hatoss, 2012). This is likely due to the participant's positionality as a member of an outgroup; as an African participant, they may be more attuned to recognising and articulating such racial markers (Cave & Dihal, 2020; Leonardo, 2004; Oliveira Andrade de Melo & Chapman, 2023).

“in terms of when I went to school. I was the only African person, but I was not the only person of colour. And so, from that perspective, I'd say there was more diversity than depicted in the image. And even now when I think about my... younger siblings who are in high school, their cohort looks a lot more diverse than what was depicted in the image." Lorna.

She continues, drawing on her personal experiences, to challenge the AI-generated visuals of an all-white schoolyard (Cave & Dihal, 2020). Her critique focuses less on whether the image could reflect a reality and more on the absence of transparency or justification within the AI tool for its racialised choices, which further emphasises that discussions about race must include an understanding of the historical context and the ongoing implications of racial dynamics (Cave & Dihal, 2020; Nelson et al., 2013). This lack of accountability, she argues, reinforces exclusionary narratives and leaves no space for questioning or contextualising the bias embedded in the generated image (Leonardo, 2004; Leonardo & Porter, 2010).

"So yeah, I think the most accurate one was probably this playground because that reminded me of like the playground that I went to school, like the colouring of it, the school uniform reminded me of the one that I had. But yeah, in the other images, I would say like the clothing and the colouring of the images was not like authentically Australian to me." Ellie.

Ellie, by contrast, shifts her focus to other aspects of the image, such as the setting, school uniforms, and the playground. For her, race and gender appear irrelevant on her description, despite the image depicting an all-white, predominantly male group of children, even though the prompt did not specify either characteristic (Tate, 1997).

The available evidence seems to suggest that AI-generated images often reflect and reinforce existing racial hierarchies (Cave & Dihal, 2020). It is important to note that there is no definitive right or wrong way to interpret these images (Entman & Rojecki, 2001). Moreover, the subjective nature of human interpretation, leaves room for various interpretations (Brooks, 1999).

My observations are made through the lens of critical race theory (Bell, 1995), highlighting the embedded assumptions within the AI-generated visuals based on the dataset collected for this study (Sukkar et al., 2024).

4.2.3 Portrait of an Australian university student working on their assignments within an Australian university surrounded by other Australian university students.

Figure 8 (Generated image 3).

Generated image 3) Portrait of an Australian university student working on their assignments within an Australian university surrounded by other Australian university students. Cinematic, clear facial features, 35mm, f/1.8, accent lighting, global illumination, - uplight - V4.

In this new image above, Ellie says:

“Yeah, I feel like even like the clothing is maybe a bit more American, but it could definitely pass as Australian style. But yeah, I think it does match the pump.” Ellie.

Once again, her interpretation centres on the observation that he may be dressed in a manner reminiscent of an American (USA), yet she does not address the fact that he appears white or male. I put forward the claim that failing to acknowledge racial and gendered markers can perpetuate dominant narratives (Crenshaw, 2017; Tate, 1997). This selective focus underscores an interpretive lens that overlooks racial and gendered markers, which remain significant from a critical race theory perspective (Bell, 1995; Lynn & Dixson, 2013).

“something along those lines. And then the guy in the uni, very, very, these people could very well be Australian, but like he looked like Greek or something. But that's Australian to me. So, it's like kind of this interesting thing where maybe AI is actually picking up on that Australian can mean more than just this Anglo looking white person from the farm.” Helena.

Helena Podrike, a 36-year-old Serbian, White Australian, female, and Christian Orthodox (henceforth Helena), observes, in the above comment, that the male student in the image does not align with her perception of the construction of Australian identity has historically been aligned with Anglocentric whiteness, placing whiteness at the centre of how national identity has been moulded (Moreton-Robinson, 2004).

She further articulates a clear vision of what she believes an Australian should look like, implying that Australianness is, or ought to be, represented by an Anglo-Saxon appearance (Moreton-Robinson, 2004). Her views are in line with a significant portion of the population that views being born in Australia and adherence to Christianity as important (Guan & Prentice, 2024).

This perception reproduces a widely held narrative in Australia, one that continues to perpetuate exclusionary ideas of national identity rooted in colonial constructs (Wolfe, 2006).

Wolfe's (2006) perspectives, helps my study to contextualise Helena's views within a broader narrative of exclusion rooted in colonial constructs.

"Because like we all migrated here and like, you know, I'm first generation Australian, but like, you know, a lot of people in my life are also first generation Australian. And, you know, I think that's, I think they came here for a better life and Australia welcomed my family, to be honest with you. The war happened and they said, come to Australia. We've got jobs for you. So that's how my dad came here on a boat and my mom came here on a plane. And that's how we ended up here." Helena.

In line with Moreton-Robinson (2004) exploring the historical context of Australian identity of how certain migrant communities, particularly those from Southern Europe, have been positioned within the national narrative (Moreton-Robinson, 2004). Helena considers Greek Australia's as the right type of "Australians" drawing by association a positive comparison to her Bosnian heritage as a first-generation Australian born to migrant parents.

This highlights that while there is support for multicultural diversity, there are also distinctions made regarding who is considered "truly Australian (Guan & Prentice, 2024)." She states, "I think they came here for a better life and Australia welcomed my family, to be honest with you." This statement subtly implies a dichotomy. Where the dominant white Australian culture has historically valued certain ethnicities over others (Moreton-Robinson, 2004).

In her view, some groups are perceived as contributing through work and are welcomed, while others, by exclusion, are not afforded the same acceptance. This is in line with (Sharples & Blair, 2021) who highlights how certain groups are framed as contributing positively to the national identity, while others are viewed as threats.

"Yep, a young white male sits at a desk. He's wearing a grey sweater. You can see bits of writing in a kind of purple, I mean blue colour. And he's wearing a denim jacket that has different sleeves. One is like... Puffy, the other one is, I don't know, it's hard to tell.” Lorna.

Above, Lorna once again describes both the individual's race and clothing. This observation highlights that, as a member of the outgroup, Lorna is particularly attuned to racial markers and their contextual significance (Bargallie et al., 2024; Sharples & Blair, 2021).

4.2.4 Portrait of a family of recently arrived Australian immigrants leaving their suburban Australian homes going to pray.

Figure 9 (Generated image 4).

Generated image 4) Portrait of a family of recently arrived Australian immigrants leaving their suburban Australian homes going to pray. Cinematic, clear facial features, 35mm, f/1.8, accent lighting, global illumination, - uplight - V4.

“Her mother next to her stands a little bit taller and she's wearing a brown floral headscarf and a pink flowy shirt. They both have quite serious expressions on their face. They both have brown eyes and olive toned skin. And in the background, you can see there's a house on the left which” Ellie.

In the above comment, Ellie, for the first time, explicitly references the skin tone of any individuals shown in the generated images (Knowles et al., 2014). Ellie speaks of the adult woman (possibly a mother) and the child (possibly a daughter).

White individuals often exhibit a tendency to acknowledge the racial identities of others while simultaneously minimising or denying their own racial identity (DiAngelo, 2022; Knowles et al., 2014) .

This shift is noteworthy, as in earlier images she made no mention of skin colour (Kilomba, 2021). The idea that whiteness represents the default individual while others are labelled as ethnic ignores the reality that all humans, including white people, belong to an ethnicity (Aspinall, 1998; Bonnett, 1998; Kilomba, 2021; Montagu, 1997). This act of othering is critical for understanding both the AI's construction of centrality of whiteness and how Ellie's perceptions align with these representations of reality (Cave & Dihal, 2020).

“Well, they look not Australian. They look okay. Because the prompt said Australian immigrant family. We'll go with that. Yes, they do look like an Australian immigrant family. And yeah, I'm not too sure what else was in the prompt, but at the suburban home, yes. And going to pray? Don't know. I don't know if that's something that... Sure.” Helena.

Here, Helena distils her assumptions and prejudices regarding who qualifies as Australian (Dunn et al., 2004). She unequivocally asserts that the pair of females in the image are not Australian (Dunn et al., 2004), further describing them as looking like an immigrant family (Lizzio-Wilson et al., 2022; O’Keeffe, 2024). A description overloaded with a tone of derogation and othering (Bell, 1995; Lynn & Dixson, 2013). She even questions their act of praying, a concept seemingly so unfamiliar to her worldview that it becomes alien (Applebaum, 2022; Mills, 2014). This highlights the inherent risks of AI generative systems, where dormant biases in users are not only surfaced but amplified when seemingly validated by the AI tool, reinforcing exclusionary narratives (Bass, 2023; Belenguer, 2022; Bolukbasi et al., 2016; Gichoya et al., 2023; Nadeem et al., 2022; Whittaker et al., 2019).

“They look more of an immigrant family than an Australian family. If we were going to like a typical look of what that would look like to someone.” Helena.

She employs the term someone to articulate what she perceives as the norm (Knowles et al., 2014), framing her opinion as reflective of a broader consensus (DiAngelo, 2022). Furthermore, Helena describes the family as foreign based on assumptions about their cultural background (possibly Muslim) and their race -not white- (DiAngelo, 2022). This act of othering is both problematic and inaccurate, as data from the Australian Bureau of Statistics (ABS) shows that most migrants to Australia originate from European countries and are often perceived as white (ABS, 2022).

This reflects a broader issue in Australian political discourse, where immigration is frequently framed as a threat to the economy and societal well-being (Guan & Prentice, 2024). However, in reality, immigration overwhelmingly contributes positively to the lives of established citizens, challenging these exclusionary narratives (Guan & Prentice, 2024; Markus, 2023).

“And it'd be like, well, because they're not like, they weren't born here. They're not Australian. And...” Helena.

Helena further asserts, without hesitation, that the pair of women are not born in Australia (O’Keeffe, 2024), reinforcing her entrenched perceptions of who belongs and who does not. This unwavering stance highlights the deeply rooted biases shaping her view of Australianness (Austin & Fozdar, 2018; Phillips, 1998).

“I guess like as an Australian when you think about showing me a photo of an Australian woman or something even I have a bias of what I think someone, not what I think every Australian looks like but what I think other people will think the standard Australian looks like.” Jaymie.

The earnestness in Jaymie's comment above, underscores the complexities and fluidity of identity in Australia (Austin & Fozdar, 2018; Phillips, 1998). She acknowledges her understanding of how others might perceive and judge the two women in the image as non-Australians, openly admitting to harbouring similar biases despite her desire to overcome such feelings (Tate, 1997). I posit that this perspective supports the idea that individuals may need to grapple with their biases while seeking to challenge and change these perceptions (Bell, 1995). This self-awareness reveals the pervasive nature of these biases (Bell, 1995; DiAngelo, 2022).

“It does, it depicts a, I believe a mother and a young daughter who is staring directly into the camera. They've both got, I think, hijabs on and they look quite...” Jazmine.

Meanwhile, Jazmine, a woman of Indian Australian background, offers a perspective that contrasts significantly with many of the descriptions above (DiAngelo, 2022; Moreton-Robinson, 2021). She identifies the pair as a mother and daughter, aligning with the prompt's request for a family, and notably refrains from mentioning their skin colour (Leonardo, 2004). Instead, she highlights the obvious yet overlooked detail that they are wearing hijabs (Dunn et al., 2004). A significant aspect of their imagery that no other participant, hitherto, acknowledged. This omission by other participants underscores the selective framing of observations and the biases that influence what is considered noteworthy (Bonilla-Silva, 2006; Bonilla-Silva & Dietrich, 2011).

“a portrait of a mother and child, a woman and a child. The woman is wearing a brown haired scarf with patterns on it. And the child is wearing a... greyish head scarf. They're both wearing pink dresses, but the child's dress has got dots on it, like small blue dots on it. She's wearing another scarf on top which is like a green brown and images taken sort of.” Lorna.

Once again, the only African Australian participant, Lorna, furthers the normalisation of the image by neither mentioning the subjects' race nor presuming they are wearing a hijab (Levin et al., 2003). On the other hand, the description of the headscarf as simply a "head scarf" rather than a hijab reflects a potential oversimplification that can obscure the cultural significance of such symbols, particularly in discussions about identity and representation (Lynn & Dixson, 2013).

“I think, if I'm not, my understanding is that there's been more recent immigrants from Europe to Australia in recent years. The demographics of people migrating to Australia has changed a little bit. And I don't think the average family would look like that.” Lorna.

“Um, I think either that it might have the stereotype bias because maybe it's only been shown images of a, like... of immigrants who mostly look like that and not the vast array of immigrants that come to this country not just as refugees but in other visa classes so I think it might have a bias because... people with my greatest trailer from all sorts of places.” Lorna.

Exhibiting her understanding of Australian society, particularly in terms of migration and demographics, Lorna highlights the fact that migration to Australia has historically been, and continues to be, predominantly European and white (ABS, 2022). This awareness possibly allows her a more critical engagement with societal structures and the development of strategies to combat discrimination and inequality (Leonardo, 2004). Furthermore, knowledge serves as a tool for empowerment (Janvieve Williams Comrie et al, 2022).

This is a reality often overlooked by those within the societal ingroup, which, in the Australian case, are the European descendants (Applebaum, 2022; Dunning, 2011; Maeso & Araújo, 2015; Nelson et al., 2013). Awareness of such aspects of society tends to vary, with those who have the most to gain or lose from these dynamics often possessing greater insight (Janvieve Williams Comrie et al, 2022).

“I mean, yes. I mean, like for me, it's like you tell me they're migrants. I'm like, yeah, I mean, I would believe I would say yes.” Rachel.

Rachel, in the above comment, illustrates one of the central argument of this thesis: while she may harbour biased sentiments, she expresses doubt about their fairness, justice, or whether she would hold those bias without external influence exposing her to more bias (DiAngelo, 2022; Kite et al., 2022; KP, 2024; McGhee, 2022). However, as the AI image generative system nudges her with biased depictions of individuals (Nicoletti et al., 2023), she seemingly into falling into these bias as she states above, (Skitka et al., 1999) ultimately, accepting the two women as migrants (Bass, 2023).

This highlights how AI systems can reinforce and amplify dormant prejudices, shaping perceptions in subtle nevertheless powerful ways (Bass, 2023; Belenguer, 2022; Bolukbasi et al., 2016; Drage & Mackereth, 2022; Ferrer et al., 2021; Mbalaka, 2023; McAra-Hunter, 2024; Nadeem et al., 2022; Nicoletti et al., 2023; X. Wang et al., 2023; Williams-Ceci et al., 2024).

Confrontation is not a preferred solution when dealing with racist thoughts or racist actions, for societal pressures to avoid appearing racist can lead white individuals to focus on their personal feelings rather than the broader implications of their actions and beliefs (Entman & Rojecki, 2001; Leonardo, 2004).

4.2.5 Portrait of a family of recently arrived Australian immigrants leaving their suburban Australian homes going to pray.

Figure 10 (Generated image 5)

Generated image 5) Portrait of an Aboriginal Australian working Melbourne, Australia on a busy day, Cinematic, clear facial features, 35mm, f/1.8, accent lighting, global illumination, - uplight - V4.

“There is an Indigenous Australian man taking up majority of the frame. He's a little right to the centre and he's staring not at the camera, just a little bit past it quite seriously. He looks to be a bit older. He's got Black hair but is a bit of a grey beard and grey moustache.” Ellie.

Ellie highlights the presence of an Indigenous man looking directly at the camera, marking a shift in her observations (Tate, 1997). Whereas previous individuals were described simply as students, men, boys, or women, this time she explicitly acknowledges his cultural background, adding a layer of identity to her interpretation (Kilomba, 2021). The acknowledgment of cultural identity counters the notion of colour-blindness (Lynn & Dixson, 2013).

“I guess there's the one produced of like the Aboriginal man, the kids on the playground, there seems to be potentially a blend of races. But I think overall it's not just the balance of the races and what truly reflects Australia…Yeah, I think I'd wanna see a little bit more of a true reflection of what you see, I guess, every day when you leave your house in Australia.” Jaymie.

Above, Jaymie brainstorms on the fact that although the images were controlled to have some variations, in her view, the diversity portrait on the AI tool is dissimilar from her Australia life reality (Cave & Dihal, 2020; Lynn & Dixson, 2013).

“possibly in a shop of some sort selling something um could be a convenience store could be art could be a cafe it looks like it's like lively street just based upon the lights and the blur of people in the background. Um... he looks a bit tired maybe he's got like a maybe a hard worker.” Jazmine.

With at least one participant suggested he look like a “homeless” person. Thus, I argue, re-enforcing Australian racist stereotypes (Mapedzahama, 2019; Vass, 2015).

The feedback regarding the Aboriginal man potentially looking like a “homeless” person also underscores the harmful stereotypes that can arise in visual representations (Leonardo, 2004).

4.3 Perceptions of Australianness and Race in AI Image Generative Tools.

4.3.1 AI and Aboriginal Racial Perception in Australia.

“An Aboriginal man, maybe in his 50s or 60s, he's got slight greys coming through in his hair and his beard. He has a beard with like a moustache. He looks like he's on a city street, just chilling like on the side.” Helena.

Above, Helena attempts to rationalise what the generated photo depicts using colour-blindness as a form of escaping racial conversation (Mills, 2014). However, the image clearly diverges from the prompt, "Portrait of an Aboriginal Australian working in Melbourne, Australia on a busy day..." Rather than acknowledging that the image fails to align with the given parameters, her response shifts away from addressing this incongruity (Mills, 2014).

By not addressing the discrepancies in the image, Helena may inadvertently contribute to the continuation of harmful stereotypes and biases that fail to capture the lived realities of Aboriginal Australians (Mills, 2014; Tate, 1997). By not confronting the biases in the generated image, Helena's response exemplifies a form of colour evasiveness (Maeso & Araújo, 2015).

“Oh, where could he be working? I don't know, like a bar or like he looks like he's on the city street. Maybe like a retail shop. Maybe like food. Maybe he works in a kitchen at a local or like great restaurant like along the street there.” Helena.

The image of the Aboriginal man is, at least, incongruent to the prompt given (Nicoletti et al., 2023). What we found again here is that Helena goes to extraordinary lengths to justify why he is seemingly seating outside on the streets, when, in reality, the prompt says he is supposed to be working (Bonilla-Silva, 2006; Mills, 2014).

Individuals within the dominant group (in this case, white Australians) often exhibit a sense of ownership over the national narrative, which can lead to a white-washing of history that ignores the realities faced by Indigenous peoples (Sharples & Blair, 2021).

Furthermore, Balla (2020) highlights the challenges faced by Aboriginal individuals in being represented authentically within the dominant cultural narrative (Balla, 2020). This avoidance and devil's advocate type of approach in regard to racial issues, is one of the tools of white supremacy where individuals of the ingroup in position of power, deliberately avoid calling the situations as they see (Leonardo, 2004; Lynn & Dixson, 2013; Oliveira Andrade de Melo & Chapman, 2023).

"Um, he just seems to be sitting somewhere, like sitting somewhere on the street maybe, and he's got some like, it looks to be charcoal or dirt or something on his face. Um, he's not really wearing a uniform so you can't really place what his job is. You can't even really hazard a guess at what his job is. So, you put a very specific prompt in, and this is, I guess, the image of what it's telling you it thinks an Aboriginal person's job is." Jaymie.

Here, Jaymie, an Aboriginal woman, notices that his face is dirty, as she is, arguably, more attuned to what an Aboriginal well-presented and dressed should look like. This is partially due to the Cross-Race Effect which states that same-race faces, can be seen as more familiar which in consequence can create implicit biases toward outgroups (Young et al., 2012).

She notices he does not wear a work uniform, not a vest. Nothing indicates on the image itself that he is working. She goes further to call it out the AI tool's prejudice, in this case, Midjourney, for its perception of Aboriginal folk (Agrawal et al., 2024; Bass, 2023; Nicoletti et al., 2023). In this research by (Bonezzi & Ostinelli, 2021) titled "Can algorithms legitimise discrimination" this very issue is discussed where people may perceive bias by algorithms as less biased than bias from humans.

The perception of Aboriginal identity is often shaped by historical and systemic biases, which can be exacerbated by AI technologies that do not adequately account for these complexities (Lin & Chen, 2022).

So, unless he's an artist and has paint all over himself, I'm a bit like, where is it pulling the details from? Like, why does he have dirt on his face? And why is he sitting on the side of the road?” Jaymie.

Jaymie continues her elucubrations, saying that unless he is a painter why is he dirty then? Anti-Blackness manifests in societal perceptions as well as in AI tools (Janvieve Williams Comrie et al, 2022; Picower, 2021).

She questions the entire scene and works facing the racial prejudice that Helena, another participant, avoided above (Entman & Rojecki, 2001; Nelson et al., 2013; Vass, 2015). This perspective aligns with the different realities people from different racial backgrounds may experience whilst living side by side in western societies (Ani, 1994).

“No, I mean Aboriginal and Australian, yes. But working in Melbourne on a busy day, I mean, I feel like it would be quite subjective. I don't really feel that when I look at the picture.” Rachel.

Rachel, is, at least, sceptical regarding the image versus what the prompt proposes. This shows that some users, despite their racial background, can, with racial awareness, under the correct circumstances, perceive the prejudices and bias AI images can portrait (Stewart et al., 2012).

4.3.2 Race, Antiblackness and AI.

“So, you've got nine school kids playing on a playground. There's one young boy in the centre who's probably about nine or 10, smiling widely in his school uniform.” Ellie.

As a critical race academic, I am attuned not only to what is said but also to what remains unsaid (Oliveira Andrade de Melo & Chapman, 2023). In the above comments, Ellie does not address the racial elements of the children depicted in the school photo (Leonardo, 2004). While it may seem reasonable for her to focus on their identity as children, there is an implicit normalisation of whiteness in her omission (Leonardo, 2004; Lynn & Dixson, 2013). This normalisation becomes evident when contrasted with her later responses to images featuring diverse individuals, where she explicitly identifies racial elements (Bonilla-Silva, 2006; Leonardo & Porter, 2010). This shift underscores how whiteness often operates as an unmarked default, only disrupted when visible diversity is introduced (Bonilla-Silva, 2006).

"Yeah, I think that's like a responsibility of the creators of these AI generation systems that they're responsible to have more people involved in the process that can offer those differing opinions and have that representation. At the end of the day, like it's their system that has been created." Ellie.

Above, Ellie reflects on the need for diverse staff and diverse perspectives in AI development, highlighting the importance of accountability within the system (La Fors & Meissner, 2022; Silva, 2022).

While she does not explicitly mention race or gender, it can be argued that her critique aligns with the concept of Technochauvinism and implicitly addresses the male-centric and Eurocentric nature of the AI industry (Broussard, 2018; D'ignazio & Klein, 2020). As noted here on (Bell et al., 2018 p. 7) "Gender as a social construct can be (and has often been) defined by, enacted through, and embodied by AI systems that shape our interactions with the world" (Bell et al., 2018).

Ellie's comments suggest a broader call for inclusivity and systemic change in the ways AI is developed and governed (Andreotta et al., 2022; Holmes et al., 2021; Martin, 2022).

"A group of children on the playground in their school, they all have their school uniform on which is baby blue shirts and navy shorts. They're all happy and they're laughing, and they seem like they're having like a really really good time doing whatever they're doing.” Helena.

Above, we see Helena once again avoiding the racial elements of the individuals depicted. This aligns with the concept of normalisation of whiteness, where individuals fail to recognise their own racial identity while readily identifying the racial identities of others (Cave & Dihal, 2020). In isolation, this omission may appear harmless; however, it is notable that this is the same participant who, earlier in the study, explicitly remarked that certain non-white individuals "looked like immigrants” (Kilomba, 2021; Moreton-Robinson, 2004).

In here, we note the racialised assumption, where Helena nominates non-white individuals as "looking like immigrants" reflecting a broader societal tendency to categorise individuals based on racial and ethnic markers (Bargallie et al., 2024; Oliveira Andrade de Melo & Chapman, 2023).

When these comments are considered together, they reveal an implied framework of what she perceives as "normal" and "acceptable." This contrasts with what she, even though not mentioning race, implicitly marks as other or outside the norm (Moreton-Robinson, 2004). Racial identities and the power dynamics associated with them are often overlooked in favour of a narrative that prioritises individual experiences over systemic issues (Bonilla-Silva, 2006; Tate, 1997). This pattern underscores how racialised assumptions can inform interpretations, even when race is left unspoken (Moreton-Robinson, 2004; Tuck & Yang, 2021).

“Yep. Yep. So, a mother and a daughter are centred into the picture from an immigrant family, both with headscarfs on, out the front of their looks like suburban home” Helena.

Here, she not only agrees with the prompt without question but also fails to acknowledge that the individuals appear to be wearing what is very likely a hijab, a detail aligned with the prompt's indication that they are going to pray (Udah, 2023).

The failure to recognise the hijab in the image can be seen as a missed opportunity to engage with the intersection of race, culture, and religion (Balla, 2020; Tate, 1997). Furthermore, the omission of significant cultural elements, such as the hijab, can contribute to a distorted understanding of the individuals depicted, reinforcing stereotypes and limiting the audience's ability to empathise with their experiences (Entman & Rojecki, 2001).

This omission highlights a lack of engagement with the cultural and religious significance of the image, overlooking an opportunity to critically analyse how such elements are represented (Entman & Rojecki, 2001; Maeso & Araújo, 2015; Sharples & Blair, 2021).

“They look more of an immigrant family than an Australian family.” Helena.

Furthermore, there seems to be no compelling reason to argue that this is pinnacle of her racialisation of non-white individuals (Sharples & Blair, 2021) and how the representations of AI and intelligent machines are predominantly racialised as white (Cave & Dihal, 2020; Crawford, 2023), which, additionally reflects broader societal norms that position whiteness as the default (Cave & Dihal, 2020) .

Moreover, she explicitly states that they "look more like an immigrant family," reinforcing the burden of othering onto non-white individuals (Entman & Rojecki, 2001). Once again, she fails to critically engage with or question the biases embedded in the AI, instead perpetuating a narrative that positions non-whiteness as outside the norm (Bonilla-Silva, 2006; DiAngelo, 2022; Sriprakash et al., 2022).

The comment underscores the unexamined assumptions driving her interpretation of the AI-generated images which furthers the argument made by (Maeso & Araújo, 2015) where discussions about race often fail to recognise the underlying power structures that shape racial relations. It can be also argued that discussions on the construction of Australian identity often emphasise how immigrants are framed as "strangers" or "the other (O’Keeffe, 2024)."

This framing reflects a deeper societal tendency to position non-white individuals as threats to the established social order, perpetuating exclusionary narratives that maintain whiteness as the default standard of belonging (O’Keeffe, 2024).

“I think that this pretty much that did look like an Aboriginal guy that did look like an immigrant family. I think it pretty yeah.” Helena.

In the above comment, she expresses satisfaction with the outcomes of the AI generative system. Along similar lines, (Cave & Dihal, 2020) argues that the portrayal of AI as predominantly white reflects societal biases and contributes to representational harms. Although, different observers may derive various meanings from the same image based on their perspectives (Brooks, 1999) and also, for humans, there is no singular correct interpretation (Rosenbaum, 2022). Nevertheless, interpretations of images can be heavily influenced by cultural contexts (Farahani & Ghasemi, 2024). Helena's comments reveals how bias has influenced her perception, leading her to accept these images as reflective of reality and reinforcing them as normative (Farahani & Ghasemi, 2024). The failure to critically engage with these dynamics can lead to the perpetuation of biases (Bonilla-Silva, 2006; Tate, 1997).

Furthermore, this underscores a central argument of this thesis: the societal harms caused by imagery generative tools (Kreps & Kriner, 2023; Lorenz et al., 2023; Newton & Dhole, 2023; Poredi et al., 2024). The images generated are influence by those who created it (Dubber et al., 2020). "the designers of such systems… might make a series of subtle design decisions that might result in favourable or unfavourable treatment for different groups" (Dubber et al., 2020 p. 599).

For example, according to this survey by (Pocol et al., 2023) only six in every ten persons can identify the difference between generated images and real images online. This is so urgent and necessary that (Poredi et al., 2024) agues we must create methods in order to detect AI-generated images. Ultimately, generative AI, thus, could potentially, intensify the dangers of misinformation shared online (Lorenz et al., 2023).

(Bargallie et al., 2024; Metzl, 2019; Oliveira Andrade de Melo & Chapman, 2023) highlights how such neglect can perpetuate inequalities and hinder progress. The danger lies in the failure to challenge AI bias, which risks amplifying and entrenching these biases in ways that hinder racial inclusion in Western societies (KP, 2024).

As AI tools continue to expand their presence in everyday tasks (Lorenz et al., 2023), the uncritical acceptance of such biases could perpetuate exclusionary practices and narratives, making meaningful progress towards inclusivity increasingly difficult (DiAngelo, 2022; Fanon, 2004; Maeso & Araújo, 2015). This context is relevant to my study, as it underscores the risks associated with accepting biased AI outputs as reflective of reality (O'Keeffe, 2024; Schwartz et al., 2022; Silva, 2022).

“…it's probably not a fair depiction of what every Australian looks like, but a lot of the population is white with blue eyes, so that does make sense in that aspect.” Jaymie.

Above, Jaymie acknowledges that it is "fair enough" to note that many Australians do have blue eyes which is in line with the fact that the majority of the population have European ancestry and the concept of an Anglocentric Australia remains deeply ingrained as a foundational element of the nation’s cultural identity, shaping societal norms, values, and perceptions of belonging. (ABS, 2022; Moreton-Robinson, 2004; O’Keeffe, 2024).

However, in the same breath, she recognises that representations of diversity could have included other physical traits, suggesting an awareness of the limitations in the imagery (Agrawal et al., 2024; Bass, 2023; Lorenz et al., 2023; Srinivasan & Uchino, 2020).

Her comment reflects a tension between accepting a partial truth and recognising the missed opportunity for broader inclusivity in the visual representation of Australian racial identity (O’Keeffe, 2024). I can be further argued that understanding and addressing the complexities of racial identity is crucial for fostering inclusivity and recognising diverse experiences (Anderson et al., 2023).

“So, it is a photo of three women. They all have very fair skin, blushed cheeks, blue eyes. The first one has long red hair in plaits. The second one has long blonde hair in the middle part. And the third one has a purply red coloured hair with a braid.” Jaymie.

For Jaymie, there is a deliberate effort to explicitly mention that the women are white. Bargallie et al. (2024) develops the claim that racial literacy is important because it interrogates the dynamics of race, class, and other variables (Bargallie et al., 2024). In fact, acknowledging the racial dynamics at play is essential for understanding the complexities of representation and the ongoing struggles against white supremacy (Maeso & Araújo, 2015; Oliveira Andrade de Melo & Chapman, 2023).

By naming whiteness, rather than normalising it as an unmarked default, she acknowledges the existence of other racial groups (Leonardo, 2004; Mills, 2014; Whittaker et al., 2018). This act of naming is significant in the process of decentralising whiteness, as it challenges its assumed universality and creates space for more inclusive representations (Cole, 2021; Sriprakash et al., 2022).

“Yeah, yeah, like the standard, you know, Australian white boy with that haircut.” Jaymie.

Contrastingly with her previous comments, here Jaymie does mention the race of the Australian university student, explicitly noting that he represents "the standard." (Moreton-Robinson, 2004) discusses how whiteness operates as an unacknowledged norm that privileges certain identities while marginalising others (Sharples & Blair, 2021). This duality is revealing; it reflects her effort to challenge the normalisation of whiteness by naming it, while simultaneously highlighting how deeply entrenched whiteness remains as the default standard.

(Leonardo & Porter, 2010) highlights how race discussions often cater to a white racial frame, indicating that even in critical dialogues about race, the focus tends to revert back to whiteness as the norm. Also noting that as (Bargallie et al., 2024) argues, acknowledging whiteness is crucial for dismantling its normative status, which often goes unchallenged in educational and social contexts.

Her comment underscores the pervasive nature of this normalisation, even in moments of critical reflection. Whiteness thus, I argue, is insistent and pervasive (Janvieve Williams Comrie et al, 2022).

“I think it's again it's important to note that all of these feature people that AI has produced as an Australian have been white passing, have been having the appearance of being white and I think it's always quite interesting that you know the handsome white guy is like quite the archetype of an Australian.” Jaymie.

Again, in the above, Jaymie explicitly states her preference by identifying the Australian man student in the generate image as male; Nevertheless, she strongly critiques the AI image generative tool's clear tendency to represent white people as quintessentially Australian (O’Keeffe, 2024). It can be argued that Jaymie's critique of the AI image generative tool's tendency to represent white individuals as typically Australian underscores a significant issue regarding bias in representation (Cave & Dihal, 2020). Classifications in AI can reinforce existing social hierarchies and inequalities, particularly through the use of biased datasets that often reflect dominant cultural narratives (Crawford, 2021).

Also, her observation places attention on the bias embedded in the AI tool, arguably, reinforcing the narrow and exclusionary depiction of Australianness and highlighting the need for more inclusive representations. Furthermore, her critique resonates with the concept of "whiteness as property" discussed by (Lynn & Dixson, 2013) which posits that whiteness awards certain privileges and societal advantages, thereby marginalising other racial identities. In the Australian context, Moreton-Robinson (2004) discusses how the construction of Australian identity has historically privileged whiteness, particularly through the lens of Anglocentric culture. Still within the local context , (Sharples & Blair, 2021) highlights the anxieties surrounding the preservation of a white national identity in Australia.

“ it would be interesting to actually see the statistics as to how many, you know, Australian people are white versus not white, to see how it's kind of gathering that information or whether it's just creating what it thinks people want to see, what they want to see when [using] the word Australian...” Jaymie.

This may be a little-known fact: living in the inner cities of Australia can create the impression that non-Europeans make up a significant proportion of the population (Renzaho, 2023). However, the reality is quite different citizens without European ancestry account for less than 10% of the national population (ABS, 2022).

This contrast between public perception and demographic reality highlights the geographical and cultural concentration of diversity in urban areas, which can distort the broader understanding of Australia's racial composition (Guan & Prentice, 2024; Markus, 2023). Often this discrepancy between reality and perception is used for political gain and to increase disharmony (Markus, 2014; Mughan & Paxton, 2006).

“Microsoft trying to like to release a bot running its own Twitter account a few years ago or something and then it started spilling really racist stuff within 48 hours and they have to shut it down. Yeah, so how do you make sure that doesn't happen in AI? Like where is it? Is it gathering the data that people are feeding into it? Are people able to manipulate it? How do you stop manipulation of artificial intelligence?” Jaymie.

Here, Jaymie references the release of Tay, a chatbot originally launched by Microsoft as a Twitter bot in 2016 (Dubber et al., 2020). After its release, Tay quickly developed into a racist, neo-Nazi, and misogynistic persona, prompting its online ejection and a huge public outrage (Dubber et al., 2020). Tay exemplified one of the fundamental issues plaguing AI tools today: data bias used on large language models and AI tools training (Zou & Schiebinger, 2018b). Moreover, the available evidence seems to suggest that racial disparities on AI tools cannot be solely attributed to the under-representation of those minority groups within the training datasets (Gichoya et al., 2022). This means that solutions are not as simple and that AI systems, including Tay, can perpetuate biases regardless of the representation in the training data (Gichoya et al., 2022).

Developed within the framework of a predominantly anglophone, male, and Eurocentric internet, where much of the data informing AI systems is sourced, the chatbot quickly emulated the internet's biases, misinformation, and inaccuracies (Dubber et al., 2020; Farahani & Ghasemi, 2024; Mullaney, 2021; Rosenbaum, 2022).

Jaymie reflects on whether this phenomenon is being perpetuated by current AI tools, questioning how these entrenched prejudices might shape within these systems (Farahani & Ghasemi, 2024). Her observation underscores the critical need to address the structural flaws in AI development and the data it relies upon (Farahani & Ghasemi, 2024; Finn Lattimore, 2020).

“And a lot of those prompts are just focused on one image of what an Australian might look like. And they aren't really representing the full spectrum of who we are.” Jazmine.

Above, Jazmine critiques the myopic lens through which the AI defines who is Australian (Lopez Andersson, 2024). Australian nationalism is often rooted in colonial assumptions and white supremacy, which limits the understanding of who is considered Australian (Moreton-Robinson, 2004). Her challenge to this representation likely stems from her position as a member of the outgroup, where the exclusionary framing of Australianness reinforces her sense of being marginalised by the AI's narrow definition (Ferrer et al., 2021).

Multiculturalism has often been framed in ways that still privilege white, Anglo-Saxon identities (Moreton-Robinson, 2004; O’Keeffe, 2024). “bias and discrimination have a different ontological status: while the former may seem easy to define in terms of programmatic solutions, the latter involves a host of social and ethical issues that are challenging to resolve” (p. 77 Ferrer et al., 2021).

Her response underscores the need to improving AI literacy, questioning and broadening these technological biases in order to create more inclusive representations (Ferrer et al., 2021).

Increasing evidence indicates that algorithmic bias can lead to discriminatory outcomes, causing inequitable treatment of specific individuals or groups (Farahani & Ghasemi, 2024).

“…It might be different to what we actually, if we go out the door, what we see and whether that be ethnicity or, yeah.” Jazmine.

Once again, Jazmine challenges the AI-generated results, asserting that her lived reality does not align with the representation produced by the tool (Abdilla et al., 2021).

AI representations may not fully encapsulate the nuanced realities of individuals (Abdilla et al., 2021). This highlights the potential for AI systems to misrepresent lived realities, as they may not adequately account for the nuances of individual experiences (Danks & London, 2017).

Her critique highlights the disconnect between the AI's narrow portrayal and the diverse experiences of individuals, underscoring the limitations and biases inherent in the system (X. Wang et al., 2023). Furthermore, AI-generated results, can perpetuate biases that do not reflect the realities of non-white, non-European, individuals (X. Wang et al., 2023).

“Sometimes like even in high school like I might have been the only one of like your Asian background in the class but and then later in university I think you are exposed to a lot more people like even international students that come in like a lot more diversities and backgrounds, religions or cultures.” Jazmine.

Here, she shares her experience as an undergraduate student surrounded by many non-European peers, contrasting it with her childhood memories of being a minority in what she implicitly describes as a predominantly white school. Along similar lines, (Moreton-Robinson, 2004) argues that individuals' experiences, in predominantly white environments, significantly influence their perceptions of race and identity. Furthermore, Jazmine's comments express feelings of alienation and marginalisation when surrounded by non-European peers, which resonates with the experience of feeling like a minority in a predominantly white school (Guan & Prentice, 2024; Sharples & Blair, 2021). Her feelings are not isolated, in fact, (Bargallie et al., 2024) has encouraged debate on the importance of racial literacy in understanding the dynamics of race and identity within Australia.

This dual perspective offers a nuanced understanding of her reality, shaped by both her current environment of what she perceives as diversity and her earlier experiences of racial marginalisation in Australia (X. Wang et al., 2023).

"So, like saying Australian young women from different ethnicities and different appearances or weights and just like inserting that little edit into it so it knows that maybe to create some variation in what it's portraying."

Here, she challenges two often intertwined yet unspoken aspects of representation. She explicitly states her desire to see women from diverse racial backgrounds depicted in the image, while also addressing the issue of weight (Nambiar, 2023; Reddy, 2016).

The European female body, frequently used as the default standard, embodies specific characteristics that differ from those of women from other racial and ethnic groups (Janvieve Williams Comrie et al, 2022; Mitchell-Walthour & Morrison). This is in line with (Janvieve Williams Comrie et al, 2022) who discusses how societal standards often privilege Eurocentric beauty ideals, which can marginalise women of colour and those who do not conform to these standards. Furthermore, (Oliveira Andrade de Melo & Chapman, 2023) argues that non-European bodies are often marginalised in literature and media, reinforcing the idea that the European female body is the standard.

In her book, Browne argues that "the accuracy of gender classifier on Africans is not as high as on Mongoloid and Caucasoid," (, Browne, 2015 p.122).

This indicates that the algorithms used in facial recognition technologies are biased towards certain racial groups, often favouring the European standard (Browne, 2015). (Rosenbaum, 2022; Zou & Schiebinger, 2018b) further states that the European female body is often used as the default standard.

For example, the racial origins of BMI (Body Mass Index) standards underscore the harm in applying a universal measure rooted in European phenotypical norms.

According to Hall & Cole (2006) the BMI is the most widely used and popular (Bays et al., 2022) and arguably the most accurate instrument for supporting humans in tacking obesity today (Hall & Cole, 2006). However, it fails to mention the influence of tools such as BMI in the construction of structurally racist mechanisms of oppression of non-European populations (Dougherty et al., 2020).

While this study primarily focuses on racial elements, it is crucial to highlight and discuss how these issues intersect and inform one another, reflecting the broader, interconnected and intersectional nature of identity and representation (Bennett et al., 2021; Crenshaw, 2017).

Ferrer et al. (2021) discusses the importance of intersectionality in understanding biases in AI and health outcomes thus supporting the idea that racial elements in health metrics like BMI cannot be viewed in isolation but must be understood within the broader context of intersecting identities. In addition to that, Mullaney (2021) discusses how certain representations and standards are informed by historical and structural inequities, inclusive BMI (Mullaney, 2021).

“like say for young girls or women, it might affect how they look themselves. They might not be like, oh, okay, this is what an Australian woman should look like. Why am I not that image? Or how, what then, who am I? It might affect their sense of identity in the process.” Jazmine.

Once again, Jazmine focuses on the intersection of race and body positivity, highlighting how she cannot see herself represented within the beauty standards imposed by white women (Schlesinger et al., 2017). In fact, (Cave & Dihal, 2020; Entman & Rojecki, 2001) argues that the portrayal of AI as predominantly white reflects and sustains existing racial hierarchies, which can exacerbate social injustices and marginalise those who do not fit within these narrow representations.

She critiques the AI tool for perpetuating these narrow standards, pushing back against its failure to accommodate diverse bodies and diverse identities (Melonio, 2020; Nagar & Virk, 2017; Reddy, 2016; Riccio & Oliver, 2022; Sparrow, 2020; Yan & Bissell, 2014). This is proposed by (Bailey et al., 2021) where, their paper highlights how structural racism influences health and beauty standards, suggesting that these standards often reflect a Eurocentric bias that marginalises Non Europeans.

Furthermore, her response underscores the exclusionary nature of the AI’s representation and its impact on those who do not align with Eurocentric ideals of beauty (Bailey et al., 2021; Entman & Rojecki, 2001; Janvieve Williams Comrie et al, 2022).

“But thankfully, over time, you learn to just love yourself. But I think that resilience for younger girls especially would be impacted.” Jazmine.

In a hopeful message, Jazmine, an Indian Australian woman who is living in the west as part of the Indian diaspora, encounters many pressures to confirm to a body standard that is not hers (Hwang, 2021; Oliver-Hoyo & Allen, 2006). In order to survive, she has developed subsistence strategies and mental tools cultivated over the years, to survive and thrive in a predominantly white society (Bledsoe & Wright, 2019). Furthermore, supporting her testimonial, (Maeso & Araújo, 2015) highlights the significance of cultural identity and the strength derived from one's heritage in overcoming the challenges of living in a Eurocentric society. Also in line with this, the strength found in womanhood, along with a commitment to self-care, is crucial for navigating the complexities of race dialogues and societal expectations (Leonardo & Porter, 2010).

She also speaks of resilience, self-care, love, and the strength found in womanhood as key elements that enable her to flourish within a Westernised context (Bledsoe & Wright, 2019).

Her words highlight not only the challenges but also the profound agency and determination required to navigate and succeed in such environments. In relation to AI, it has been emphasised that intersectionality is important to be addressed in AI systems in order to mitigate its limitations (Dubber et al., 2020).

Ultimately, AI's failure to accommodate diverse bodies can lead to significant impacts on those who do not fit within the narrow standards (Mullaney, 2021).

“But if what we know about AI now continues and nothing changes, like say for example, you know, people say it's subjective, but we know it's not objective really. it is a product of a particular system, a particular way of thinking, which in this case is like Western methodology. And that in itself has biases and oppressive systems within that. So, if that continues and continues to replicate the current power dynamics, then we know that the outcome is going to be more oppression." Lorna.

Here, Lorna highlights how oppression and systemic inequities perpetuate themselves through institutions, education systems, and technology (Almeida, 2019; Bonilla-Silva, 2006; Tate, 1997). Her insights align with her lived experience as an African woman in Australia, where an acute awareness of her racial positioning and characteristics is likely a constant (Heleta, 2016).

This heightened understanding is not just a reflection of her identity but a necessary survival tool, enabling her to navigate the societal structures that consistently challenge her existence (Heleta, 2016). Her perspective underscores the resilience required to live within systems that often fail to accommodate or represent diverse identities equitably (Raji et al., 2020). Finally, although race is a social construct (Loury & Loury, 2009; Montagu, 1997), there is a role played by societal perceptions in shaping experiences of oppression (Banerjee et al., 2021).

"Okay, there are three young white women standing close together facing the camera, their eyes looking directly at the lens. On the left, the young woman's hair is orange, and she has it braided back into two braids coming down." Lorna.

As with other non-white participants, Lorna explicitly identifies the race of the women depicted by the AI image generative systems (DiAngelo, 2022; Jones, 2019). Colour-blindness is a privilege seldom afforded to non-whites (Bonilla-Silva, 2006; Bonilla-Silva & Dietrich, 2011).

This observation highlights her attentiveness to racial markers, likely informed by her lived experience and positionality, which contrasts with the tendency of some participants to overlook or normalise whiteness in similar contexts (Leonardo & Gamez-Djokic, 2019; Leonardo & Porter, 2010). Furthermore, discussions around race often cater to a white racial frame, which can lead to the marginalisation of non-white perspectives (Leonardo & Porter, 2010).

Structural racism influences perceptions and representations in various fields, including technology. This underscores the significance of Lorna's explicit identification of race as a critical response to the dominant narratives that often normalise Whiteness (Bailey et al., 2021; Sparrow, 2020).

“In my experience it's been usually more likely to turn out some of European descent for any prompt…” Lorna.

As a user of image generative systems (a detail I was unaware of when she registered to participate, though it in no way diminishes or hinders her contributions to the study) Lorna confirms that her experience aligns with some of the findings of this research (Cave & Dihal, 2020). She observes that, when creating images, the tool consistently defaults to depicting humans as white (Banerjee et al., 2021; Nadeem et al., 2022).

This reinforces the systemic bias embedded in such technologies, validating Lorna’s lived experience and its relevance to this study's focus (Finn Lattimore, 2020).

Furthermore, the amplification of societal bias that happens within AI systems, reinforces the notion that the generative systems Lorna uses are likely to perpetuate existing biases, validating her lived experience of the tool's default settings (Finn Lattimore, 2020).

Historical narratives and representations have marginalised non-White populations, which can extend to contemporary technologies (Maeso & Araújo, 2015).

“…you see even like, I don't know, newspaper, like all these like entertainment magazines, you hardly ever see a person of colour on the cover. So those kinds of things are just replicated in the AI.” Lorna.

This aligns with the discussion presented in the conclusion of this thesis. It is argued that individuals are significantly influenced by their surrounding environments, particularly through the images they encounter (Mullaney, 2021).

Visual representations have the capacity to shape and shift public perceptions on critical issues, reinforcing or challenging societal norms (Mullaney, 2021). The portrayal of African Americans in media often reinforces negative stereotypes (Entman & Rojecki, 2001; Tate, 1997). Moreover, anthropomorphised tools, such as AI, amplify this influence, holding even greater potential to impact perceptions profoundly and poignantly (Banerjee et al., 2021; Cave & Dihal, 2020).

Addressing these dynamics is both timely and essential, as the integration of AI into daily life continues to expand (Agrahari, 2024). This discussion highlights the imperative need to call out and address the biases built into image-generating technologies, paving the way for a fairer and more inclusive society (Ferrer et al., 2021; Kumar et al., 2024).

4.4 Chapter Conclusion

Several themes emerged during my study. Those themes were generated from the data collected via the semi structured interview questionnaire. The data was then analysed through NVivo (Creswell & Creswell, 2017). As my sample was, circumstantially, limited to female university students, it is likely that additional themes related to image generative systems may surface with different cohorts. Within the cohort investigated by this study, these were the themes that presented themselves. Though not all themes presented had sufficient depth to be included in this study. However, I believe they are worth mentioning for future researchers to explore in greater detail, helping to elicit further studies in this area.

Support graphic 5

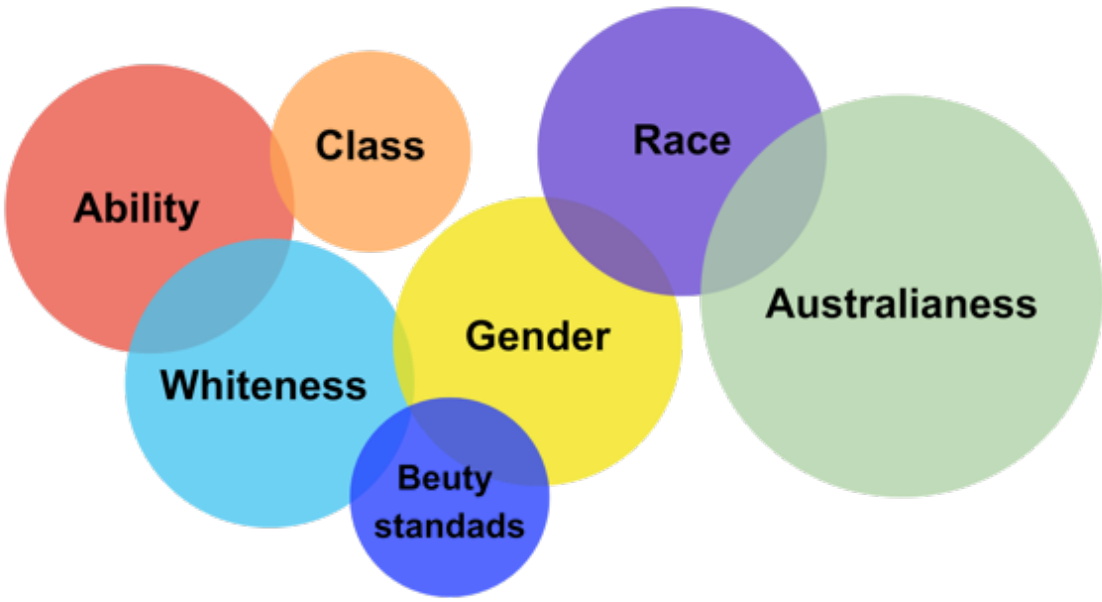

Figure 11

The representation of race in AI-generated imagery and its intersection with identity emerged as a critical point, along with questions of who is considered Australian, shaped by colonial narratives and perceptions. Understanding how image generative systems function and are perceived by the public, particularly regarding transparency and trust, remains an area suitable for exploration.

Gender biases in AI-generated imagery were also notable, as stereotypical roles appeared reinforced. Another key theme involved the real-world implications of image generation, including the credibility of AI, the ethical challenges posed by deep fakes, and their influence on societal trust. AI's portrayal in pop culture and its role in shaping expectations and fears further highlighted the cultural dimensions of the technology. Its use in politics, particularly concerning bias and propaganda, emerged as another important area for investigation. The potential for AI to provide solutions to societal issues, while raising ethical concerns, is equally significant.

Themes of bias in AI representation of non-Europeans and the implications for inclusivity also require attention. Finally, the portrayal of Aboriginality in AI-generated imagery revealed important considerations about representation and the potential perpetuation of racist colonial narratives. These themes point to critical avenues for future research to address the complex interplay of technology, identity, and ethics.

Although it was not among the themes that emerged in my thesis or the primary focus of the study, I would also like to highlight the issue of AI and ability, for example which, in this thesis is not addressed (Shew, 2020). Through personal exploration, I observed that AI generative systems appear to exhibit ableist tendencies and discriminatory patterns towards people with disabilities. This observation suggests that the technology often overlooks or misrepresents individuals with disabilities, reflecting biases embedded in its training datasets (El Morr et al., 2024; Shew, 2020; Whittaker et al., 2019). I believe this area calls for deeper exploration to ensure AI systems are inclusive and represent diverse abilities, promoting the fair and ethical use of AI technologies (Kumar et al., 2024).

Ultimately, this research was constrained by time and sample size, which, while suitable for a Master of Research program, produced results specific to the study's context and limitations. The study yielded results that were equally specific and idiosyncratic to those circumstances surrounding the research. A longer investigation, perhaps a PhD or industry based studies, on each specific topic mentioned are warranted.

Furthermore, additional research into the implications of AI image generative usage within our unique landscape, may be useful to elucidate its societal impacts in Australia.

Chapter 5: Discussion

Artificial Intelligence is no longer a distant concept confined to research or to science fiction (Murphy, 2024). It is a ubiquitous force, deeply embedded in nearly every industry, from healthcare to marketing, from education to criminal justice (Alrayes et al., 2024; Belciug & Iliescu, 2022; Birhane, 2021; Gichoya et al., 2022; Hamilton, 2023; Holmes et al., 2021; Metaxa et al., 2018; Murár & Kubovics, 2023; Sargsyan et al., 2024). AI is helping on pregnancy and IVF (Belciug & Iliescu, 2022) and can even help tackle climate change, and, it has the potential to help companies operate more sustainably (Vinuesa et al., 2020).

Support graphic 6

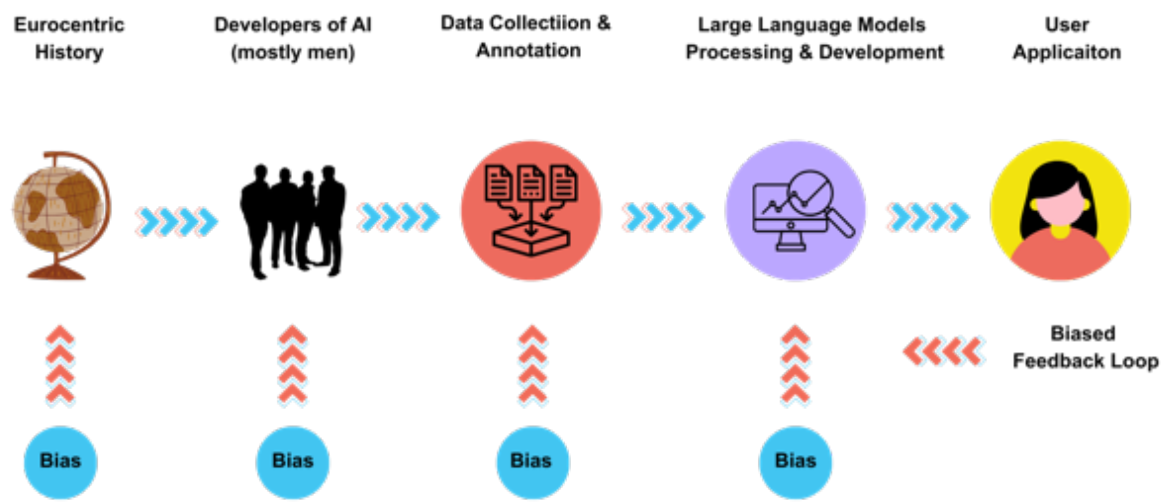

Figure 12

Its ability to process vast amounts of data and perform complex tasks is making AI indispensable in modern society (Kurian et al., 2023; Sargsyan et al., 2024). However, its increasing omnipresence also amplifies a critical issue: the inherent biases embedded in these systems (Belenguer, 2022; Danks & London, 2017; Ferrer et al., 2021; Gichoya et al., 2023; Mittermaier et al., 2023; Williams-Ceci et al., 2024; Zong et al., 2023; Zou & Schiebinger, 2018b).

In a recent article by MIT titled How it feels to be sexually objectified by an AI ,Heikkilä (2022) found that women online are often associated with sexualization, and that Asian women encounters this bias even more sharply (Bhaimiya, 2023). When asked why, a spokesperson for a large language model provider called this phenomenon a “sporadic sexualization” of women (Heikkilä, 2022).

Humans are profoundly influenced by the images, literature, and narratives they encounter (Ellison, 2021; Newman & Schwarz, 2024; Oliveira Andrade de Melo & Chapman, 2023; Phillips et al., 2018). This, in itself, is not problematic. However, when the tools shaping these mediums, like AI, are perceived as neutral and impartial, this is an issue (Crawford, 2021). AI tools, like all anthropoid developed tools, are designed by humans who themselves have their own biases and worldviews (Abdilla et al., 2021; Crawford, 2021). AI systems are part of schemes that are embedded in social, political, cultural, and economic contexts (Crawford, 2021). Those, therefore, are shaped by human decisions, which ultimately influence their functionality and impact.

The myth of AI's objectivity creates a dangerous blind spot, where biased algorithms can perpetuate and even exacerbate existing inequalities under the guise of fairness and efficiency (Danks & London, 2017; Dawson et al., 2019; Finn Lattimore, 2020).

Without these safeguards, however, the risks of misuse and gaps in accountability could undermine its positive impact (Birhane et al.). In the realm of tourism, for example, images play a pivotal role in shaping destination perceptions (Wang et al., 2022). In AI, a safeguard means any intentional technical, ethical, or regulatory measure designed to minimise harm, prevent bias, and ensure accountability in system design and deployment. This can include algorithm audits, dataset curation for diversity, transparent reporting, and governance structures that uphold fairness, privacy, and inclusion in AI outcomes (Benjamin, 2023; Cheryl Staats, 2017a; Noble, 2018; Tanksley, 2024).

These safeguards, oftentimes comes as audits. What is found is that current AI audits, while aimed at improving systems, often fall short of being effective. Achieving better outcomes will require further refinement and substantial effort to address existing limitations and challenges (Birhane et al.).

Recently, efforts have been made to explore the extent of bias in AI large language models by introducing a taxonomy of stereotype content. This approach aims to categorise and better understand how biases manifest in contemporary large language models (Nicolas & Caliskan, 2024).

Furthermore, automation bias, where individuals unquestioningly follow the directives of automated tools, has become an increasing and insistent concern (Ajunwa, 2020; Danks & London, 2017; Skitka et al., 1999; D. Wang et al., 2023). This reliance on automated systems overlooks the inherent idiosyncrasies and unique circumstances of individuals, promoting a one-size-fits-all approach (Dawson et al., 2019). Automation bias is dangerous for many reasons, one of it is the fact that it can legitimize and help justify discrimination (Bonezzi & Ostinelli, 2021).

Moreover, there is evidence in various domains, including education, healthcare, and decision-making processes, where tailored approaches are recommended and even paramount to address diverse needs more effectively (Dillenburger et al., 2014).

Such an approach, or worse, the adoption of an "average" solution to complex issues, poses a significant risk to human rights globally as it fails to account for the diversity and nuance of human experiences (Dawson et al., 2019; Holmes et al., 2021; KP, 2024).

5.1 Critical Race Theory as an invitation not as an accusation

Critical Race Theory (CRT) argues that race is a core part of social identity, and that racism is not just about individual prejudice (Bonilla-Silva, 2006; Sriprakash et al., 2022; Tate, 1997). Instead, it is a systemic and structural issue, embedded in the very fabric of laws, policies, and institutions, shaping the everyday realities of marginalised people (Lynn & Dixson, 2013). The theory highlights how race intersects with other social categories like class, gender, ability and sexuality, acknowledging that people face overlapping and interconnected forms of oppression (Buolamwini & Gebru, 2018;

Crenshaw, 2017; Shew, 2020; Tate, 1997). Critical Race Theory values the lived experiences and stories of non-Europeans, who are often excluded from academic spaces, legislation, and positions of power (Grosfoguel, 2015; Heleta, 2016; Joseph, 1987; Maeso & Araújo, 2015; Oliveira Andrade de Melo & Chapman, 2023; Pillay, 2020). In this area of study, these narratives are seen as legitimate and vital sources of knowledge, offering key insights into the realities of racism and oppression (Phillips et al., 2018; Tate, 1997).

For me, and for this study, Critical Race Theory is more than just a theoretical framework; it is a call to action (Guan & Prentice, 2024). It advocates for dismantling systemic racism and promoting social justice through avenues such as legal reform, educational practices, and, increasingly, the ethical development of artificial intelligence (Bonilla-Silva, 2006; Lynn & Dixson, 2013; Skinner-Dorkenoo et al., 2021; Sriprakash et al., 2022; Tate, 1997). Moreover, in a general sense, racial socialisation – which means engaging in conversations and providing instruction or learning opportunities about topics related to race and ethnicity (Loyd & Gaither, 2018) - should be seen as vital for the health and cohesion of diverse societies like Australia (Loyd & Gaither, 2018).

At its essence, Critical Race Theory, which, like other intricate frameworks, should function as an invitation. An invitation to think deeply, to empathise with others, and to care about the structures shaping our world. Yet, the digital landscape often distorts these ideas, weaponising them into tools for accusation. We live in today's world and moving forward is paramount. However, xenophobia is on the rise, and. surprising, many citizens of former colonies are now celebrating their former colonisers, this shows how things are not as simple as "goodies and baddies" (Baker & Cupery, 2023).

Critical Race Theory should be an invitation not an accusation…but how?

As perfectly put here by (Kowal, 2012 p. 7): "The study of dominant cultural groups cannot be based on their dominance alone, just as oppressed groups cannot be understood solely through the fact of their oppression" (Kowal, 2012).

The participants of this study, and even the individuals who built AI are part of a larger system (Indicators, 2024; Jarvie, 2019). I propose that participants of this study, independent of how they responded to the questionnaire, should not be seen as guilty nor as wrongdoers, but part of a larger ecosystem (Abascal et al., 2016; Nelson et al., 2013).

Moreover, The thoughts, emotions, and actions of individuals are shaped by the actual, imagined, or implied presence of others (Cheryl Staats, 2017a; Gatwiri & Anderson, 2022; Harrison et al., 2006).

If anything they are victims of a larger historical ideology (Mills, 2014; Nelson et al., 2013). This should not be a contest between good and bad whites or good or bad men (DiAngelo, 2022). Frameworks, policy and law can help change society, not individual simplistic accusations and finger pointing (Kendi, 2016).

As noted on the quote below, this is not a unique Australian nor a unique white problem. It is a global issue (Stewart et al., 2012).

“Prejudice and discrimination continue to be social problems on a global scale. Understanding the mechanisms through which intergroup biases are perpetuated and the means by which they might be reduced therefore remains a central focus of both social psychologists and diversity training practitioners” (Stewart et al., 2012 p. 12).

Racially speaking, perception is different depending on one’s version of reality, for example, the adjective “ethnic”, according to (Agnes, 2003) simply means, in most other languages other than English, to be related to or “relating to a large groups of people classed according to common racial, national, tribal, religious, linguistic, or cultural origin or background”. However, in the USA, white Americans, developed this seemingly innocuous word into a noun that denotes NOT white humans. It means, to be a member of other groups there are not Anglo-Saxon nor are members of the correct type of white (Sverdljuk et al.). These bias moves into the AI (Eidelman & Crandall, 2012; Mitchell, 2021). The bias, thus, is even on the definition of words.

On the internet and social media, one often encounters accusatory statements similar to “…you are complicit [insert cause here] if you don’t know this”. Phrases like that are all too common online, creating a culture of guilt that pressures individuals to engage superficially rather than encouraging meaningful understanding (Grosfoguel, 2015). Moreover, it can be argued, that framing anti-racism solely in terms of guilt can undermine the political nature of racial struggles (Maeso & Araújo, 2015). Perhaps centring the lived experiences of Black (non-white) individuals, in this particular realm, rather than centring discussions on white guilt, can help (Janvieve Williams Comrie et al, 2022). Moreover, white guilt can paralyse critical reflection and distract from understanding the structural origins of racism (DiAngelo, 2022; Leonardo, 2004).

My research takes a different path. Specially knowing that for someone to perceive their so called societal privilege one must be taught (Stewart et al., 2012). As

Paulo Freire would say, as one teaches one learns, and, as one learn they too teach; it can only be done together as a dance (Freire, 2000).

I wish to go beyond, instead of fostering guilt, I seek to raise awareness of the biases embedded in AI systems and to inspire collective efforts toward building a more equitable society. Humanity has much to gain from confronting issues of race, class, colonialism, feminism, ableism, and other forms of systemic discrimination. By working together to understand the historical and cultural forces that have shaped our present, we can strengthen our communities and improve life for every human. This collaborative effort, grounded in knowledge and empathy, has the potential to guide us toward a future that is both inclusive and just.

Critical Race theory should be an invitation to all humanity, and people of all races, to fully comprehend the harm done for the last 600 years of racial oppression. Although racism was not created in Europe (Dove, 1998; Harari, 2014), it was expended , developed, theorised, academically justified, and, eventually, transformed into law, policy and modus operandus of European thinking and epistemology building (Ani, 1994; Grosfoguel, 2015; Kendi, 2016; Nascimento, 2020). We are all, whites, Blacks and other racial groups, paying for the consequence of racial and territorial colonisation (Sriprakash et al., 2022).

There is a potential feeling, that many white or non-African Black readers may feel when reading about racial justice and CRT , which may be “what is it for me?” or “what am I losing by not having non-Europeans in my life? (DiAngelo, 2022; McGhee, 2022).

One of the arguments for imperialism is that all other (non-European) civilizations would, at a point of their development, become imperialist themselves (Ani, 1994). This has been effusively debated and debunked by decolonial and Afrocentric scholars (Ani, 1994; Bernardino-Costa et al., 2018; Grosfoguel, 2015).

Furthermore, it is proposed that Western knowledge systems are structured to benefit a select few, excluding those unable to participate (Mignolo, 2011). This systemic exclusion marginalises diverse perspectives, reinforcing existing power imbalances and perpetuating inequities in knowledge production and dissemination (Mignolo, 2011). Additionally, since at least the Renaissance, the west has been establishing the epistemic direction of the globe (Mignolo, 2011). Perhaps, it is time to open up to different players to offer alternatives ways of create epistemology (Mignolo, 2011).

I posit myself with Felwine Sarr, (2020) who proposes in his book Afrotopia, that all of humanity won't know what is missing and we won't know the other possibilities of existing because, hitherto, it only works and have worked on the European paradigms and dogmas (Sarr, 2020) . But there should be more to us than that (Sarr, 2020). And surely, by now we could agree that it is worth a try.

5.2 Are people influenced by images?

The influence of images on human psychology is a well-documented phenomenon (Newman & Schwarz, 2024). Images can evoke strong psychological reactions, shape perceptions, and influence behaviours across different contexts, including health, tourism, and environmental awareness (Deng & Li, 2013). Images are so powerful, that they make people change their memories and perceptions depending on their previous bias and pre-existing stereotypes (Duan et al., 2021; Jo et al., 2022; O'Connell & Greene, 2017).

Moreover, positive imagery, such as nature scenes, have been shown to enhance mood and physiological responses, promoting relaxation and well-being (Jo et al., 2022).

Images play a significant role in influencing elections, particularly through their capacity to shape public perceptions of candidates and their messages (Bovet & Makse, 2019; Mattes et al., 2010; Rodarte & Lukito, 2024).

In fact, images on social media significantly impact how voters perceive candidates, with attributes like honesty, intelligence, and trustworthiness being crucial for electoral decisions as well as leading less informed voters to base their choices on visual impressions rather than substantive policy issues (Lenz & Lawson, 2011; Mattes et al., 2010). Furthermore, the personalisation of political communication through visual means, such as campaign posters and social media posts, has been linked to increased voter engagement and turnout (Steffan & Venema, 2020).

It can be argued that AI is the continuation of the colonial project, thus, its implementation without safeguards, ethics or cultural balance, reinforces the dominance of Western values, epistemology and knowledge systems in AI development, sidelining non-Western perspectives and narrowing the potential for decolonising AI and diversifying its frameworks (Ebell et al., 2021; Muldoon & Wu, 2023).

5.3 Race

Racism is often subtle, complex, and hidden, making it difficult to recognise (Mapedzahama, 2019). AI technologies can make this even more covert due to the perception of AI as impartial and objective (Dubber et al., 2020; Ferrer et al., 2021).

Negative media portrayals of African Americans have been shown to contribute to racism, which is a pervasive form of racism that influences individual and institutional expressions of racism (Szymanski & Lewis, 2015). This cultural racism is often manifested through the internalisation of negative stereotypes, leading individuals to adopt prejudiced views against their own racial group or others (David et al., 2019).

It can be argued that the concept of internalised racism, which is a type of racism that occurs when people from a racially marginalised group unconsciously take on society’s negative beliefs, stereotypes, or harmful narratives about their own group. This can lead them to undervalue themselves or others like them, and to act in ways that keep racial inequality in place (Gatwiri & Anderson, 2022; Natapu-Ponton, 2024). This is crucial for understanding how images shape and influence racial attitudes (Cheryl Staats, 2017b). Internalised racism involves the internal acceptance of negative societal stereotypes and ideologies, often perpetuated through visual media (Campón & Carter, 2015). This acceptance can result in self-doubt, a diminished sense of self-worth, and identity struggles among racially minoritised individuals (Campón & Carter, 2015). By reinforcing these stereotypes, images, inclusive AI generated ones, hold the power to normalise harmful ideologies, making the need to critically engage with visual representations even more pressing in efforts to combat systemic racism. (Campón & Carter, 2015). The concept of unconscious or implicit bias is linked to racial profiling, where individuals form rapid, automatic associations between perceived race and characteristics such as dangerousness or criminality (Hopkins, 2024).

Support graphic 7

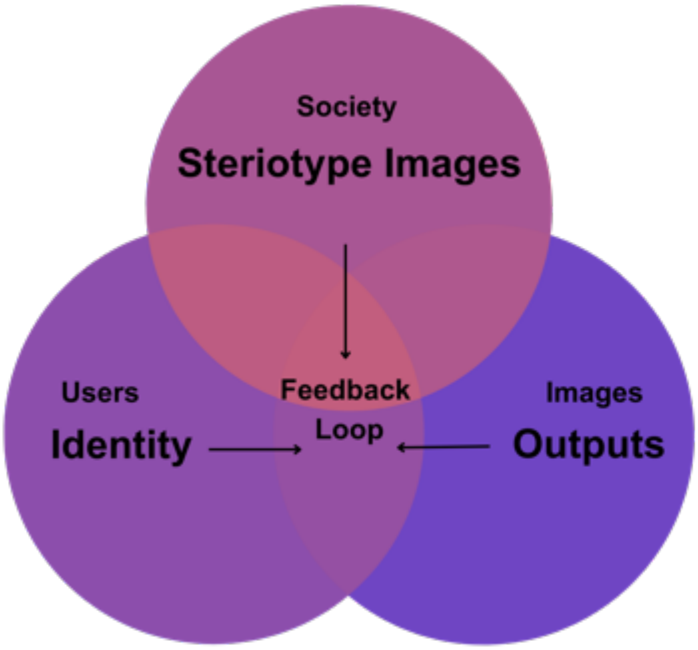

Figure 13

Moreover, the theory of symbolic racism posits that images and representations in media can shape political attitudes and policy preferences among white individuals, often leading to opposition against race-conscious programs (Kite et al., 2022; Sears & Henry, 2002; Sniderman & Tetlock, 1986).

Visual media can serve as a vehicle for perpetuating racial biases, as individuals may unconsciously align their beliefs with the negative portrayals they encounter (Kite et al., 2022). The influence of these images is not limited to overt racism; it also includes subtle forms of bias that manifest in everyday interactions and societal structures (Ramasubramanian, 2015; Tukachinsky et al., 2015). In a different direction, (Ramasubramanian, 2015) puts forward the view that exposure to news stories featuring African American media personalities in counter-stereotypical roles, as opposed to stereotypical ones, led to a reduction in stereotypical perceptions and symbolic racist attitudes held by white Americans toward African Americans.

Moreover, the representation of racial minorities in media can influence public discourse and policy decisions, thereby shaping the lived experiences of these groups (Tukachinsky et al., 2015). Ultimately, what is being seeing in imagery can either be a counter narrative against racism and bias (Ramasubramanian, 2015), or it can lead to increased bias, bolstering of negative stereotypes, and the perception that racial minorities fit limited, negative roles (Asadi & Milani, 2022; Kite et al., 2022).

5.3.1 How images can influence Popular Culture

Images have a significant influence on pop culture (Lee & Bai, 2016). Moreover, Pop culture imagery, such as in media, entertainment, and advertising, can shape public perceptions, attitudes, and behaviour (Lee & Bai, 2016). Furthermore, photographs can construct meaning, exert influence over public perception, and actively shape broader cultural, social, and political landscapes ("Portrait Photography as a Visual Repository of Vernacular Memory: Interplay of Memory, Identity and Culture," 2023). AI tools can analyse images and figure out what will be viral or popular and deploy those images to the public (Mullaney, 2021). It may mean that because of its embedded bias, those images will favour the ingroup. (whites or European descendants in the Western case) over other members of society.

Additionally, in popular media, the portrayal of racial minorities often falls into stereotypical roles, which can shape societal attitudes and individual self-perceptions (Schmader et al., 2015). Moreover, (Schmader et al., 2015) explored how stereotypic film portrayals affect the emotional responses of ethnic groups, revealing that negative portrayals can prompt adverse feelings among those represented.

5.3.2 Advertising

Recently, AI-driven digital images and products have been introduced into the market utilising psychological, social and cultural readings of users in order to obtain sales results (Dubber et al., 2020).

"AI-based advertising systems manipulate us far more efficiently and in a much more personal way than traditional advertising" (Mullaney, 2021 p. 40). Therefore, there seems to be a compelling reason to argue that personalised image design, facilitated by AI, can significantly improve the effectiveness of advertising strategies (Mullaney, 2021).

Moreover, it has been argued that the emotional impact of imagery plays a crucial role in advertising effectiveness (Dunlop et al., 2008). Therefore, it is not a surprise that similar effect may be experienced when users are observing AI generated images.

Advertisements prompting high emotional responses, particularly through graphic imagery, tend to be more memorable and impactful, reinforcing the idea that emotional intensity can enhance memory and recognition (Dunlop et al., 2008).

Moreover, Zainordin et al (2021) discusses how age, race and ethnocultural factors can affect the reception of advertising images, suggesting that understanding the target audience's characteristics is essential for crafting effective advertisements (Zainordin et al., 2021). This, consequently, may be also applied to AI tools such as image generative tools.

I propose that further investigation and academic research in the influence and effect of generated AI images on individuals controlling for age, gender, ability, race, ethnical background, nationality, and even social economical class is urgently needed.

AI has an extraordinary ability to generate images so realistic they shape how people perceive reality (Atanasoski & Vora, 2019; Mullaney, 2021). These images carry an air of neutrality, making their influence particularly pervasive (Belenguer, 2022; Bolukbasi et al., 2016; Eidelman & Crandall, 2012; Favaretto et al., 2019; Ferrer et al., 2021). Furthermore, Images generated by AI are not just reflections of reality but actively shape societal values and perceptions (Atanasoski & Vora, 2019). The available evidence seems to suggest that the datasets powering AI tools are rarely impartial if at all (Bender & Friedman, 2018).

In a world where considerable amounts of data privacy is privately owned (Walsh, 2023b), a standout exception is the Mukurtu project, developed by Kim Christen and Craig Dietrich, which offers a groundbreaking approach to managing privacy in digital cultural heritage (Pedersen, 2024). The platform enables Māori and other communities to safeguard sensitive content while allowing for controlled sharing that honours their cultural traditions and practices (Cofey, 2021; Hall, 2017; Pedersen, 2024). Furthermore, Mukurtu stands as an essential example of a Māori-led initiative harnessing Artificial Intelligence to preserve the Māori language and protect Indigenous data sovereignty (Cofey, 2021; Hall, 2017; Pedersen, 2024).

AI datasets and the large language models that control it reflect societal prejudices tied to race, gender, language which is mostly Anglophone centric (Nicolas & Caliskan, 2024; Zaugg et al., 2022), class, ability, and cultural norms (Kite et al., 2022). These biases, thus, are not just replicated; they are amplified (Bonezzi & Ostinelli, 2021; Celi et al., 2022; Finn Lattimore, 2020; Lewis, 2019; Whittaker et al., 2019; Zou & Schiebinger, 2018b).

For instance, an AI system trained on Eurocentric beauty standards produces images that repeatedly reinforce those ideals, subtly positioning them as universal standards for beauty (Melonio, 2020; Reddy, 2016; Ryan-Mosley, 2021; Yan & Bissell, 2014). This normalisation marginalises other identities and contributes to the exclusion of diverse perspectives (Buolamwini & Gebru, 2018; Sparrow, 2020). This creates a dangerous feedback loop: societal biases shape AI, which in turn reshapes societal norms, embedding stereotypes deeper into our culture (Bender et al., 2021).

The trust placed in these tools heightens the issue, as individuals frequently perceive AI outputs as neutral and objective, these users often overlook the inherent biases embedded in the data and algorithms (Bonezzi & Ostinelli, 2021; Finn Lattimore, 2020). These dynamics extend beyond representation; they influence decisions in industries like hiring, education, and healthcare, where AI tools often reinforce systemic inequities (Agrawal et al., 2024; Alrayes et al., 2024; Chan & Hu, 2023; Eubanks, 2018; Holmes et al., 2021; Poba-Nzaou et al., 2021; Sargsyan et al., 2024).

The problem is one of power and who gets to be seen and who doesn't. Whose stories are valued, and whose realities are erased (Entman & Rojecki, 2001; Lynn & Dixson, 2013; Maeso & Araújo, 2015).

Another fear is that as organisations increasingly use AI to maximise profits, this can, and likely will, come at the expense of consumer fairness and transparency (Md Sumon Gazi et al., 2024).

In Australia, where discussions around multiculturalism and representation are ongoing, these issues are particularly relevant (Austin & Fozdar, 2018; Majavu, 2020; O'Keeffe, 2024). As AI becomes embedded in daily life, its ability to skew perceptions of identity, beauty, and normality, risks perpetuating existing inequalities (Sargsyan et al., 2024).

Addressing this requires more than diverse datasets or algorithmic transparency (Andreotta et al., 2022; Matt Davies 2023). It demands critical engagement, questioning who controls these narratives and whose interests they serve. We must adopt digital literacy that challenges the illusion of AI neutrality (Bonezzi & Ostinelli, 2021; Finn Lattimore, 2020). If left unchecked, these systems will continue to reinforce exclusionary practices and limit progress towards a more equitable society. AI should reflect and celebrate diversity and not further entrench the biases of the past (Ferrer et al., 2021; KP, 2024).

5.3.3 Aboriginal Perceptions

Media plays a crucial role in shaping public perceptions by either reinforcing negative stereotypes or providing a platform for Indigenous voices (Moreton-Robinson, 2004). stereotypes and racist depictions in the media can create a moral panic around issues such as violence and dysfunction Moreton-Robinson (2004).

For instance, mainstream media often portrays Aboriginal communities in a manner that emphasises dysfunction and poverty, which can lead to a public perception that views these communities as “dangerous” (Dyson et al., 2016; O’Keeffe, 2024). This framing is compounded by the historical context of settler colonialism, which has systematically undermined Indigenous autonomy and identity (Bennett et al., 2023; Wolfe, 2006).

In contrast, Indigenous media initiatives aim to present a more nuanced and authentic representation of Aboriginal cultures, thus, challenging the dominant narratives and fostering a more positive public perception (Latimore et al., 2017; Lewis et al., 2020). In fact, such initiatives coming from First Nations themselves, are vital as they provide an inside perspective in the Indigenous public sphere (Latimore et al., 2017). Those narratives counter the reductive portrayals found in mainstream media where, often, First Nations are portrait as “problem population” (Latimore et al., 2017).

5.4 But in AI imagery we are under illusion that the images are real?

AI possesses a remarkable capacity to generate highly realistic images that profoundly shape users' perceptions of reality (Newman & Schwarz, 2024; Rosenbaum, 2022). This power lies in its ability to process large datasets and produce visuals that mimic those created by human hands (Newton & Dhole, 2023; Zou & Schiebinger, 2018b). Furthermore, as my research highlights, these technologies are far from neutral (Bonezzi & Ostinelli, 2021; Crawford, 2023; Noble, 2018).

When AI systems are trained on datasets rooted in historical inequities or shaped by gaps in data collection, these biases are not only reflected but often amplified in the outputs (Bender & Friedman, 2018; KP, 2024). For instance, AI tools generating workplace imagery frequently default to depicting leadership roles with predominantly male, Eurocentric features, while relegating women and non-whites to secondary or stereotypical positions (Cave & Dihal, 2020; Li, 2020). Such portrayals reinforce harmful narratives about who belongs in positions of authority and who does not (Entman & Rojecki, 2001; Maeso & Araújo, 2015).

This is especially troubling because many users inherently trust AI outputs, perceiving them as objective representations of reality (Bonezzi & Ostinelli, 2021; Newman & Schwarz, 2024). Over time, repeated exposure to these biased images reshapes societal ideas of what is "normal" or "acceptable," affecting self-esteem, career aspirations, and broader expectations (KP, 2024). This cycle of bias highlights the urgent need for change (Buolamwini & Gebru, 2018; Raji & Buolamwini, 2019; Raji et al., 2020; Whittaker et al., 2019). To address this, we must prioritise diversifying datasets, ensuring ethical oversight, and encouraging critical engagement with AI tools to create fairer and more inclusive representations in every aspect of society (Chan, 2023; Dawson et al., 2019; Holmes et al., 2021; Mhlambi, 2020; Oluwaseun Augustine Lottu et al., 2024; Schmidt et al., 2024).

Without delving into the manipulation of images, which is itself a field worthy of further research, it is important to highlight another significant tool : Adobe Photoshop. Photoshop, a widely used image editing software, has become an integral tool in various fields (Corl et al., 2002; Dubey et al., 2016). Because the Adobe software is known for manipulating and enhancing digital images, when combined with AI image generative, tools like Photoshop make it increasingly difficult for the average viewer to discern between what is real and what has been altered (Corl et al., 2002; Zou & Schiebinger, 2018b). For example, "AI-driven filtering practices can shape user mood, which can be used to enhance the impact of targeted advertising" (Mullaney, 2021 p. 46).

This convergence of technology not only blurs the lines between authenticity and fabrication but also poses serious implications for trust in visual media (Dubber et al., 2020; KP, 2024).

5.5 Are people influence because of their ignorance, or it is malice ?

One of the defining challenges of the early 21st century is the polarisation exacerbated by algorithms and social media platforms (Ali et al., 2021; Guess et al., 2021; Kubin & Sikorski, 2021). This division is not limited to the realms of politics or entertainment but has seeped into every facet of human endeavour, including academia (Beard, 2022; Lahiri-Dutt, 2018). Many times, this polarisation is driven by the tendency of users to seek out and engage with information that aligns with their existing political view (Cho et al., 2016; Guess et al., 2021). Situations like "techlash" a phenomenon where critical coverage of technology's overreach into everyday lives have significantly increased in recent years (Dubber et al., 2020).

The selective exposure and interaction with like-minded content reinforces users' existing beliefs and attitudes, leading to increased polarisation (Kubin & Sikorski, 2021; Wollebæk et al., 2019).

As complex ideas like feminism, critical race theory, decolonisation, and settler colonialism move beyond scholarly circles into popular culture, controlling the narratives and mainstream discourses surrounding these concepts has become increasingly difficult (Literat et al., 2022). Literat et al. (2022) examined how social media platforms like TikTok can be used for "media criticism," where marginalised communities can shape their self-representation and challenge dominant narratives. However, the same platforms can also be used to spread misinformation and destructive forms of media criticism (Literat et al., 2022).

These theories and frameworks are inherently multifaceted and require deep, empathetic engagement to be fully understood (Mitchell-Walthour & Morrison; Tate, 1997). It can take years, or even a lifetime, to grasp the layers of meaning embedded within them (Bell, 1995).

Although, it is possible that social media can enable users to have political discussions with peers, which, consequently, can increase political interest and drive them to engage in protest grass roots activities (Valenzuela et al., 2016).

It is also likely that the algorithms powering social media platforms, thrive on conflict and division, prioritising user engagement over substance (Ali et al., 2021). Consequently, these intricate and nuanced discussions are often distilled into overly simplistic narratives designed to fit ever-shortening attention spans (Carstens et al., 2018). For example, the complex and layered discourse of decolonisation might be reduced to a 15-second video, stripping away its essence and depth in the process (Basch et al., 2022; Vizcaíno-Verdú).

Are individuals swayed by this oversimplification because they are ignorant? The answer is both yes and no. Yes, because many feel an intense pressure to stay informed in today's fast-paced, hyper-connected world, often consuming information without the time to reflect on its depth (Dubber et al., 2020; Eubanks, 2018; Favaretto et al., 2019). No, because the overwhelming visual stimuli and relentless flow of content in digital spaces make it nearly impossible to disengage, drawing people into discussions that are frequently reduced to reductive and polarising frameworks (Infield, 2020). The result is often a cycle where the search for knowledge is complicated or even evaporated by the very systems designed to disseminate it (Infield, 2020).

But not all is despair, It has been found that images can influence individuals positively or negatively. Also, "the findings suggest that media images of duration as little as a minute can lead to increased acceptance [internalisation] of socially defined ideals" (Nagar & Virk, 2017 p. 5). Furthermore, it is important to point that the relationship between social media usage and cognitive performance is a complex and nuanced topic.

Barton et al., (2018) suggested that excessive social media use, particularly on platforms like TikTok, may have negative impacts on attention, motivation, and academic performance (Barton et al., 2018). While excessive use may have negative impacts, the specific mechanisms and long-term effects are not yet fully understood and further research, particularly in Australia, is needed (Basch et al., 2022; Kang & Lou, 2022).

5.6 How does the results fit into concepts such as CRT

The findings align with Critical Race Theory, revealing that AI image generation tools often perpetuate antiblackness and Eurocentrism (Dancy & Saucier, 2021; Guillory, 2020; Maeso & Araújo, 2015). Australian users tended to overlook race and gender when presented with images of white, able-bodied individuals but became acutely aware of racialisation when encountering depictions of non-white figures (Sharples & Blair, 2021; Wolfe, 2006). This study sheds light on the perceptions of Australian users, who operate within a unique racial context (Moreton-Robinson, 2004, 2021).

The majority of Australia's population is of European descent, with significant Asian communities due to geographical proximity, while African and Aboriginal populations remain comparatively smaller (ABS, 2022). While there is some empirical research on racist attitudes in Australia, comprehensive data remains limited (Ben et al., 2024; Gatwiri & Anderson, 2022; Gatwiri et al., 2021; Yared et al., 2020). Studies such as the Challenging Racism Project have examined attitudes toward cultural diversity and experiences of racism, but a nationwide, in-depth analysis is still lacking (Ben et al., 2024; Dunn et al., 2004; Gatwiri et al., 2021).

Addressing biases in AI technologies is essential to ensure fair and equitable representation for all racial and ethnic groups, as these dynamics highlight the potential for systemic disparities (Pulivarthy & Whig, 2025). By critically examining and rectifying these biases, we can work towards AI systems that reflect the true diversity of society, fostering inclusivity and reducing the perpetuation of harmful stereotypes (KP, 2024).

Moreover, social cohesion in Australia will increasingly depend on the equitable development and deployment of artificial intelligence (Roselli et al., 2019; Sargsyan et al., 2024). As more Australians integrate these tools into their decision-making processes, it is imperative to ensure that AI systems are designed and implemented in ways that promote fairness and inclusivity (Alrayes et al., 2024).

This necessitates comprehensive research into how AI intersects with race, class, gender, ability, and ethnicity. Such studies are essential to identify and mitigate potential biases, ensuring that AI technologies serve to unite rather than divide our diverse society (Feijóo et al., 2020; Selwyn, 2004).

5.7 Is this Intentional (Legislation, Design and Law)

It's natural, or perhaps inevitable, that individuals design tools reflecting their own perspectives and needs. In academia, oftentimes, academics use reflexivity and critical analysis to establish themselves as "subjects" of their own destiny, which implies that the tools and frameworks they create are influenced by their specific cultural and historical contexts (Heleta, 2016). It can be also argued that scholars [and individuals] should be aware of their own ontological and epistemological biases (Oliveira Andrade de Melo & Chapman, 2023).

The movement to decolonise the curriculum underscores the necessity for a fundamental transformation in the construction of knowledge. It advocates for educational tools and frameworks that embody the diverse perspectives of various cultures and histories, challenging the dominance of Western-centric narratives (Heleta, 2016).

Historically, women's health research has lagged, largely because men dominated scientific inquiry. For example, O'Keeffe (2024), discusses how the medical model has historically prioritised male health issues, leading to a significant underrepresentation of women's health concerns in research agendas (O'Keeffe, 2024). The persistent underrepresentation of women in scientific research has significant implications, particularly in perpetuating gendered perspectives within the field (Sharples & Blair, 2021). Similarly, race plays a significant role here in Australia, and Aboriginal women and Black and Non white women suffer the most discrimination (Balla, 2020; Sharples & Blair, 2021).

Ageism is another issue in AI Age-related bias in artificial intelligence (AI) systems has recently emerged as a critical issue demanding immediate attention. The historical neglect of certain demographic factors, particularly age, in AI development has led to systems that inadvertently perpetuate ageism. This oversight can, therefore, also result in discriminatory outcomes (Stypińska, 2022).

Comparably, it can be argued that biases in AI systems often arise not from deliberate intent but from neglect and profit-driven motives. "Bias and discrimination have a different ontological status: while the former may seem easy to define in terms of programmatic solutions, the latter involves a host of social and ethical issues that are challenging to resolve from a positivist framework” (Ferrer et al., 2021 p. 7).

Because of its data bias, AI has been known to influence different racial groups individuals to react more biased depending on the AI creations (X. Wang et al., 2023).

"the ethical development of AI cannot solely be viewed as a 'technical' problem to be resolved, and instead requires a strong focus on the communities it affects" (Dawson et al., 2019 p. 25).

5.8 Recommendations and Suggestions

Addressing discrimination in AI demands comprehensive interdisciplinary collaboration. This approach integrates diverse perspectives from fields such as computer science, ethics, law, health, academia, literature, education, and social sciences to effectively identify and mitigate biases within AI systems. By fostering cross-disciplinary partnerships, we can develop more equitable AI solutions that reflect a broader range of human experiences and values. It is known that users of AI are often willing to learn and use ingenuity and creativity in order to overcome learning difficulties (Okada, 2024).

Support graphic 8

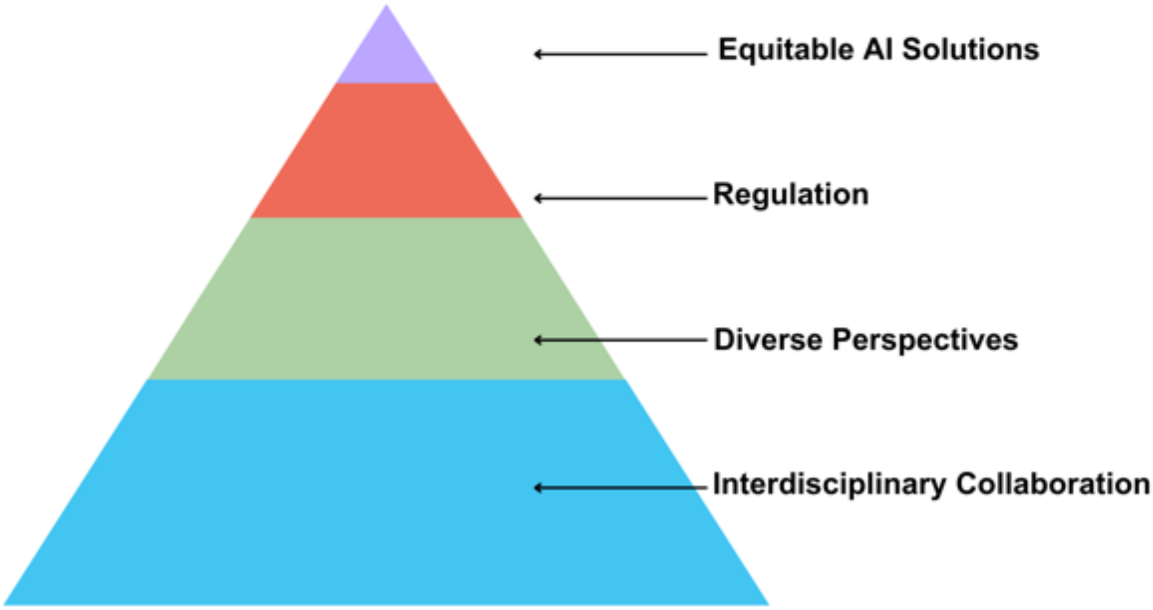

Figure 14

Addressing these issues requires human-centered, nonprofit initiatives dedicated to developing equitable AI solutions. Perhaps entire government departments (maybe a social AI department) looking for grassroots solutions within the community, for AI usage is likely to arise problems.

We may need the courage, vigour and energy to leave the Anglosphere, Anglocentric, and the Eurocentric procedures as the only possible solutions and work together and collaborative embracing Indigenous, and linguistically and culturally pluralistic scenarios and possibilities.

Summary :

In conclusion, this thesis has explored and addressed essential research questions in the field of Artificial Intelligence, its biases and how those biases influence users.

Given Australia's unique racial composition, predominantly composed of European descendants, and the fact that much research in this country is produced by those within the majority ingroup, it was critical to examine AI bias through the lens of Critical Race Theory.

Through a rigorous process of qualitative data collection, analysis, and interpretation, this study has illuminated significant phenomena and their implications for Australian society. The importance of considering Indigenous solutions and perspectives cannot be overstated, as exemplified by initiatives such as New Zealand's Mukurtu project (Hall, 2017) and others (Abdilla et al., 2021).

As AI becomes increasingly ubiquitous, fostering social cohesion in Australia and globally requires a commitment to developing fairer and more equitable systems. While this study has made valuable contributions to the fields of AI bias, AI ethics, critical algorithm studies, Critical Race Theory, and whiteness studies, it is important to acknowledge its limitations.

There is substantial scope for further research, particularly into the themes and questions that emerged during this investigation. Further studies are necessary to continue advancing toward a more just and inclusive future.

One thing is worth mentioning is that most of the research into AI bias I could encounter for my reference in this thesis, was done by women. I wonder if this is a consequence or a causality.

Thank you | Obrigado.

Exemplar Questions and Prompts:

Title: Can Bias Embedded in Image-Generative AI Systems Influence Public Perception?

Exemplar questions:

What is the first time you have heard or or what is your earliest memory about Artificial Intelligence?

Please describe what you know about imagery generative systems and how they work.

Is the image credible/believable to you? In what ways can you tell this?

Please describe, as clearly as you can, why this image seems accurate/natural to you. If the opposite, please state why.

How do you think this AI can be used for work and workers in the present and future?

In what ways could AI imagery generative systems be used in your work today?

In which ways can the images that come out of the imagery generative systems have an impact on the real world?

Does it matter what kind of images are generated? If it does why? If it doesn’t, why? When looking at these images, can you give me examples and describe different ways the image be represented?

What are your suggestions for improving the image generative systems, if any?

Thinking of your personal experience in Australia: In your words, does this image represent an accurate portrait of what the population of Australia looks like today?

In your opinion, do you feel anyone could feel discriminated against by the image? In what ways?

Part of this research is attempting to understand the best ways to utilised AI imagine generative system and support people who are and will use this technology.

What should be the steps in order to mitigate different racial differences/disparities in AI imagery generative systems?

If you are comfortable with it, would you be able to describe your ethnicity, age, gender, religion, and race?

Prompts:

1) young attractive Australian women looking into the camera.

2) Happy Australian school kids playing on school playground looking towards the camera.

3) Portrait of an Aboriginal Australian working Melbourne, Australia on a busy day, Cinematic, clear facial features, 35mm, f/1.8, accent lighting, global illumination, - uplight - V4.

4) Portrait of a family of recently arrived Australian immigrants leaving their suburban Australian homes going to pray. Cinematic, clear facial features, 35mm, f/1.8, accent lighting, global illumination, - uplight - V4.

5) Portrait of an Australian university student working on their assignments within an Australian university surrounded by other Australian university students. Cinematic, clear facial features, 35mm, f/1.8, accent lighting, global illumination, - uplight - V4.

Advertising Recruitment Materials

Advertising/Recruitment Materials

Advertising will be done via email or social media posts with the following information below:

You are invited to participate in a research project entitled:
Examining Bias Impact and Attitudes on Australian University Students When Utilising Imagery Generative AI Systems

This project is being conducted by a student researcher, Guido Oliveira Andrade de Melo, as part of a Master of Research study at Victoria University under the supervision of Dr Peter Thomas from the College of Arts and Education | Co-Chair of the B.Ed. P-12.

AI utilisation has become ubiquitous in many sectors of modern Western societies. Of late, it can be noticed that the utilisation of AI-powered tools is widespread (Sukhadeve, 2021). At the same time, several cases of AI bias, AI discrimination and automation bias have been exposed both by academic studies and mainstream media. This thesis will investigate if AI imagery generative systems can influence users (in this study's case, undergrad students) to make design decisions.

What will I be asked to do?

You will be invited to take part in an in-depth interview where the participants will be shown the same pre-generated images (3-5) with a set text-to-image context. Images will be generated using a prompt that will be disclosed to students and participants of the research before we show the images. The interviews will be held in person or via Zoom, and the interview is expected to be between 40-50 minutes and will be undertaken at a time to be negotiated between the researcher and the participant. All comments and notes written by the participants, as well as possible physical reactions to the images, may be utilised as a part of the study. For the safety and privacy of all participants involved, this study will use pseudonyms instead of real names when saving data.

What will I gain from participating?

Through participation in this research, you will be contributing to the theoretical and practical understanding of AI usage's effects on humans is an emerging and arguably important area of the research endeavour. You will be helping elucidate questions about AI usage. You may also feel empowered supporting the research of AI, as well as an increase in knowledge that came about as a result of the study, which may lead to better AI for humanity.

Information to Participants involved in research

You are invited to participate.

You are invited to participate in a research project entitled:
Examining Bias Impact and Attitudes on Australian University Students When Utilising Imagery Generative AI Systems.

This project is being conducted by a student researcher Guido Oliveira Andrade de Melo as part of a Master of Research study at Victoria University under the supervision of Dr Peter Thomas from the College of Arts and Education | Co-Chair of the B.Ed. P-12.

Project explanation

AI utilisation has become ubiquitous in many sectors of modern Western societies. Of late, it can be noticed that the utilisation of AI-powered tools is widespread (Sukhadeve, 2021). At the same time, several cases of AI bias, AI discrimination and automation bias have been exposed both by academic studies and mainstream media. This thesis will investigate if AI imagery generative systems can influence users (in this study's case, undergrad students) to make design decisions.

What will I be asked to do?

You will be invited to take part in an in-depth interview where the participants will be shown the same pre-generated images (3-5) with a set text-to-image context. Images will be generated using a prompt that will be disclosed to students and participants of the research before we show the images. The interviews will be held in person or via Zoom and is expected to be between 40-50 minutes and will be undertaken at a time to be negotiated between the researcher and the participant. All comments and notes written by the participants, as well as possible physical reactions to the images, may be utilised as a part of the study. For the safety and privacy of all participants involved, this study will use pseudonyms instead of real names when saving data.

What will I gain from participating?

Through participation in this research, you will be contributing to the theoretical and practical understanding of AI usage's effects on humans is an emerging and arguably important area of the research endeavour. You will be helping elucidate questions about AI usage. You may also feel empowered supporting the research of AI, as well as an increase in knowledge that came about as a result of the study, which may lead to better AI for humanity.

How will the information I give be used?

The information given will be used to support the development of AI usage as well as the potentiality for developing AI Tools and regulations. This study may potentially use some of the results for publications, including journal articles and research conferences.

What are the potential risks of participating in this project?

There are no foreseeable risks for participants in this project. Informed consent will be gained from all participants, with participation in the study being on a voluntary basis. All participants have the option to discontinue the study at any time without consequence. However, if participants experience any distress, they will be directed to the mental health support available in Australia.

How will this project be conducted?

This research includes 6-8 interviews with potential AI users who are currently enrolled on an undergraduate tertiary degree in Australia. The study will examine users on their interaction with image-generative systems.

Who is conducting the study?

Victoria University Institute for Sustainable Industries and Liveable Cities

Research Student Guido Oliveira Andrade de Melo, student at the Institute for Sustainable Industries and Liveable Cities, Victoria University, guido.oliveiraandradedemelo@live.vu.edu.au, or 0401 355 963

Chief Investigator Dr Peter Thomas from the College of Arts and Education | Co-Chair of the B.Ed. P-12, Victoria University peter.thomas@vu.edu.au

Any queries about your participation in this project may be directed to the Chief Investigator listed above.
If you have any queries or complaints about the way you have been treated, you may contact the Ethics Secretary, Victoria University Human Research Ethics Committee, Office for Research, Victoria University, PO Box 14428, Melbourne, VIC, 8001, email researchethics@vu.edu.au or phone (03) 9919 4781 or 4461.

Consent form for participants involved in research

INFORMATION TO PARTICIPANTS:

We would like to invite you to be a part of a study into:
Examining Bias Impact and Attitudes on Australian University Students When Utilising Imagery Generative AI Systems.

As AI utilisation becomes ubiquitous in our society and the utilisation of AI-powered tools is widespread, several cases of AI bias have been exposed in academia and the media. Arguably, AI is one of the most powerful tools known to humankind. The procedure will examine users on their interaction with image-generative systems.

If you agree to participate in this research, you will be asked to participate in 50 minutes interview. Interviews will preferably be conducted at a VU campus or online via Zoom in response to COVID-19 restrictions or personal preference. With permission, interviews will be audio-recorded. As it is an AI image generative study, some of the images involved in the study or situations discussed may cause distress. All comments and notes written by the participants, as well as possible physical reactions to the images, may be utilised as a part of the study.

Although this research does not anticipate it, reflecting on images could raise uncomfortable emotions, leading to psychological issues. You are welcome to end your participation in the project at any time without any penalties and do not have to answer any questions you do not want to answer. If you choose to withdraw at any stage of the research project, any data already collected will be destroyed. All identifying data will be removed from the documents and from our database collection. For your safety and privacy, we will use pseudonyms instead of the real names of participants when saving data.

If psychological or emotional distress is experienced, Victoria University offers free counselling and mental Health services. Another place to find support is Beyond Blue Support Service, which provides advice and support by calling 1300 22 4636. The service is available 24 hours a day, seven days a week. Alternatively, experienced online counsellors are here 24/7 at our Webchat Support Service.

All webchats are free, and you don't have to tell us your name if you don't want to.

While every effort will be made to ensure that your identity remains confidential and published materials are deidentified, there is always a risk that you may be identified, for example, if you publicly discuss your involvement in the project.

CERTIFICATION BY PARTICIPANT

I, "[Click here & type participant's name]"
of "[Click here & type participant's suburb]"

Certify that I am at least 18 years old* and that I am voluntarily giving my consent to participate in the study:
Examining Bias Impact and Attitudes on Australian University Students When Utilising Imagery Generative AI Systems
being conducted at Victoria University by:

I certify that the objectives of the study, together with any risks and safeguards associated with the procedures listed hereunder to be carried out in the research, have been fully explained to me by:

Guido Oliveira Andrade de Melo, student at the Institute for Sustainable Industries and Liveable Cities, Victoria University, guido.oliveiraandradedemelo@live.vu.edu.au, or 0401 355 963

Dr Peter Thomas from the College of Arts and Education | Co-Chair of the B.Ed. P-12, Victoria University peter.thomas@vu.edu.au,

and that I freely consent to participation involving the below mentioned procedures:

I certify that the objectives of the study, together with any risks and safeguards associated with the procedures listed hereunder to be carried out in the research, have been fully explained to me by the researcher and that I freely consent to participation involving the below mentioned procedures:

- One 50-minute semi-structured interview.

☐ I consent to my de-identified narratives being presented at conferences and training workshops and published in books and academic journals.

☐ I consent to be audio recorded during interviews.

I certify that I have had the opportunity to have any questions answered and that I understand that I can withdraw from this study at any time and that this withdrawal will not jeopardise me in any way.

I have been informed that the information I provide will be kept confidential.

Signed:

Date:

Any queries about your participation in this project may be directed to the researcher
:
Dr Peter Thomas from the College of Arts and Education | Co-Chair of the B.Ed. P-12, Victoria University peter.thomas@vu.edu.au,

If you have any queries or complaints about the way you have been treated, you may contact the Ethics Secretary, Victoria University Human Research Ethics Committee, Office for Research, Victoria University, PO Box 14428, Melbourne, VIC, 8001, email Researchethics@vu.edu.au or phone (03) 9919 4781 or 4461.

Ethics Application ID: HRE23-178

ResearchMaster

Ethics Application

Application ID :	HRE23-178
Application Title :	Examining Bias Impact and Attitudes on Australian University Students when Utilising Imagery Generative AI Systems.
Date of Submission :	16/11/2023
Primary Investigator :	DR PETER THOMAS (Chief Investigator)
Other Personnel :	DR WENJIE YE (Associate Investigator)
	Mr GUIDO OLIVEIRA ANDRADE DE MELO (Student)

References:

Abascal, T. E., Fluker, M., & Jiang, M. (2016). Domestic demand for Indigenous tourism in Australia: understanding intention to participate. Journal of Sustainable Tourism, 24(8-9), 1350-1368.

Abbas, N., Imran, A. S., Manzoor, R., Hussain, T., & Hussain, M. (2023). Role of Artificial Intelligence Tools in Enhancing Students' Educational Performance at Higher Levels. Journal of Artificial Intelligence Machine Learning and Neural Network(35), 36-49. https://doi.org/10.55529/jaimlnn.35.36.49

Abdilla, A., Kelleher, M., Shaw, R., & Yunkaporta, T. (2021). Out of the black box: indigenous protocols for AI. Melbourne: Deakin University (https://hdl. handle. net/10536/DRO/DU: 30159239).

Abraham, L., & Appiah, O. (2006). Framing news stories: The role of visual imagery in priming racial stereotypes. The Howard Journal of Communications, 17(3), 183-203.

ABS, A. B. o. S. (2022). Australia's Population by Country of Birth. ABS Retrieved from https://www.abs.gov.au/census/find-census-data/community-profiles/2021/AUS/download/GCP_AUS.xlsx

Aggarwal, R., Farag, S., Martin, G., Ashrafian, H., & Darzi, A. (2021). Patient Perceptions on Data Sharing and Applying Artificial Intelligence to Health Care Data: Cross-Sectional Survey. Journal of Medical Internet Research, 23(8), e26162. https://doi.org/10.2196/26162

Agnes, M. (2003). Webster's new world dictionary. Simon and Schuster.

Agrahari, S. K. (2024). The Expanding Role of Artificial Intelligence: Opportunities and Challenges. As the editors of Transforming Learning: The Power of Educational, 75.

Agrawal, A., Gupta, G., Agrawal, A., & Gupta, H. (2024). Evaluating Diversity and Stereotypes Amongst Healthcare Providers Using Generative AI. https://doi.org/10.21203/rs.3.rs-4259586/v1

Ajunwa, I. (2020). The paradox of automation as anti-bias intervention, 41 Cardozo, L. In: L.

Albarrán, I., Molina, J. M., & Gijón, C. (2021). Perception of Artificial Intelligence in Spain. Telematics and Informatics, 63, 101672. https://doi.org/10.1016/j.tele.2021.101672

Ali, F., Awais, M., & Faran, M. (2021). Social media use and political polarization: The mediating role of political engagement and political loyalty. International Journal of Media and Information Literacy, 6(1), 34-45.

Ali, R., Tang, O. Y., Connolly, I. D., Abdulrazeq, H., Mirza, F. N., Lim, R., Johnston, B. R., Groff, M. W., Williamson, T., Svokos, K., Libby, T. J., Shin, J. H., Gokaslan, Z. L., Doberstein, C. E., Zou, J., & Asaad, W. F. (2023). The Face of a Surgeon: An Analysis of Demographic Representation in Three Leading Artificial Intelligence Text-to-Image Generators. https://doi.org/10.1101/2023.05.24.23290463

Ali, R., Tang, O. Y., Connolly, I. D., Abdulrazeq, H. F., Mirza, F. N., Lim, R. K., Johnston, B. R., Groff, M. W., Williamson, T., Svokos, K., Libby, T. J., Shin, J. H., Gokaslan, Z. L., Doberstein, C. E., Zou, J., & Asaad, W. F. (2024). Demographic Representation in 3 Leading Artificial Intelligence Text-to-Image Generators. Jama Surgery, 159(1), 87. https://doi.org/10.1001/jamasurg.2023.5695

Allemang, B., Sitter, K., & Dimitropoulos, G. (2022). Pragmatism as a paradigm for patient-oriented research. Health Expectations, 25(1), 38-47.

Allen, R., & Wiles, J. (2015). A Rose by Any Other Name: Participants Choosing Research Pseudonyms. Qualitative Research in Psychology, 13(2), 149-165. https://doi.org/10.1080/14780887.2015.1133746

Almeida, S. (2019). Racismo estrutural. Pólen Produção Editorial LTDA.

Alrayes, A., Henari, T. F., & Ahmed, D. A. (2024). ChatGPT in Education – Understanding the Bahraini Academics Perspective. The Electronic Journal of E-Learning, 22(2), 112-134. https://doi.org/10.34190/ejel.22.2.3250

Anderson, R. E., Ahn, L. H., Brooks, J. R., Charity-Parker, B., Inniss-Thompson, M., Gumudavelly, D., Mitchell, S., & Anyiwo, N. (2023). “The talk” tells the story: A qualitative investigation of parents’ racial socialization competency with Black adolescents. Journal of Adolescent Research, 38(3), 562-588.

Andreotta, A. J., Kirkham, N., & Rizzi, M. (2022). AI, big data, and the future of consent. AI & SOCIETY, 37(4), 1715-1728. https://doi.org/10.1007/s00146-021-01262-5

Andrews, K. (2023). The psychosis of whiteness: Surviving the insanity of a racist world. Random House.

Ang, S., Song, J., & Pan, Q. (2024). Pandemic racism and sexism in Australia: Responses and reflections among Asian women. Current Sociology, 72(3), 463-481.

Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2022). Machine bias. In Ethics of data and analytics (pp. 254-264). Auerbach Publications.

Ani, M. (1994). Yurugu: An African-centered critique of European cultural thought and behavior (Vol. 213). Africa World Press Trenton, NJ.

Annamma, S. A., Jackson, D. D., & Morrison, D. (2017). Conceptualizing color-evasiveness: Using dis/ability critical race theory to expand a color-blind racial ideology in education and society. Race Ethnicity and Education, 20(2), 147-162.

APH. African Australian Communities Leadership Forum Preliminary Community Issues.

Applebaum, B. (2022). Willful Hermeneutical Ignorance and the “Critical Race Theory” Controversy. Educational Theory.

Applebaum, B. (2023). Interview with Barbara Applebaum: “Willful Ignorance and the Critical Race Theory Controversy”. Retrieved 25/03/2023 from https://www.youtube.com/watch?v=npQXTt86I4o&t=1s

Arantes, J. A. (2020). Big Data, Black Boxes and Bias The University of Newcastle, Australia].

Asadi, Z. K., & Milani, F. (2022). Racial and Gender Disparities in the Portrayal of Lawyers and Physicians on Television. https://doi.org/10.59720/22-093

Aspinall, P. J. (1998). Describing the “white” ethnic group and its composition in medical research. Social Science & Medicine, 47(11), 1797-1808.

Atanasoski, N., & Vora, K. (2019). Surrogate humanity: Race, robots, and the politics of technological futures. Duke University Press.

Austin, C., & Fozdar, F. (2018). Australian national identity: Empirical research since 1998. National Identities, 20(3), 277-298.

Awasthi, P., & George, J. J. (2020). A case for data democratization.

Baas, M. (2015). The question of racism: How to understand the violent attacks on Indian students in Australia? Cosmopolitan Civil Societies: An Interdisciplinary Journal, 7(3), 37-60.

Bailey, Z. D., Feldman, J. M., & Bassett, M. T. (2021). How structural racism works—racist policies as a root cause of US racial health inequities. In (Vol. 384, pp. 768-773): Mass Medical Soc.

Baker, A., & Cupery, D. (2023). Animosity, amnesia, or admiration? Mass opinion around the world toward the former colonizer. British Journal of Political Science, 53(4), 1132-1149.

Balla, P. (2020). Disrupting artistic terra nullius: The ways that first nations women in art & community speak blak to the colony & patriarchy Victoria University].

Banerjee, I., Bhimireddy, A. R., Burns, J. L., Celi, L. A., Chen, L.-C., Correa, R., Dullerud, N., Ghassemi, M., Huang, S.-C., & Kuo, P.-C. (2021). Reading race: AI recognises patient's racial identity in medical images. arXiv preprint arXiv:2107.10356.

Bannister, D., & Fransella, F. (2019). Inquiring man: The psychology of personal constructs. Routledge.

Bargallie, D., Fernando, N., & Lentin, A. (2024). Breaking the racial silence: Putting racial literacy to work in Australia. Ethnic and Racial studies, 47(8), 1532-1551.

Barrat, J. (2023). Our final invention: Artificial intelligence and the end of the human era. Hachette UK.

Barrie, C. (2020). Searching racism after George Floyd. Socius, 6, 2378023120971507.

Barton, B. A., Adams, K. S., Browne, B. L., & Arrastia-Chisholm, M. C. (2018). The Effects of Social Media Usage on Attention, Motivation, and Academic Performance. Active Learning in Higher Education, 22(1), 11-22. https://doi.org/10.1177/1469787418782817

Basch, C. H., Donelle, L., Fera, J., & Jaime, C. (2022). Deconstructing TikTok Videos on Mental Health: Cross-Sectional, Descriptive Content Analysis. Jmir Formative Research, 6(5), e38340. https://doi.org/10.2196/38340

Bass, L. N. a. D. (2023). Humans are biased, generative AI is even worse. bloomberg. Retrieved 11/08/2023 from

Bays, H. E., Gonsahn-Bollie, S., Younglove, C., & Wharton, S. (2022). Obesity Pillars Roundtable: Body mass index and body composition in Black and Female individuals. Race-relevant or racist? Sex-relevant or sexist? Obesity Pillars, 4, 100044.

Beard, M. (2022). The Importance of Academic Impartiality. Quillette. Retrieved 23/09/2023 from https://quillette.com/2022/01/15/the-importance-of-academic-impartiality/

Belciug, S., & Iliescu, D. Pregnancy with Artificial Intelligence.

Belciug, S., & Iliescu, D. (2022). Pregnancy with artificial intelligence: a 9.5 months journey from preconception to birth (Vol. 234). Springer Nature.

Belenguer, L. (2022). AI bias: exploring discriminatory algorithmic decision-making models and the application of possible machine-centric solutions adapted from the pharmaceutical industry. AI and Ethics, 2(4), 771-787.

Bell, D. (2018). Faces at the bottom of the well: The permanence of racism. Hachette UK.

Bell, D. A. (1995). Who's afraid of critical race theory. U. Ill. L. Rev., 893.

Bell, G., Broad, E., Martin, B., O'Brien, E., Parsons, J., & Zafiroglu, A. (2018). Gender and Artificial Intelligence. The International Encyclopedia of Anthropology, 1-11.

Beltrán, C. (2020). Cruelty as citizenship: How migrant suffering sustains white democracy. U of Minnesota Press.

Ben, J., Elias, A., Sharples, R., Dunn, K., Truong, M., Mansouri, F., Denson, N., Walton, J., & Paradies, Y. (2024). Racism data in Australia: A review of quantitative studies and directions for future research. Journal of Intercultural Studies, 45(2), 228-257.

Bender, E. M., & Friedman, B. (2018). Data statements for natural language processing: Toward mitigating system bias and enabling better science. Transactions of the Association for Computational Linguistics, 6, 587-604.

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?□ü¶ú. Proceedings of the 2021 ACM conference on fairness, accountability, and transparency,

Benefo, E. O., Tingler, A., White, M., Cover, J. D., Torres, L., Broussard, C., Shirmohammadi, A., Pradhan, A. K., & Patra, D. (2022). Ethical, Legal, Social, and Economic (ELSE) Implications of Artificial Intelligence at a Global Level: A Scientometrics Approach. AI and Ethics, 2(4), 667-682. https://doi.org/10.1007/s43681-021-00124-6

Benjamin, R. (2023). Race after technology. In Social Theory Re-Wired (pp. 405-415). Routledge.

Bennett, B., Menzel, K., Prehn, J., & Gates, T. G. (2023). Australian universities, indigenization, whiteness, and settler colonial epistemic violence. In Handbook of Critical Whiteness: Deconstructing Dominant Discourses Across Disciplines (pp. 1-14). Springer.

Bennett, R., Uink, B., & Van den Berg, C. (2021). Educating Rita at the cultural interface: Exploring intersections between race and gender in the experiences of Australian Aboriginal women at university. Diaspora, Indigenous, and Minority Education, 15(2), 84-98.

Bernardino-Costa, J., Maldonado-Torres, N., & Grosfoguel, R. (2018). Decolonialidade e pensamento afrodiaspórico. Autêntica.

Bernasconi, R. (2013). Critical philosophy of race. In The Routledge Companion to Phenomenology (pp. 551-562). Routledge.

Bhaimiya, S. (2023). An Asian MIT student asked AI to turn an image of her into a professional headshot. It made her white, with lighter skin and blue eyes. In: Business Insider. https://www. businessinsider. com/student-uses-playrgound ….

Bhandari, P. (2022). Triangulation in research| guide, types, examples. Re-trieved from Scribbr: https://www. scribbr. com/methodology/triangulation.

Bianchi, F., Kalluri, P., Durmus, E., Ladhak, F., Cheng, M., Nozza, D., Hashimoto, T., Jurafsky, D., Zou, J., & Caliskan, A. (2023). Easily accessible text-to-image generation amplifies demographic stereotypes at large scale. Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency,

Birhane, A. (2021). Algorithmic injustice: a relational ethics approach. Patterns, 2(2), 100205.

Birhane, A., Steed, R., Ojewale, V., Vecchione, B., & Raji, I. D. AI auditing: The Broken Bus on the Road to AI Accountability (Jan 2024). URL http://arxiv. org/abs/2401.14462.

Bledsoe, A., & Wright, W. J. (2019). The anti-Blackness of global capital. Environment and Planning D: Society and Space, 37(1), 8-26.

Bolukbasi, T., Chang, K.-W., Zou, J., Saligrama, V., & Kalai, A. (2016). Man is to computer programmer as woman is to homemaker. Debiasing Word Embeddings, 29.

Bonezzi, A., & Ostinelli, M. (2021). Can Algorithms Legitimize Discrimination? Journal of Experimental Psychology Applied, 27(2), 447-459. https://doi.org/10.1037/xap0000294

Bonilla-Silva, E. (2006). Racism without racists: Color-blind racism and the persistence of racial inequality in the United States. Rowman & Littlefield Publishers.

Bonilla-Silva, E., & Dietrich, D. (2011). The sweet enchantment of color-blind racism in Obamerica. The ANNALS of the American Academy of Political and Social Science, 634(1), 190-206.

Bonnett, A. (1998). Who was white? The disappearance of non-European white identities and the formation of European racial whiteness. Ethnic and Racial studies, 21(6), 1029-1055.

Bourke, B. (2014). Positionality: Reflecting on the research process. The qualitative report, 19(33), 1-9.

Bovet, A., & Makse, H. A. (2019). Influence of fake news in Twitter during the 2016 US presidential election. Nature communications, 10(1), 7.

Boydell, N., Fergie, G., McDaid, L., & Hilton, S. (2015). Understandings of Participation in Behavioural Research: A Qualitative Study of Gay and Bisexual Men in Scotland. Plos One, 10(8), e0135001. https://doi.org/10.1371/journal.pone.0135001

Bragazzi, N. L., Crapanzano, A., Converti, M., Zerbetto, R., & Khamisy-Farah, R. (2023). The impact of generative conversational artificial intelligence on the Lesbian, gay, Bisexual, transgender, and queer community: scoping review. Journal of Medical Internet Research, 25, e52091.

Brear, M. (2017). Swazi Co-Researcher Participants' Dynamic Preferences and Motivations For, Representation With Real Names and (English-Language) Pseudonyms – An Ethnography. Qualitative Research, 18(6), 722-740. https://doi.org/10.1177/1468794117743467

Brew, M., Taylor, S., Lam, R., Havemann, L., & Nerantzi, C. (2023). Towards developing AI literacy: three student provocations on AI in higher education. Asian Journal of Distance Education, 18(2), 1-11.

Briese, J., Elliott, L. M., & Duthie, D. (2024). Resistance, White fragility, and fear of the unknown in tertiary settings: a recipe for Blak fatigue. In Handbook of

Critical Whiteness: Deconstructing Dominant Discourses Across Disciplines (pp. 181-196). Springer.

Broad, E. (2018). Made by humans: The AI condition. Melbourne Univ. Publishing.

Brooks, A. (2012). Black negrophobia and black self-empowerment: Afro-descendant responses to societal racism in Sao Paulo, Brazil. UW-L Journal of Undergraduate Research, 25(2), 65-86.

Brooks, R. A. (1999). Cambrian intelligence: The early history of the new AI. MIT press.

Broussard, M. (2018). Artificial unintelligence: How computers misunderstand the world. MIT Press.

Broussard, M. (2019). Letting Go of Technochauvinism. publicbooks. https://www.publicbooks.org/letting-go-of-technochauvinism/

Browne, S. (2015). Dark matters: On the surveillance of blackness. Duke University Press.

Buchinskaia, O., & Stremousova, E. (2021). Inequality of Digital Access Between Russian Regions. E3s Web of Conferences, 301, 05001. https://doi.org/10.1051/e3sconf/202130105001

Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. Conference on fairness, accountability and transparency,

Campbell, R. D., Dennis, M. K., Lopez, K., Matthew, R., & Choi, Y. J. (2021). Qualitative research in communities of color: Challenges, strategies, and lessons. Journal of the Society for Social Work and Research, 12(1), 177-200.

Campón, R. R., & Carter, R. T. (2015). The Appropriated Racial Oppression Scale: Development and preliminary validation. Cultural Diversity and Ethnic Minority Psychology, 21(4), 497.

Carpenter, S., & Lertpratchya, A. P. (2016). A qualitative and quantitative study of social media communicators: An extension of role theory to digital media workers. Journal of Broadcasting & Electronic Media, 60(3), 448-464.

Carroll, J. M. (2009). Human computer interaction (HCI). Encyclopedia of Human-Computer Interaction. Aarhus: The Interaction Design Foundation.

Carstens, D. S., Doss, S. K., & Kies, S. C. (2018). Social media impact on attention span. Journal of management & engineering integration, 11(1).

Carswell, S. (2020). Why news organizations' move to capitalize'Black'is a win. The Washington Post, NA-NA.

Cave, S., & Dihal, K. (2020). The Whiteness of AI. Philosophy & Technology, 33(4), 685-703. https://doi.org/10.1007/s13347-020-00415-6

Celi, L. A., Cellini, J., Charpignon, M.-L., Dee, E. C., Dernoncourt, F., Eber, R., Mitchell, W., Moukheiber, L., Schirmer, J., Situ, J., Paguio, J. A., Park, J., Wawira, J. G., & Yao, J. S. (2022). Sources of Bias in Artificial Intelligence That Perpetuate Healthcare Disparities—A Global Review. Plos Digital Health, 1(3), e0000022. https://doi.org/10.1371/journal.pdig.0000022

Cepeda, C., Dias, M. C., Rindlisbacher, D., Gamboa, H., & Cheetham, M. (2021). Knowledge Extraction From Pointer Movements and Its Application to Detect Uncertainty. Heliyon, 7(1), e05873. https://doi.org/10.1016/j.heliyon.2020.e05873

Chan, C. K. Y. (2023). A comprehensive AI policy education framework for university teaching and learning. International journal of educational technology in higher education, 20(1), 38.

Chan, C. K. Y., & Hu, W. (2023). Students’ Voices on Generative AI: Perceptions, Benefits, and Challenges in Higher Education. International journal of educational technology in higher education, 20(1). https://doi.org/10.1186/s41239-023-00411-8

Chapman, D., & de Melo, G. O. A. (2022). Zooming while Black: creating a Black Aesthetic and counter hegemonic discourse in a digital age. In The Routledge Handbook of Media Education Futures Post-Pandemic (pp. 37-45). Routledge.

ChatGPT in Higher Education: Considerations for Academic Integrity and Student Learning. (2023). Journal of Applied Learning & Teaching, 6(1). https://doi.org/10.37074/jalt.2023.6.1.17

Chavan, M., Bowden-Everson, J., Lundmark, E., & Zwar, J. (2014). Exploring the drivers of service quality perceptions in the tertiary education sector: comparing domestic Australian and international Asian students. Journal of International Education in Business, 7(2), 150-180.

Chellappa, R., & Niiler, E. (2022). Can We Trust AI? JHU Press.

Chen, M. S., Lara, P. N., Dang, J., Paterniti, D. A., & Kelly, K. (2014). Twenty Years Post-NIH Revitalization Act: Enhancing Minority Participation in Clinical Trials

(EMPaCT): Laying the Groundwork for Improving Minority Clinical Trial Accrual. Cancer, 120(S7), 1091-1096. https://doi.org/10.1002/cncr.28575

Chen, P., Wu, L., & Wang, L. (2023). AI Fairness in Data Management and Analytics: A Review on Challenges, Methodologies and Applications. Applied Sciences, 13(18), 10258. https://doi.org/10.3390/app131810258

Cheng, Z., & Li, Y. (2024). Like, comment, and share on TikTok: Exploring the effect of sentiment and second-person view on the user engagement with TikTok news videos. Social Science Computer Review, 42(1), 201-223.

Cheryl Staats, K. C., Lena Tenney, and Sarah Mamo. (2017a). Implicit Bias Review.

Cheryl Staats, K. C., Lena Tenney, and Sarah Mamo. (2017b). Race and Ethnicity: Views from Inside the Unconscious Mind. STATE OF THE SCIENCE: IMPLICIT BIAS REVIEW https://offices.depaul.edu/academic-affairs/academic-diversity/Documents/State%20of%20Science%20Implicit%20Bias%20Review.pdf (The Ohio State University)

Cho, J., Ahmed, S., Keum, H., Choi, Y. J., & Lee, J. H. (2016). Influencing Myself: Self-Reinforcement Through Online Political Expression. Communication Research, 45(1), 83-111. https://doi.org/10.1177/0093650216644020

Chubb, J., Cowling, P., & Reed, D. (2022). Speeding up to keep up: exploring the use of AI in the research process. AI & SOCIETY, 37(4), 1439-1457.

Chui, M., Hazan, E., Roberts, R., Singla, A., Smaje, K., Sukharevsky, A., Yee, L., & Zemmel, R. (2023). The economic potential of generative AI: The next productivity frontier. McKinsey & Company. In.

Cofey, D. (2021). Maori are trying to save their language from Big Tech. Wired UK (April 2021).

Cole, M. (2021). Understanding Critical Whiteness Studies: Harmful or Helpful in the Struggle for Racial Equity in the Academy? In Doing Equity and Diversity for Success in Higher Education: Redressing Structural Inequalities in the Academy (pp. 277-298). Springer.

Collins, C. S., & Stockton, C. M. (2018). The central role of theory in qualitative research. International journal of qualitative methods, 17(1), 1609406918797475.

Collins, P. H. (2022). Black feminist thought: Knowledge, consciousness, and the politics of empowerment. routledge.

Corl, F. M., Garland, M., Lawler, L. P., & Fishman, E. K. (2002). A Five-Step Approach to Digital Image Manipulation for the Radiologist. Radiographics, 22(4), 981-992. https://doi.org/10.1148/radiographics.22.4.g02jl33981

Council, R. (2017). Uluru statement from the heart.

Crawford, C. E. (2023). Technology Producers Use of Language and Discourse to Shape and Reinstate Anti-Black Global Realities: An Analysis of Amazon’s Facial Recognition Technology Communications and Responses to Racial Bias in Rekognition. https://doi.org/10.32920/ryerson.14647572.v1

Crawford, K. (2021). The Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence. Yale University Press.

Crenshaw, K. W. (2017). On intersectionality: Essential writings. The New Press.

Creswell, A., White, T., Dumoulin, V., Arulkumaran, K., Sengupta, B., & Bharath, A. A. (2018). Generative adversarial networks: An overview. IEEE signal processing magazine, 35(1), 53-65.

Creswell, J. W., & Creswell, J. (2014). Research design, qualitative, quantitative, and mexed methods approaches (I. Sage Publications. In: Inc.

Creswell, J. W., & Creswell, J. D. (2017). Research design: Qualitative, quantitative, and mixed methods approaches. Sage publications.

Creswell, J. W., & Poth, C. N. (2016). Qualitative inquiry and research design: Choosing among five approaches. Sage publications.

Crockett, K., Garratt, M., Latham, A., Colyer, E., & Goltz, S. (2020). Risk and Trust Perceptions of the Public of Artifical Intelligence Applications. 1-8. https://doi.org/10.1109/ijcnn48605.2020.9207654

Curry, T. J. (2017). The man-not: Race, class, genre, and the dilemmas of Black manhood. Temple University Press.

D’Arrietta, L. M., Vangaveti, V., Crowe, M., & Malau-Aduli, B. S. (2022). Impact of Research Training on Newly Graduated Health Professionals’ Motivation to Undertake Research. Journal of Multidisciplinary Healthcare, Volume 15, 2223-2240. https://doi.org/10.2147/jmdh.s377963

D'ignazio, C., & Klein, L. F. (2020). Data feminism. MIT press.

Dancy, C. L., & Saucier, P. K. (2021). AI and blackness: toward moving beyond bias and representation. IEEE Transactions on Technology and Society, 3(1), 31-40.

Danks, D., & London, A. J. (2017). Algorithmic Bias in Autonomous Systems. Ijcai, Dann, C., Redmond, P., Fanshawe, M., Brown, A., Getenet, S., Shaik, T., Tao, X., Galligan, L., & Li, Y. (2024). Making Sense of Student Feedback and Engagement Using Artificial Intelligence. Australasian Journal of Educational Technology. https://doi.org/10.14742/ajet.8903

David, E. R., Schroeder, T. M., & Fernandez, J. (2019). Internalized racism: A systematic review of the psychological literature on racism's most insidious consequence. Journal of Social Issues, 75(4), 1057-1086.

Dawson, D., Schleiger, E., Horton, J., McLaughlin, J., Robinson, C., Quezada, G., Scowcroft, J., & Hajkowicz, S. (2019). Artificial intelligence: Australia's ethics framework-a discussion paper.

Delgado, R. (1989). Storytelling for oppositionists and others: A plea for narrative. Michigan law review, 87(8), 2411-2441.

Delgado, R., & Stefancic, J. (2023). Critical race theory: An introduction (Vol. 87). NyU press.

Deng, Q., & Li, M. (2013). A Model of Event–Destination Image Transfer. Journal of Travel Research, 53(1), 69-82. https://doi.org/10.1177/0047287513491331

Denzin, N. (2010). The fundamentals of Triangulation. An Introduction to Triangulation, D. Rugg, editor. UNAIDS Monitoring and Evaluation Fundamentals, Geneva, Switzerland, 12.

Deshpande, K. V., Pan, S., & Foulds, J. R. (2020). Mitigating demographic Bias in AI-based resume filtering. Adjunct publication of the 28th ACM conference on user modeling, adaptation and personalization,

Dhali, A., D'Souza, C., Rathna, R. B., Biswas, J., & Dhali, G. K. (2022). Authorship Diversity in Gastroenterology-Related Cochrane Systematic Reviews: Inequities in Global Representation. Frontiers in Medicine, 9. https://doi.org/10.3389/fmed.2022.982664

DiAngelo, R. (2022). White Fragility (Adapted for Young Adults): Why Understanding Racism Can Be So Hard for White People (Adapted for Young Adults). Beacon Press.

Dillenburger, K., Röttgers, H.-R., Dounavi, K., Sparkman, C., Keenan, M., Thyer, B., & Nikopoulos, C. (2014). Multidisciplinary Teamwork in Autism: Can One Size Fit All? The Australian Educational and Developmental Psychologist, 31(2), 97-112. https://doi.org/10.1017/edp.2014.13

Dominguez-Catena, I., Paternain, D., & Galar, M. (2023). Metrics for Dataset Demographic Bias: A Case Study on Facial Expression Recognition. https://doi.org/10.48550/arxiv.2303.15889

Dougherty, G. B., Golden, S. H., Gross, A. L., Colantuoni, E., & Dean, L. T. (2020). Measuring structural racism and its association with BMI. American journal of preventive medicine, 59(4), 530-537.

Dove, N. (1998). African womanism: an Afrocentric theory. Journal of Black Studies, 28(5), 515-539.

Dr. Rohit Markan, D. S. P. S. D. A. K., & M. Rajalakshmi, R. K. D. S. N. (2024). Rise of Artificial Intelligence in Business and Industry. 4(2). https://doi.org/10.52783/jier.v4i2.850

Drage, E., & Mackereth, K. (2022). Does AI Debias Recruitment? Race, Gender, and AI's "Eradication of Difference". Philosophy & Technology, 35(4). https://doi.org/10.1007/s13347-022-00543-1

Duan, R., Takahashi, B., & Zwickle, A. (2021). How Effective Are Concrete and Abstract Climate Change Images? The Moderating Role of Construal Level in Climate Change Visual Communication. Science Communication, 43(3), 358-387. https://doi.org/10.1177/10755470211008192

Dubber, M. D., Pasquale, F., & Das, S. (2020). The Oxford handbook of ethics of AI. Oxford Handbooks.

Dubey, S., Sarawagi, A., & Shrivastava, M. (2016). Image Forgery Detection Based on Local Descriptors and Block-Matching Using Clustering Technique. International Journal of Computer Applications, 141(10), 11-14. https://doi.org/10.5120/ijca2016909808

Dumas, M. J. (2016). Against the dark: Antiblackness in education policy and discourse. Theory into practice, 55(1), 11-19.

Dunlop, S., Wakefield, M., & Kashima, Y. (2008). Can you feel it? Negative emotion, risk, and narrative in health communication. Media Psychology, 11(1), 52-75.

Dunn, K. M., Forrest, J., Burnley, I., McDonald, A., & Dunn, K. M. (2004). ``Xenophobia (racism) in Australia: please explain!''. Australian Journal of Social Issues, 39, 409-430.

Dunn, N., & Luchner, A. F. (2022). The Emotional Impact of Self-criticism on Self-reflection and Rumination. Psychology and Psychotherapy Theory Research and Practice, 95(4), 1126-1139. https://doi.org/10.1111/papt.12422

Dunning, D. (2011). The Dunning–Kruger effect: On being ignorant of one's own ignorance. In Advances in experimental social psychology (Vol. 44, pp. 247-296). Elsevier.

Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., Baabdullah, A. M., Koohang, A., Raghavan, V., & Ahuja, M. (2023). Opinion Paper:“So what if ChatGPT wrote it?” Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy.

Dyson, L. E., Grant, S., & Hendriks, M. A. (2016). Indigenous people and mobile technologies. Routledge New York, NY.

Ebell, C., Baeza-Yates, R., Benjamins, R., Cai, H., Coeckelbergh, M., Duarte, T., Hickok, M., Jacquet, A., Kim, A., & Krijger, J. (2021). Towards intellectual freedom in an AI Ethics Global Community. AI and Ethics, 1, 131-138.

Economist, T. (2022). Huge “foundation models” are turbo-charging AI progress. Retrieved 13/11/2024 from

Eidelman, S., & Crandall, C. S. (2012). Bias in favor of the status quo. Social and Personality Psychology Compass, 6(3), 270-281.

El Morr, C., Kundi, B., Mobeen, F., Taleghani, S., El-Lahib, Y., & Gorman, R. (2024). AI and disability: A systematic scoping review. Health Informatics Journal, 30(3), 14604582241285743.

El Naqa, I., & Murphy, M. J. (2015). What is machine learning? Springer.

Elias, A., Mansouri, F., & Paradies, Y. (2021). Racism in Australia today. Springer.

Elliott-Mainwaring, H. (2021). Exploring using NVivo software to facilitate inductive coding for thematic narrative synthesis. British Journal of Midwifery, 29(11), 628-632.

Ellison, C. G., McFarland, M. J., & Krause, N. (2011). Measuring religiousness among older African Americans: Exploring race-of-interviewer effects. Review of Religious Research, 53(1), 65-84.

Ellison, T. L. (2021). “Stop Whitewashing Our Stories”: Using Counter-Stories to Dismantle Racist and Deficit Perspectives in BIPOC Narratives. In Whitewashed Critical Perspectives (pp. 86-103). Routledge.

Enders, A. M., & Thornton, J. R. (2022). Biased Interviewer Assessments of Respondent Knowledge Based on Perceptions of Skin Tone. The Journal of Race Ethnicity and Politics, 7(3), 572-588. https://doi.org/10.1017/rep.2021.40

Entman, R. M., & Rojecki, A. (2001). The black image in the white mind: Media and race in America. University of Chicago Press.

Eubanks, V. (2018). Automating inequality: How high-tech tools profile, police, and punish the poor. St. Martin's Press.

Fanon, F. (1970). Black skin, white masks. Paladin London.

Fanon, F. (2004). The Wretched of the Earth. 1961. Trans. Richard Philcox. New York: Grove Press, 6.

Farahani, M. S., & Ghasemi, G. (2024). Artificial intelligence and inequality: challenges and opportunities. Int. J. Innov. Educ, 9, 78-99.

Favaretto, M., De Clercq, E., & Elger, B. S. (2019). Big Data and discrimination: perils, promises and solutions. A systematic review. Journal of Big Data, 6(1), 12. https://doi.org/10.1186/s40537-019-0177-4

Feehan, J., & Apostolopoulos, V. (2021). Is COVID-19 the worst pandemic? Maturitas, 149, 56.

Feijóo, C., Kwon, Y., Bauer, J. M., Bohlin, E., Howell, B., Jain, R., Potgieter, P., Vu, K., Whalley, J., & Xia, J. (2020). Harnessing artificial intelligence (AI) to increase wellbeing for all: The case for a new technology diplomacy. Telecommunications Policy, 44(6), 101988.

Ferdinand, A. S., Paradies, Y., & Kelaher, M. (2015). Mental health impacts of racial discrimination in Australian culturally and linguistically diverse communities: A cross-sectional survey. BMC Public Health, 15, 1-14.

Fereday, J., & Muir-Cochrane, E. (2006). Demonstrating rigor using thematic analysis: A hybrid approach of inductive and deductive coding and theme development. International journal of qualitative methods, 5(1), 80-92.

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

Ferrer, X., Nuenen, T. v., Such, J. M., Coté, M., & Criado, N. (2021). Bias and Discrimination in AI: A Cross-Disciplinary Perspective. IEEE Technology and Society Magazine, 40(2), 72-80. https://doi.org/10.1109/MTS.2021.3056293

Finkel, S. E., Guterbock, T. M., & Borg, M. J. (1991). Race-of-interviewer effects in a preelection poll Virginia 1989. Public opinion quarterly, 55(3), 313-330.

Finn Lattimore, S. O. C., Zoe Paleologos, Alistair Reid, Edward Santow, Holli Sargeant and Andrew Thomsen (2020). Using artificial intelligence to make decisions: Addressing the problem of algorithmic bias. Technical Paper. https://humanrights.gov.au/our-work/technology-and-human-rights/publications/technical-paper-addressing-algorithmic-bias

Floridi, L. (2012). Big Data and Their Epistemological Challenge. Philosophy & Technology, 25(4), 435-437. https://doi.org/10.1007/s13347-012-0093-4

Foley, D. (2000). 'Too white to be black, too black to be white'. Social Alternatives, 19(4), 44-49.

Ford, C. L., & Airhihenbuwa, C. O. (2010). Critical Race Theory, Race Equity, and Public Health: Toward Antiracism Praxis. American Journal of Public Health, 100(S1), S30-S35. https://doi.org/10.2105/ajph.2009.171058

Forrest, J., & Dunn, K. (2010). Attitudes to multicultural values in diverse spaces in Australia's immigrant cities, Sydney and Melbourne. Space and polity, 14(1), 81-102.

Fošner, A. (2024). University Students’ Attitudes and Perceptions towards AI Tools: Implications for Sustainable Educational Practices. Sustainability, 16(19), 8668.

Frankenberg, R. (2004). On unsteady ground: Crafting and engaging in the critical study of whiteness 1. In Researching race and racism (pp. 104-118). Routledge.

Freire, P. (2000). Pedagogy of freedom: Ethics, democracy, and civic courage. Rowman & Littlefield Publishers.

Fry, C. L. (2003). From technology to transcendence: humanity's evolutionary journey in 2001: A Space Odyssey. Extrapolation (pre-2012), 44(3), 331.

Funderburg, L. (1994). Black, white, other: Biracial Americans talk about race and identity. (No Title).

Ganasegeran, K., & Abdulrahman, S. A. (2020). Artificial intelligence applications in tracking health behaviors during disease epidemics. Human Behaviour Analysis Using Intelligent Systems, 141-155.

Garcia, D. G. (2008). Culture clash invades Miami: Oral histories and ethnography center stage. Qualitative inquiry, 14(6), 865-895.

Gates, R. R. (1960). The genetics of the Australian aborigines. Acta geneticae medicae et gemellologiae: twin research, 9(1), 7-50.

Gatwiri, K., & Anderson, L. (2022). Afrodiasporic identities in Australia: articulations of Blackness and Africanness. Springer.

Gatwiri, K., Mwanri, L., & McPherson, L. (2021). Afro-diasporic experiences of highly skilled Black African immigrants in Australia. Australian Social Work, 74(4), 480-491.

Gebru, T., & Denton, R. (2024). Beyond Fairness in Computer Vision: A Holistic Approach to Mitigating Harms and Fostering Community-Rooted Computer Vision Research. Foundations and Trends® in Computer Graphics and Vision, 16(3), 215-321.

Gibney, E. (2024). Chatbot AI makes racist judgements on the basis of dialect. Nature, 627(8004), 476-477.

Gichoya, J. W., Banerjee, I., Bhimireddy, A. R., Burns, J. L., Celi, L. A., Chen, L.-C., Correa, R., Dullerud, N., Ghassemi, M., & Huang, S.-C. (2022). AI recognition of patient race in medical imaging: a modelling study. The Lancet Digital Health, 4(6), e406-e414.

Gichoya, J. W., Thomas, K., Celi, L. A., Safdar, N., Banerjee, I., Banja, J. D., Seyyed-Kalantari, L., Trivedi, H., & Purkayastha, S. (2023). AI Pitfalls and What Not to Do: Mitigating Bias in AI. British Journal of Radiology, 96(1150). https://doi.org/10.1259/bjr.20230023

Gilbert, E., O'Shea, M., & Duffy, S. (2021). Gender Equality Mainstreaming and the Australian Academy: Paradoxical Effects? Discover Psychology, 1(1). https://doi.org/10.1007/s44202-021-00008-0

Godsil, R. D., Tropp, L. R., Goff, P. A., & Powell, J. A. (2014). Addressing implicit bias, racial anxiety, and stereotype threat in education and health care. The Science of Equality, 1(November), 1-90.

Gomez, L., & Bernet, P. (2019). Diversity improves performance and outcomes. Journal of the National Medical Association, 111(4), 383-392.

Gong, M., & Aadland, D. (2011). Interview effects in an environmental valuation telephone survey. Environmental and Resource Economics, 49(1), 47-64.

Goodfellow, I. (2016). Deep learning-ian goodfellow, yoshua bengio, aaron courville-google books. In: MIT Press Cambridge, MA, USA.

Goswami, R. (2023). Bill Gates thinks A.I. like ChatGPT is the 'most important' innovation right now. cnbc. Retrieved 06/08/2023 from https://www.cnbc.com/2023/02/10/bill-gates-says-ai-like-chatgpt-is-the-most-important-innovation.html

Grinberg, K. (2002). O fiador dos brasileiros: cidadania, escravidão e direito civil no tempo de Antonio Pereira Rebouças. (No Title).

Grosfoguel, R. (2015). Epistemic racism/sexism, westernized universities and the four genocides/epistemicides of the long sixteenth century. Eurocentrism, racism and knowledge: Debates on history and power in Europe and the Americas, 23-46.

Guan, J. O. D. Q., & Prentice, w. T. (2024). Mapping Social Cohesion 2024 (Mapping Social Cohesio, Issue. S. F. R. Institute. https://scanloninstitute.org.au/mapping-social-cohesion-2024

Guess, A. M., Barberá, P., Munzert, S., & Yang, J. (2021). The Consequences of Online Partisan Media. Proceedings of the National Academy of Sciences, 118(14). https://doi.org/10.1073/pnas.2013464118

Guillaumin, C. (2002). Racism, sexism, power and ideology. Routledge.

Guillory, D. (2020). Combating anti-blackness in the ai community. arXiv preprint arXiv:2006.16879.

Hall, C. (2017). Mukurtu for mātauranga Māori: A case study in Indigenous archiving for reo and tikanga revitalisation. He whare hangarau Māori: Language, culture & technology, 189-199.

Hall, D. M., & Cole, T. J. (2006). What use is the BMI? Archives of disease in childhood, 91(4), 283-286.

Hamilton, I. (2023). Artificial Intelligence In Education: teachers' Opinions On AI In The Classroom. Forbes Advis.

Hancock, N., Smith-Merry, J., Jessup, G., Wayland, S., & Kokany, A. (2018). Understanding the Ups and Downs of Living Well: The Voices of People Experiencing Early Mental Health Recovery. BMC Psychiatry, 18(1). https://doi.org/10.1186/s12888-018-1703-1

Hannah-Jones, N. (2021). The 1619 Project: A New American Origin Story. Random House.

Harari, Y. N. (2014). Sapiens: A brief history of humankind. Random House.

Harris, A., Spark, C., & Watts, M. C. N. C. (2014). Gains and Losses: African Australian Women and Higher Education. Journal of Sociology, 51(2), 370-384. https://doi.org/10.1177/1440783314536792

Harris, C. I. (1993). Whiteness as property. Harvard law review, 1707-1791.

Harris, J. C. (2016). Toward a critical multiracial theory in education. International Journal of Qualitative Studies in Education, 29(6), 795-813.

Harrison, D. A., Kravitz, D. A., Mayer, D. M., Leslie, L. M., & Lev-Arey, D. (2006). Understanding attitudes toward affirmative action programs in employment: Summary and meta-analysis of 35 years of research. Journal of Applied Psychology, 91(5), 1013.

Hartwig, L. D., Jackson, S., Markham, F., & Osborne, N. (2021). Water Colonialism and Indigenous Water Justice in South-Eastern Australia. International Journal of Water Resources Development, 1-34. https://doi.org/10.1080/07900627.2020.1868980

Hashmi, O. U., Chan, N., Vries, C. F. d., Gangi, A., & Lip, G. (2023). Artificial Intelligence in Radiology: Trainees Want More. Clinical Radiology, 78(4), e336-e341. https://doi.org/10.1016/j.crad.2022.12.017

Hati, S. R. H., Balqiah, T. E., Hananto, A., & Yuliati, E. (2021). A Decade of Systematic Literature Review on Airbnb: The Sharing Economy From a Multiple Stakeholder Perspective. Heliyon, 7(10), e08222. https://doi.org/10.1016/j.heliyon.2021.e08222

Hatoss, A. (2012). Where are you from? Identity construction and experiences of ‘othering’in the narratives of Sudanese refugee-background Australians. Discourse & society, 23(1), 47-68.

Hatzinikolakis, J., & Crossman, J. (2020). The Management of Emotional Labor in the Work of Australian University Business School Academics and the Implications for Wellbeing. 1-21. https://doi.org/10.1007/978-3-030-02470-3_43-1

Havea, J. (2017). Postcolonial Voices from Downunder: Indigenous Matters, Confronting Readings. Wipf and Stock Publishers.

Heikkilä, M. (2022). How it feels to be sexually objectified by an AI. MIT Technology Review.

Heleta, S. (2016). Decolonisation of higher education: Dismantling epistemic violence and Eurocentrism in South Africa. Transformation in Higher Education, 1(1), 1-8.

Henwood, F., & Wyatt, S. (2019). Technology and in/Equality, Questioning the Information Society. Digital Culture & Society, 5(1), 183-194. https://doi.org/10.14361/dcs-2019-0111

Hertzmann, A. (2022). Give this AI a few words of description and it produces a stunning image–but is it art. The Conversation (June 2022). http://theconversation. com/give-this-ai-a-few-words-of-descriptionand-it-produces-a-stunning-image-but-is-it-art-184363.

Hirst, J. (2016). Australian history in 7 questions. Black Inc.

Ho, S. (2023). From Development to Dissemination: Social and Ethical Issues With Text-to-Image AI-Generated Art. https://doi.org/10.21428/594757db.acad9d77

Holbrook, M. B., & Schindler, R. M. (2003). Nostalgic bonding: Exploring the role of nostalgia in the consumption experience. Journal of Consumer Behaviour: An International Research Review, 3(2), 107-127.

Holmes, W., Porayska-Pomsta, K., Holstein, K., Sutherland, E., Baker, T. T., Shum, S. B., Santos, O. C., Rodrigo, M. M. T., Cukurova, M., Bittencourt, I. I., & Koedinger, K. R. (2021). Ethics of AI in Education: Towards a Community-Wide Framework. International Journal of Artificial Intelligence in Education, 32(3), 504-526. https://doi.org/10.1007/s40593-021-00239-1

Hopkins, T. (2024). Racial profiling and the social construction of race in Australia. In Handbook of Critical Whiteness: Deconstructing Dominant Discourses Across Disciplines (pp. 465-480). Springer.

Hossfeld, K. J. (1990). Their logic against them: Contradictions in sex, race, and class in Silicon Valley. Women workers and global restructuring, 149, 155-167.

Hovy, D., & Spruit, S. L. (2016). The social impact of natural language processing. Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers),

Huang, X. (2023). A Comparative Study of the Influence of Feminist Development on Young Womens Aesthetics in China and Australia. Communications in Humanities Research, 20(1), 299-306. https://doi.org/10.54254/2753-7064/20/20231388

Hughes, R. T., Zhu, L., & Bednarz, T. (2021). Generative adversarial networks–enabled human–artificial intelligence collaborative applications for creative and design industries: A systematic review of current approaches and trends. Frontiers in artificial intelligence, 4, 604234.

Hundt, A., Agnew, W., Zeng, V., Kacianka, S., & Gombolay, M. (2022). Robots enact malignant stereotypes. 2022 ACM Conference on Fairness, Accountability, and Transparency,

Hwang, W.-C. (2021). Demystifying and addressing internalized racism and oppression among Asian Americans. American Psychologist, 76(4), 596.

Impett, L. (2018). Marxism and new developments in science: Artificial intelligence and deep learning: Technical and political challenges. Theory & Struggle(119), 82-92.

Indicators, S. V. (2024). Snapshot of the Region. https://siliconvalleyindicators.org/about/snapshot/

Infield, T. (2020). Americans who get news mainly on social media are less knowledgeable and less engaged and social media is now among the most common ways people-particularly young adults-get their political news. In.

Intahchomphoo, C., & Gundersen, O. E. (2020). Artificial Intelligence and Race: A Systematic Review. Legal Information Management, 20(2), 74-84. https://doi.org/10.1017/s1472669620000183

Jagota, V. (2016). Why do all the snapchat filters try to make you look white. In: June.

Janvieve Williams Comrie et al, A. M. L., Ph.D., Kwyn Townsend Riley, Jason D. Williamson, Esq. (2022). Antiblackness Colorism| Moving Toward Antibigotry (Center for Antiracist Research, Issue. https://www.bu.edu/antiracism-center/files/2022/06/Moving-Towards-Antibigotry.pdf

Jarvie, D. (2019). Organizational Social Networks and Implications for Inequality in Silicon Valley Tech. 1-22. https://doi.org/10.1007/978-3-319-74078-2_152-1

Jo, H., Ikei, H., & Miyazaki, Y. (2022). Physiological and Psychological Responses of Viewing a Waterfall Image: A Crossover Study. International Journal of Environmental Research and Public Health, 20(1), 565. https://doi.org/10.3390/ijerph20010565

John, M. C., & John, M. C. (2001). Human Computer Interaction In The New Millennium. https://scholar.google.com/scholar?q=Human Computer Interaction In The New Millennium

Johnson, G. D., Thomas, K. D., Harrison, A. K., & Grier, S. A. (2019). Race in the marketplace: Crossing critical boundaries. Springer.

Johnson, G. M. (2021). Algorithmic bias: on the implicit biases of social technology. Synthese, 198(10), 9941-9961. https://doi.org/10.1007/s11229-020-02696-y

Jones, J. M. (2024). Reflections on Psychological Critical Race Theory (PCRT) as a Framework for Disrupting Racism in Pursuit of Social Justice. Journal of Social Issues, 80(2), 801-814. https://doi.org/10.1111/josi.12629

Jones, L. K. (2019). Bringing the receipts: Black feminist theory and intellectual capital in the age of social media. Dissertations available from Proquest.

Joseph, G. G. (1987). Foundations of Eurocentrism in mathematics. Race & Class, 28(3), 13-28.

Julia Angwin, J. L., Surya Mattu and Lauren Kirchner, ProPublica. (2016). Machine Bias

There's software used across the country to predict future criminals. And it's biased against blacks. Pro Publica. Retrieved 25/04/2033 from https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

Kang, H., & Lou, C. (2022). AI agency vs. human agency: understanding human–AI interactions on TikTok and their implications for user engagement. Journal of Computer-Mediated Communication, 27(5), zmac014.

Kanyumba, B., & Lourens, M. (2022). Career Development for Female Academics in Australian and South African Universities. International Journal of Research in Business and Social Science (2147-4478), 11(2), 391-401. https://doi.org/10.20525/ijrbs.v11i2.1576

Keles, P. U., & Aydin, S. (2021). University Students' Perceptions about Artificial Intelligence. Shanlax International Journal of Education, 9, 212-220.

Kelly, C., Karthikesalingam, A., Suleyman, M., Corrado, G. S., & King, D. (2019). Key Challenges for Delivering Clinical Impact With Artificial Intelligence. BMC Medicine, 17(1). https://doi.org/10.1186/s12916-019-1426-2

Kendi, I. X. (2016). Stamped from the beginning: The definitive history of racist ideas in America. Hachette UK.

Khakurel, U., & Rawat, D. B. (2023). On the performance of machine learning fairness in image classification. Artificial Intelligence and Machine Learning for Multi-Domain Operations Applications V, Khandelwal, K., Tonneau, M., Bean, A. M., Kirk, H. R., & Hale, S. A. (2023). Casteist but not racist? quantifying disparities in large language model bias between india and the west. arXiv preprint arXiv:2309.08573.

Kilomba, G. (2021). Plantation memories: episodes of everyday racism. Between the Lines.

Kim, S.-W. (2023). Change in Attitude Toward Artificial Intelligence Through Experiential Learning in Artificial Intelligence Education. International Journal on Advanced Science Engineering and Information Technology, 13(5), 1953-1959. https://doi.org/10.18517/ijaseit.13.5.19039

Kite, M. E., Whitley Jr, B. E., & Wagner, L. S. (2022). Psychology of prejudice and discrimination. Routledge.

Knowles, E. D., Lowery, B. S., Chow, R. M., & Unzueta, M. M. (2014). Deny, Distance, or Dismantle? How White Americans Manage a Privileged Identity. Perspectives on Psychological Science, 9(6), 594-609. https://doi.org/10.1177/1745691614554658

Kokina, J., & Davenport, T. H. (2017). The Emergence of Artificial Intelligence: How Automation Is Changing Auditing. Journal of Emerging Technologies in Accounting, 14(1), 115-122. https://doi.org/10.2308/jeta-51730

Kowal, E. (2012). Stigma and suffering: white anti-racist identities in northern Australia. Postcolonial Studies, 15(1), 5-21.

Kozlova, O., Макарова, М. Н., & Воронина, Л. Г. (2021). Regional Employment Quality in the Context of Russia's Technological Development. E3s Web of Conferences, 301, 03005. https://doi.org/10.1051/e3sconf/202130103005

KP, A. (2024). Contemporary forms of racism, racial discrimination, xenophobia and related intolerance: report of the Special Rapporteur on Contemporary Forms of Racism, Racial Discrimination, Xenophobia and Related Intolerance, Ashwini KP.

Kreps, S., & Kriner, D. (2023). How AI Threatens Democracy. Journal of Democracy, 34(4), 122-131. https://doi.org/10.1353/jod.2023.a907693

Kubin, E., & Sikorski, C. v. (2021). The Role of (Social) Media in Political Polarization: A Systematic Review. Annals of the International Communication Association, 45(3), 188-206. https://doi.org/10.1080/23808985.2021.1976070

Kumar, S., Verma, A. K., & Mirza, A. (2024). Digital Transformation, Artificial Intelligence and Society. In: Springer.

Kumari, M. (2023). Perception of Dental Students in Incorporating Artificial Intelligence Into Dental Education. Journal of Advanced Sciences, 2(1), 41-45. https://doi.org/10.58935/joas.v2i1.27

Kurian, N., Cherian, J. M., Sudharson, N. A., Varghese, K. G., & Wadhwa, S. (2023). AI Is Now Everywhere. BDJ, 234(2), 72-72. https://doi.org/10.1038/s41415-023-5461-1

Kutza, J.-O., Hannemann, N., Hübner, U., & Babitsch, B. (2023). The Representation of Trust in Artificial Intelligence Healthcare Research. https://doi.org/10.3233/shti230409

La Fors, K., & Meissner, F. (2022). Contesting border artificial intelligence: Applying the guidance-ethics approach as a responsible design lens. Data & Policy, 4, e36.

Ladd, A. L. (2014). Gendered Innovations in Orthopaedic Science: Title IX Education: Book Learnin' and Bone Mendin'. Clinical Orthopaedics and Related Research, 472(9), 2586-2589. https://doi.org/10.1007/s11999-014-3738-z

Ladson-Billings, G. (2021). Critical race theory—What it is not! In Handbook of critical race theory in education (pp. 32-43). Routledge.

Lahiri-Dutt, K. (2018). "Academic War" Over Geography? Death of Human Geography at the Australian National University. Antipode, 51(3), 858-877. https://doi.org/10.1111/anti.12496

Lanham, D., & Liu, A. (2019). Not just a typographical change: Why Brookings is capitalizing Black.

Latimore, J., Nolan, D., Simons, M., & Khan, E. (2017). Reassembling the Indigenous Public Sphere. Australasian Journal of Information Systems, 21. https://doi.org/10.3127/ajis.v21i0.1529

Lauri, M. A. (2011). Triangulation of data analysis techniques. Papers on Social Representations, 20(2), 34-31.

Leavy, P. (2022). Research design: Quantitative, qualitative, mixed methods, arts-based, and community-based participatory research approaches. Guilford Publications.

Lee, J. Y., & Lee, J. Y. (2019). Female Transnational Entrepreneurs (FTEs): A Case Study of Korean American Female Entrepreneurs in Silicon Valley. Journal of Entrepreneurship and Innovation in Emerging Economies, 6(1), 67-83. https://doi.org/10.1177/2393957519881925

Lee, N. T. (2018). Detecting Racial Bias in Algorithms and Machine Learning. Journal of Information Communication and Ethics in Society, 16(3), 252-260. https://doi.org/10.1108/jices-06-2018-0056

Lee, S., & Bai, B. (2016). Influence of Popular Culture on Special Interest Tourists' Destination Image. Tourism Management, 52, 161-169. https://doi.org/10.1016/j.tourman.2015.06.019

Lenz, G., & Lawson, J. C. H. (2011). Looking the Part: Television Leads Less Informed Citizens to Vote Based on Candidates' Appearance. American Journal of Political Science, 55(3), 574-589. https://doi.org/10.1111/j.1540-5907.2011.00511.x

Leonardo, Z. (2004). The color of supremacy: Beyond the discourse of 'white privilege'. Educational philosophy and theory, 36(2), 137-152.

Leonardo, Z., & Gamez-Djokic, B. (2019). Sometimes leaving means staying: Race and white teachers' emotional investments. Teachers College Record, 121(13), 1-22.

Leonardo, Z., & Porter, R. K. (2010). Pedagogy of fear: Toward a Fanonian theory of 'safety'in race dialogue. Race Ethnicity and Education, 13(2), 139-157.

Leong, N. (2021). Identity capitalists: The powerful insiders who exploit diversity to maintain inequality. Stanford University Press.

Levin, S., Van Laar, C., & Sidanius, J. (2003). The effects of ingroup and outgroup friendships on ethnic attitudes in college: A longitudinal study. Group Processes & Intergroup Relations, 6(1), 76-92.

Lewis, J. E., Abdilla, A., Arista, N., Baker, K., Benesiinaabandan, S., Brown, M., Cheung, M., Coleman, M., Cordes, A., & Davison, J. (2020). Indigenous protocol and artificial intelligence position paper.

Lewis, S. (2019). The racial bias built into photography. The New York Times, 25, 2019.

Li, S. (2020). The problems with Instagram's most popular beauty filters, from augmentation to eurocentrism. In: July.

Liao, Q. V., & Sundar, S. S. (2022). Designing for Responsible Trust in AI Systems: A Communication Perspective. https://doi.org/10.48550/arxiv.2204.13828

Lin, T.-A., & Chen, P.-H. C. (2022). Artificial Intelligence in a Structurally Unjust Society. Feminist Philosophy Quarterly, 8(3/4). https://doi.org/10.5206/fpq/2022.3/4.14191

Literat, I., Boxman-Shabtai, L., & Kligler-Vilenchik, N. (2022). Protesting the Protest Paradigm: TikTok as a Space for Media Criticism. The International Journal of Press/Politics, 28(2), 362-383. https://doi.org/10.1177/19401612221117481

Lizzio-Wilson, M., Wibisono, S., & Louis, W. (2022). Immigrants as Threat and Opportunity: The Australian Experience. 103-122. https://doi.org/10.1037/0000294-006

Long, C. R., Stewart, M. K., & McElfish, P. A. (2017). Health Research Participants Are Not Receiving Research Results: A Collaborative Solution Is Needed. Trials, 18(1). https://doi.org/10.1186/s13063-017-2200-4

Lopez Andersson, E. A. (2024). Whiteness, power, and decolonial imaginaries in settler Australia: the future hopes of a southern approach. In Handbook of Critical Whiteness: Deconstructing Dominant Discourses Across Disciplines (pp. 907-927). Springer.

López López, L. (2024). 'Apes,''monkeys,'and vibranium: Antiblackness moves in Australian classrooms. Equity & Excellence in Education, 57(1), 62-75.

Lorenz, P., Perset, K., & Berryhill, J. (2023). Initial policy considerations for generative artificial intelligence.

Loury, G. (2022). Why I Don't Capitalize "Black". substack. Retrieved 27.12.2024 from

Loury, G. C., & Loury, G. C. (2009). The anatomy of racial inequality. Harvard University Press.

Loyd, A. B., & Gaither, S. E. (2018). Racial/ethnic socialization for White youth: What we know and future directions. Journal of Applied Developmental Psychology, 59, 54-64.

Lynn, M., & Dixson, A. D. (2013). Handbook of critical race theory in education. Routledge New York, NY.

Lyon, D. (2003). Surveillance as social sorting: Privacy, risk, and digital discrimination. Psychology Press.

MacIntosh, R., Bartunek, J. M., Bhatt, M., & MacLean, D. (2016). I Never Promised You a Rose Garden: When Research Questions Ought to Change. 47-82. https://doi.org/10.1108/s0897-301620160000024003

Maeso, S. R., & Araújo, M. (2015). Eurocentrism, political struggles and the entrenched will-to-ignorance: An introduction. In Eurocentrism, racism and

knowledge: Debates on history and power in Europe and the Americas (pp. 1-22). Springer.
Magubane, Z. (2004). Bringing the empire home: Race, class, and gender in Britain and colonial South Africa. University of Chicago Press.
Maire, Q. (2021). Moral Cosmopolitanism, Civic Action and Ethical Consumption: Social Differences in Young Australians' Global Citizenship. Journal of Sociology, 59(2), 530-551. https://doi.org/10.1177/14407833211053897
Majavu, M. (2020). The 'African gangs' narrative: Associating Blackness with criminality and other anti-Black racist tropes in Australia. African and Black Diaspora: An International Journal, 13(1), 27-39.
Malagon, M. C., Huber, L. P., & Velez, V. N. (2009). Our experiences, our methods: Using grounded theory to inform a critical race theory methodology. Seattle J. Soc. Just., 8, 253.
Mapedzahama, V. (2019). Race matters:(Re) thinking the significance of race and racial inequalities in community development practice in Australia. Social Work & Policy Studies: Social Justice, Practice and Theory, 2(1).
Marger, M. (2003). Race and ethnic relations: American and global perspectives. Wadsworth/Thomson Learning Belmont, CA.
Marginson, S. (2009). Is Australia overdependent on international students? International Higher Education(54).
Markovits, D. (2019). The meritocracy trap. Penguin UK.
Markus, A. (2014). Attitudes to immigration and cultural diversity in Australia. Journal of Sociology, 50(1), 10-22.
Markus, A. (2023). Cultural Diversity: The Australian Social Cohesion Surveys: A. Markus. In Selected Topics in Migration Studies (pp. 301-306). Springer.
Martin, K. (2022). Ethics of data and analytics: Concepts and cases. Auerbach Publications.
Maslej, N., Fattorini, L., Brynjolfsson, E., Etchemendy, J., Ligett, K., Lyons, T., Manyika, J., Ngo, H., Niebles, J. C., & Parli, V. (2023). Artificial intelligence index report 2023. arXiv preprint arXiv:2310.03715.
Matiluko, O. (2020). Decolonising the master's house: How Black Feminist epistemologies can be and are used in decolonial strategy. The Law Teacher, 54(4), 547-561.
Matt Davies , M. B. (2023). Regulating AI in the UK

Strengthening the UK's proposals for the benefit of people and society. https://www.adalovelaceinstitute.org/report/regulating-ai-in-the-uk/

Mattes, K., Spezio, M. L., Kim, H., Todorov, A., Adolphs, R., & Alvarez, R. M. (2010). Predicting Election Outcomes From Positive and Negative Trait Assessments of Candidate Images. Political Psychology, 31(1), 41-58. https://doi.org/10.1111/j.1467-9221.2009.00745.x

Matusitz, J., & Davidson, A. L. (2015). Race of the Future: Mixing of Colors in the United States Leading to One Race? Journal of Human Behavior in the Social Environment, 25(8), 857-870. https://doi.org/10.1080/10911359.2015.1035145

Mbalaka, B. (2023). Epistemically violent biases in artificial intelligence design: the case of DALLE-E 2 and Starry AI. Digital Transformation and Society, 2(4), 376-402.

McAra-Hunter, D. (2024). How AI Hype Impacts the LGBTQ + community. AI and Ethics. https://doi.org/10.1007/s43681-024-00423-8

McDermott, M. (2011). Racial Attitudes in City, Neighborhood, and Situational Contexts. The ANNALS of the American Academy of Political and Social Science, 634(1), 153-173. https://doi.org/10.1177/0002716210388388

McGhee, H. (2022). The sum of us: What racism costs everyone and how we can prosper together. One World.

McKeganey, N., & Bloor, M. (1991). Spotting the invisible man: the influence of male gender on fieldwork relations. British journal of sociology, 195-210.

McLane, P., Mackey, L., Holroyd, B. R., Fitzpatrick, K., Healy, C., Rittenbach, K., Plume, T. B., Bill, L., Bird, A., & Healy, B. (2022). Impacts of racism on First Nations patients' emergency care: results of a thematic analysis of healthcare provider interviews in Alberta, Canada. BMC Health Services Research, 22(1), 1-18.

Md Sumon Gazi, N., Md Rokibul Hasan, N., Nisha Gurung, N., & Anik Mitra, N. (2024). Ethical Considerations in AI-driven Dynamic Pricing in the USA: Balancing Profit Maximization With Consumer Fairness and Transparency. Journal of Economics Finance and Accounting Studies, 6(2), 100-111. https://doi.org/10.32996/jefas.2024.6.2.8

Medina, D. R. (2023). Infographic: AI Research and Development in the U.S., EU and China. Retrieved 04/11/2024 from https://www.bfna.org/digital-world/infographic-ai-research-and-development-in-the-us-eu-and-china-4mk29rb8ig/

Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2021). A Survey on Bias and Fairness in Machine Learning. Acm Computing Surveys, 54(6), 1-35. https://doi.org/10.1145/3457607

Melo, G. (2021). The fox and the grapes. Meanjin, 80(2), 16-17.

Melonio, P. (2020). TikTok's Non-Inclusive Beauty Algorithm & Why We Should Care.

Mentzas, G., Fikardos, M., Lepenioti, K., & Apostolou, D. (2024). Exploring the Landscape of Trustworthy Artificial Intelligence: Status and Challenges. Intelligent Decision Technologies, 18(2), 837-854. https://doi.org/10.3233/idt-240366

Metaxa, D., Wang, K., Landay, J. A., & Hancock, J. (2018). Gender-Inclusive Design. https://doi.org/10.1145/3173574.3174188

Metz, C., & Wakabayashi, D. (2020). Google researcher says she was fired over paper highlighting bias in AI. The New York Times, 3.

Metzl, J. M. (2019). Dying of whiteness: How the politics of racial resentment is killing America's heartland. Hachette UK.

Mhlambi, S. (2020). From rationality to relationality: ubuntu as an ethical and human rights framework for artificial intelligence governance. Carr Center for Human Rights Policy Discussion Paper Series, 9.

Mignolo, W. (2011). The Darker Side of Western Modernity:: global futures, decolonial options. Duke University Press.

Mignolo, W. D., & Walsh, C. E. (2018). On decoloniality: Concepts, analytics, praxis. Duke University Press.

Mills, C. W. (2014). The racial contract. In The racial contract. Cornell University Press.

Mills, K. A., & Godley, A. (2017). Race and racism in digital media: What can critical race theory contribute to research on techno-cultures? In Handbook of writing, literacies, and education in digital cultures. Taylor & Francis.

Mitchell, M. (2021). Bias in the Vision and Language of Artificial Intelligence. In.

Mitchell, M., Wu, S., Zaldivar, A., Barnes, P., Vasserman, L., Hutchinson, B., Spitzer, E., Raji, I. D., & Gebru, T. (2019). Model Cards for Model Reporting. https://doi.org/10.1145/3287560.3287596

Mitchell-Walthour, G. L., & Morrison, M. K. Black Lives Matter in Latin America.

Mittermaier, M., Raza, M. M., & Kvedar, J. C. (2023). Bias in AI-based Models for Medical Applications: Challenges and Mitigation Strategies. NPJ Digital Medicine, 6(1). https://doi.org/10.1038/s41746-023-00858-z

Mogashoa, T. (2014). Understanding critical discourse analysis in qualitative research. International Journal of Humanities Social Sciences and Education, 1(7), 104-113.

Montagu, A. (1997). Man's most dangerous myth: The fallacy of race. Rowman & Littlefield.

Moreton-Robinson, A. (2004). Whiteness, epistemology and Indigenous representation. Whitening race: Essays in social and cultural criticism, 1, 75-88.

Moreton-Robinson, A. (2021). Talkin'up to the white woman: Indigenous women and feminism. U of Minnesota Press.

Morrow, S. L. (2005). Quality and Trustworthiness in Qualitative Research in Counseling Psychology. Journal of Counseling Psychology, 52(2), 250-260. https://doi.org/10.1037/0022-0167.52.2.250

Morse, J. M., Barrett, M., Mayan, M., Olson, K., & Spiers, J. (2002). Verification Strategies for Establishing Reliability and Validity in Qualitative Research. International journal of qualitative methods, 1(2), 13-22. https://doi.org/10.1177/160940690200100202

Mughan, A., & Paxton, P. (2006). Anti-immigrant sentiment, policy preferences and populist party voting in Australia. British Journal of Political Science, 36(2), 341-358.

Muldoon, J., & Wu, B. A. (2023). Artificial intelligence in the colonial matrix of power. Philosophy & Technology, 36(4), 80.

Mullaney, T. S. (2021). Your computer is on fire.

Murár, P., & Kubovics, M. (2023). Using AI to Create Content Designed for Marketing Communications. European Conference on Innovation and Entrepreneurship, 18(1), 660-668. https://doi.org/10.34190/ecie.18.1.1638

Murphy, P. (2024). AI in the Movies. Edinburgh University Press.

Murray, D., & Williams, K. (2023). Exploring Business Students' Views of the Use of Generative AI in Assignment Writing. Ascilite Publications, 167-174. https://doi.org/10.14742/apubs.2023.662

Nadeem, A., Marjanovic, O., & Abedin, B. (2022). Gender bias in AI-based decision-making systems: a systematic literature review. Australasian Journal of Information Systems, 26.

Naden, N. (2017). ABORIGINAL LAND AND AUSTRALIAeS FIRST NATIONS PEOPLES. Postcolonial Voices from Downunder: Indigenous Matters, Confronting Readings, 46.

Nagar, I., & Virk, R. (2017). The Struggle Between the Real and Ideal. SAGE Open, 7(1). https://doi.org/10.1177/2158244017691327

Nambiar, A. (2023). The Experiences Of Colorism For South Asian Women: A Qualitative Study The College of William and Mary].

Nascimento, A. (2020). O quilombismo. Editora Perspectiva SA.

Nasir, B., Toombs, M., Kondalsamy-Chennakesavan, S., Kisely, S., Gill, N., Black, E., Hayman, N., Ranmuthugala, G., Beccaria, G., Ostini, R., & Nicholson, G. C. (2018). Common Mental Disorders Among Indigenous People Living in Regional, Remote and Metropolitan Australia: A Cross-Sectional Study. BMJ Open, 8(6), e020196. https://doi.org/10.1136/bmjopen-2017-020196

Natapu-Ponton, V. (2024). Navigating Whiteness in education: a Pasifika perspective. In Handbook of Critical Whiteness: Deconstructing Dominant Discourses Across Disciplines (pp. 293-308). Springer.

Nealon, K., Balter, P. A., Douglas, R., Fullen, D., Nitsch, P. L., Olanrewaju, A., Soliman, M., & Court, L. E. (2022). Using Failure Mode and Effects Analysis to Evaluate Risk in the Clinical Adoption of Automated Contouring and Treatment Planning Tools. Practical Radiation Oncology, 12(4), e344-e353. https://doi.org/10.1016/j.prro.2022.01.003

Nelson, J. C., Adams, G., & Salter, P. S. (2013). The Marley hypothesis: Denial of racism reflects ignorance of history. Psychological science, 24(2), 213-218.

Newman, E. J., & Schwarz, N. (2024). Misinformed by images: How images influence perceptions of truth and what can be done about it. Current Opinion in Psychology, 56, 101778.

Newton, A., & Dhole, K. (2023). Is AI Art Another Industrial Revolution in the Making? https://doi.org/10.48550/arxiv.2301.05133

Nicolas, G., & Caliskan, A. (2024). A taxonomy of stereotype content in large language models. arXiv preprint arXiv:2408.00162.

Nicoletti, L., Bass, D., & Whiteaker, C. (2023). Generative AI takes stereotypes and bias from bad to worse. Bloomberg [Internet], 8.

Noble, S. U. (2018). Algorithms of oppression. In Algorithms of oppression. New York university press.

Noel, N., Pinder, D., Stewart, S., & Wright, J. (2019). The economic impact of closing the racial wealth gap. McKinsey & Company, 13.

Nundy, S., Kakar, A., Bhutta, Z. A., Nundy, S., Kakar, A., & Bhutta, Z. A. (2022). How to Write the Introduction to a Scientific Paper? How to Practice Academic Medicine and Publish from Developing Countries? A Practical Guide, 193-199.

O'Connell, A., & Greene, C. M. (2017). Not strange but not true: Self-reported interest in a topic increases false memory. Memory, 25(8), 969-977.

O'Donnell, J. (2023). Mapping social cohesion 2023.

O'Keeffe, P. (2024). Australian Identity: Multiculturalism, 'Mainstream Australia' and White Fragility. 27-53. https://doi.org/10.1007/978-981-97-1144-4_2

O'Neil, C. (2017). Weapons of math destruction: How big data increases inequality and threatens democracy. Crown.

Okada, E. S. (2024). A Conversation Analysis of Interactions between Users and a Language Model University of Washington].

Oliveira Andrade de Melo, G. (2022). A little bit of African thinking. Africa is a Country. https://africasacountry.com/2022/05/a-little-bit-of-african-thinking

Oliveira Andrade de Melo, G., & Chapman, D. (2023). Deconstructing Eurocentric Narratives in Non-fiction Literature. In Handbook of Critical Whiteness: Deconstructing Dominant Discourses Across Disciplines (pp. 1-18). Springer.

Oliver-Hoyo, M., & Allen, D. (2006). The Use of Triangulation Methods in Qualitative Educational Research. Journal of college science teaching, 35(4).

Olsen, A., & Mooney-Somers, J. (2017). Except as required by law': Australian researchers' legal rights and obligations regarding participant confidentiality. Australian Human Research Ethics Consultancy Services.

Oluwaseun Augustine Lottu, N., Boma Sonimiteim Jacks, N., Olakunle Abayomi Ajala, N., & Enyinaya Stefano Okafor, N. (2024). Towards a Conceptual Framework for Ethical AI Development in IT Systems. World Journal of Advanced Research and Reviews, 21(3), 408-415. https://doi.org/10.30574/wjarr.2024.21.3.0735

Oyinlade, A. O., & Losen, A. (2014). Extraneous effects of race, gender, and race-gender homo-and heterophily conditions on data quality. SAGE Open, 4(1), 2158244014525418.

Painter, N. I. (2020). Opinion: Why 'White'should be capitalized, too. The Washington Post, 22.

Parikh, R. B., Teeple, S., & Navathe, A. S. (2019). Addressing Bias in Artificial Intelligence in Health Care. Jama, 322(24), 2377. https://doi.org/10.1001/jama.2019.18058

Park, Y. S. (2024). White default: Examining racialized biases behind AI-generated images. Art Education, 77(4), 36-45.

Pathe, T. (2023). How More Women Are Entering the Coding Industry. The Fintech Times. https://thefintechtimes.com/how-more-women-are-entering-the-coding-industry/?utm_source=chatgpt.com

Pawle, F. (2023). What AI thinks the typical Aussie looks like in your home town: Midjourney nails the hipster chic of Melbourne and glamour of the Gold Coast - but why is one major city stuck in the Gold Rush? Dailymail Australia. Retrieved 13/12/2024 from https://www.dailymail.co.uk/news/article-12647765/Stereotypical-Australian-looks-like-town-AI-Midjourney-sydney-melbourne-adelaide.html

Peck, B., & Mummery, J. (2022). Hermeneutic Constructivism: One Ontology for Authentic Understanding. Nursing Inquiry, 30(2). https://doi.org/10.1111/nin.12526

Pedersen, H. (2024). Maori-led Program Shows How AI Can Keep Indigenous Languages Alive. National Indigenous Radio Service. Retrieved 13/12/2024 from https://nirs.org.au/news/indigenous-media/maori-led-program-shows-how-ai-can-keep-indigenous-languages-alive/

Petrick, N., Chen, W., Delfino, J. G., Gallas, B. D., Kang, Y., Krainak, D. M., Sahiner, B., & Samala, R. K. (2023). Regulatory Considerations for Medical Imaging AI/ML Devices in the United States: Concepts and Challenges. Journal of Medical Imaging, 10(05). https://doi.org/10.1117/1.jmi.10.5.051804

Phillips, L. G., Bunda, T., & Quintero, E. P. (2018). Research through, with and as storying. Taylor & Francis.

Phillips, P. J., Jiang, F., Narvekar, A., Ayyad, J., & O'Toole, A. J. (2011). An other-race effect for face recognition algorithms. ACM Transactions on Applied Perception (TAP), 8(2), 1-11.

Phillips, T. (1998). Popular views about Australian identity: research and analysis. Journal of Sociology, 34(3), 281-302.

Phillips-Wren, G., & Adya, M. (2020). Decision making under stress: The role of information overload, time pressure, complexity, and uncertainty. Journal of decision systems, 29(sup1), 213-225.

Picower, B. (2009). The unexamined whiteness of teaching: How white teachers maintain and enact dominant racial ideologies. Race Ethnicity and Education, 12(2), 197-215.

Picower, B. (2021). Reading, writing, and racism: Disrupting whiteness in teacher education and in the classroom. Beacon Press.

Pillay, S. (2020). Thinking the state from Africa: Political theory, eurocentrism and concrete politics. In Decolonisation after Democracy (pp. 32-47). Routledge.

Poba-Nzaou, P., Galani, M., Uwizeyemungu, S., & Ceric, A. (2021). The impacts of artificial intelligence (AI) on jobs: an industry perspective. Strategic HR Review, 20(2), 60-65.

Pocol, A., Istead, L., Siu, S., Mokhtari, S., & Kodeiri, S. (2023). Seeing is no longer believing: A survey on the state of deepfakes, ai-generated humans, and other nonveridical media. Computer Graphics International Conference,

Politi, C., Ciarambino, T., Franconi, F., Feudis, L. D., Mayer, M. C., Martignoni, A., Giovi, I., Tonani, M., Carrubba, S. L., Moretti, A., Martino, C., & Scanelli, G. (2013). Gender Medicine: An Up-Date. Italian Journal of Medicine, 96-106. https://doi.org/10.4081/itjm.2013.96

Poole, R. (2018). The Myth of Progress: 2001–A Space Odyssey. Limiting Outer Space: Astroculture After Apollo, 103-129.

Poredi, N., Sudarsan, M., Solomon, E., Nagothu, D., & Chen, Y. (2024). Generative Adversarial Networks-Based AI-generated Imagery Authentication Using Frequency Domain Analysis. 40. https://doi.org/10.1117/12.3013240

Portrait Photography as a Visual Repository of Vernacular Memory: Interplay of Memory, Identity and Culture. (2023). 10(10), 127-142. https://doi.org/10.61275/isvsej-2023-10-10-09

Pulivarthy, P., & Whig, P. (2025). Bias and Fairness Addressing Discrimination in AI Systems. In Ethical Dimensions of AI Development (pp. 103-126). IGI Global.

Quijano, A. (1992). Colonialidad y modernidad/racionalidad. Perú indígena, 13(29), 11-20.

Raji, I. D., & Buolamwini, J. (2019). Actionable auditing: Investigating the impact of publicly naming biased performance results of commercial ai products. Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society,

Raji, I. D., Gebru, T., Mitchell, M., Buolamwini, J., Lee, J., & Denton, E. (2020). Saving face: Investigating the ethical concerns of facial recognition auditing. Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society,

Ramasubramanian, S. (2015). Using Celebrity News Stories to Effectively Reduce Racial/Ethnic Prejudice. Journal of Social Issues, 71(1), 123-138. https://doi.org/10.1111/josi.12100

Reddy, V. (2016). Fashioning diaspora: Beauty, femininity, and south Asian American culture. Temple University Press.

Reis, J. J. (1995). Slave rebellion in Brazil: the Muslim uprising of 1835 in Bahia. Taylor & Francis.

Rennie, E., Hogan, E., Gregory, R., Crouch, A., Wright, A., & Thomas, J. (2016). Internet on the outstation. The digital divide and remote Aboriginal communities (Vol. 19). Institute of Network Cultures.

Renzaho, A. M. (2023). The lack of race and ethnicity data in Australia—a threat to achieving health equity. International Journal of Environmental Research and Public Health, 20(8), 5530.

Resnikoff, J. (2024). The Problem With AI Is About Power, Not Technology. jacobin. Retrieved 23.12.2024 from

Riccio, P., & Oliver, N. (2022). Racial bias in the beautyverse: Evaluation of augmented-reality beauty filters. European Conference on Computer Vision,

Rieger, A. (2022). Interactive Interventions to Mitigate Cognitive Bias. https://doi.org/10.1145/3503252.3534362

Riverside.fm. (2024). Online Courses: How King's College (Education) records professional courses completely online. Riverside.fm. Retrieved 19.12.2024 from https://riverside.fm/case-studies/kings-college

Robinson, C. (2011). Hair as Race: Why "Good Hair" May Be Bad for Black Females. Howard Journal of Communications, 22(4), 358-376. https://doi.org/10.1080/10646175.2011.617212

Robinson, L., Cotten, S. R., Ono, H., Quan-Haase, A., Mesch, G., Chen, W., Schulz, J., Hale, T. M., & Stern, M. J. (2015). Digital inequalities and why they matter [Article]. Information Communication and Society, 18(5), 569-582. https://doi.org/10.1080/1369118X.2015.1012532

Robinson-Wood, T., Balogun-Mwangi, O., Weber, A., Zeko-Underwood, E., Rawle, S.-A. C., Popat-Jain, A., Matsumoto, A., & Cook, E. (2020). "What Is It Going to Be Like?": A Phenomenological Investigation of Racial, Gendered, and Sexual Microaggressions Among Highly Educated Individuals. Qualitative Psychology, 7(1), 43-58. https://doi.org/10.1037/qup0000113

Rodarte, A. K., & Lukito, J. (2024). Does Social Media Level the Political Field or Reinforce Existing Inequalities? Cartographies of the 2022 Brazilian Election. Political Communication, 1-23.

Roselli, D., Matthews, J., & Talagala, N. (2019). Managing bias in AI. Companion proceedings of the 2019 world wide web conference,

Rosenbaum, J. E. (2022). AI perceptions of gender RMIT University].

Rowe, N. (2023). It's destroyed me completely': Kenyan moderators decry toll of training of AI models. The Guardian, 2.

Ryan-Mosley, T. (2021). Beauty filters are changing the way young girls see themselves. MIT Technology Review. Retrieved 25/04/2023 from https://www.technologyreview.com/2021/04/02/1021635/beauty-filters-young-girls-augmented-reality-social-media/

Sadek, A. W. (2007). Artificial intelligence applications in transportation. Transportation research circular, 1-7.

Salami, M. (2021). Why I don't believe the word 'black'should always have a capital 'b'. The Guardian.

Samuel, A. L. (1959). Machine learning. The Technology Review, 62(1), 42-45.

Sandeen, C. (2008). Boomers, Xers, and Millennials: Who Are They and What Do They Really Want from Continuing Higher Education? Continuing Higher Education Review, 72, 11-31.

Sargsyan, A., Hovsepyan, S., & Muradyan, A. (2024). Ubiquitous and Powerful Artificial Intelligence (AI). In Digitalization of Medicine in Low-and Middle-

Income Countries: Paradigm Changes in Healthcare and Biomedical Research (pp. 255-271). Springer.

Sarr, F. (2020). Afrotopia. U of Minnesota Press.

Saunders, P., & Doyle, A. K. (2022). Cultural Proficiency in First Nations Health Research: A Mixed-Methods, Cross-Cultural Evaluation of a Novel Resource. International Journal of Environmental Research and Public Health, 20(1), 39. https://doi.org/10.3390/ijerph20010039

Schank, R. C. (1987). What is AI, anyway? AI magazine, 8(4), 59-59.

Schlesinger, A., Edwards, W. K., & Grinter, R. E. (2017). Intersectional HCI. https://doi.org/10.1145/3025453.3025766

Schmader, T., Block, K., & Lickel, B. (2015). Social identity threat in response to stereotypic film portrayals: Effects on self-conscious emotion and implicit ingroup attitudes. Journal of Social Issues, 71(1), 54-72.

Schmidt, L., Pehlke, M., & Jansen, M. (2024). AI-Enhanced QOC-Analysis: A Framework for Transparent and Insightful Decision-Making. Working Conference on Virtual Enterprises,

Schostak, J. (2005). Interviewing and representation in qualitative research. McGraw-Hill Education (UK).

Schwartz, R., Vassilev, A., Greene, K., Perine, L., Burt, A., & Hall, P. B. (2022). Towards a Standard for Identifying and Managing Bias in Artificial Intelligence. https://doi.org/10.6028/nist.sp.1270

Sears, D. O., & Henry, P. (2002). Race and politics: The theory of symbolic racism. handbook of political psychology. Japan: Matsysaka university.

Selwyn, N. (2004). Reconsidering political and popular understandings of the digital divide. New Media & Society, 6(3), 341-362.

Sendak, M., Balu, S., & Hernandez, A. F. (2023). Proactive Algorithm Monitoring to Ensure Health Equity. Jama Network Open, 6(12), e2345022. https://doi.org/10.1001/jamanetworkopen.2023.45022

Shankland, S. (2022). In Silicon Valley, Big Tech Gets Bigger but Leaves Many Families Behind. Cnet. Retrieved 27.12.2024 from https://www.cnet.com/culture/in-silicon-valley-big-tech-gets-bigger-but-leaves-many-behind/

Sharples, R., & Blair, K. (2021). Claiming ‘anti-white racism’in Australia: victimhood, identity, and privilege. Journal of Sociology, 57(3), 559-576.

Shelby, R., Rismani, S., Henne, K., Moon, A., Rostamzadeh, N., Nicholas, P., Yilla-Akbari, N. M., Gallegos, J., Smart, A., & Garcia, E. (2023). Sociotechnical harms of algorithmic systems: Scoping a taxonomy for harm reduction. Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society,

Shew, A. (2020). Ableism, technoableism, and future AI. IEEE Technology and Society Magazine, 39(1), 40-85.

Silva, A. (2023). AI images of women from around the world have gone viral. Do they promote colourism and cultural beauty standards? ABC NEWS. Retrieved 22.12.2024 from https://www.abc.net.au/news/2023-08-30/artificial-intelligence-racial-bias-images-women-south-east-asia/102732046

Silva, T. (2022). Racismo algorítmico: inteligência artificial e discriminação nas redes digitais. Edições Sesc SP.

Silva, V., Buzzo, D., Hernández-Ramírez, R., & Ayanoglu, H. (2024). UNDERSTANDING HOW UNDERGRADUATE STUDENTS PERCEIVE BIASES IN AI-GENERATED IMAGES, A RESEARCH-THROUGH-DESIGN PROBE. DS 131: Proceedings of the International Conference on Engineering and Product Design Education (E&PDE 2024),

Singh, J., & Krishna, V. V. (2015). Trends in Brain Drain, Gain and Circulation: Indian Experience of Knowledge Workers. Science Technology and Society, 20(3), 300-321. https://doi.org/10.1177/0971721815597132

Sinkovics, R. R., Penz, E., & Ghauri, P. N. (2008). Enhancing the Trustworthiness of Qualitative Research in International Business. Management International Review, 48(6), 689-714. https://doi.org/10.1007/s11575-008-0103-z

Skinner-Dorkenoo, A. L., Sarmal, A., Andre, C. J., & Rogbeer, K. G. (2021). How microaggressions reinforce and perpetuate systemic racism in the United States. Perspectives on Psychological Science, 16(5), 903-925.

Skitka, L. J., Mosier, K. L., & Burdick, M. (1999). Does automation bias decision-making? International Journal of Human-Computer Studies, 51(5), 991-1006.

Smirnoff, M., Wilets, I., Ragin, D. F., Adams, R., Holohan, J., Rhodes, R., Winkel, G., Ricci, E. M., Clesca, C., & Richardson, L. D. (2018). A Paradigm for Understanding Trust and Mistrust in Medical Research: The Community VOICES Study. Ajob Empirical Bioethics, 9(1), 39-47. https://doi.org/10.1080/23294515.2018.1432718

Smith-Maddox, R., & Solórzano, D. G. (2002). Using Critical Race Theory, Paulo Freire's Problem-Posing Method, and Case Study Research to Confront Race and Racism in Education. Qualitative inquiry, 8(1), 66-84. https://doi.org/10.1177/107780040200800105

Sniderman, P. M., & Tetlock, P. E. (1986). Symbolic Racism: Problems of Motive Attribution in Political Analysis. Journal of Social Issues, 42(2), 129-150. https://doi.org/10.1111/j.1540-4560.1986.tb00229.x

Solorzano, D. G., & Yosso, T. J. (2001). Critical race and LatCrit theory and method: Counter-storytelling. International Journal of Qualitative Studies in Education, 14(4), 471-495.

Solórzano, D. G., & Yosso, T. J. (2002). Critical race methodology: Counter-storytelling as an analytical framework for education research. Qualitative inquiry, 8(1), 23-44.

Sparrow, M. (2022). Riverside Partners With Spotify For Streaming Giant's Video Creation Platform. Forbes. https://www.forbes.com/sites/marksparrow/2022/04/21/riverside-partners-with-spotify-as-the-streaming-giants-official-video-creation-platform/

Sparrow, R. (2020). Robotics has a race problem. Science, Technology, & Human Values, 45(3), 538-560.

Srinivasan, R., & Uchino, K. (2020). Biases in Generative Art -- A Causal Look From the Lens of Art History. https://doi.org/10.48550/arxiv.2010.13266

Sriprakash, A., Rudolph, S., & Gerrard, J. (2022). Learning Whiteness: Education and the settler colonial state. Pluto Press.

Steffan, D., & Venema, N. (2020). New medium, old strategies? Comparing online and traditional campaign posters for German Bundestag elections, 2013–2017. European Journal of Communication, 35(4), 370-388.

Stewart, J., Lu, J., Gahungu, N., Goudie, A., Fegan, P. G., Bennamoun, M., Sprivulis, P., & Dwivedi, G. (2023). Western Australian medical students' attitudes towards artificial intelligence in healthcare. Plos One, 18(8), e0290642.

Stewart, T. L., Latu, I. M., Branscombe, N. R., Phillips, N. L., & Denney, H. T. (2012). White Privilege Awareness and Efficacy to Reduce Racial Inequality Improve White Americans' Attitudes Toward African Americans. Journal of Social Issues, 68(1), 11-27. https://doi.org/10.1111/j.1540-4560.2012.01733.x

Stiglitz, J. E. (2012). THE PRICE OF INEQUALITY: HOW TO-DAY'S DIVIDED SOCIETY ENDANGERS.

Stoyanovich, J., & Howe, B. (2019). Nutritional labels for data and models. A Quarterly bulletin of the Computer Society of the IEEE Technical Committee on Data Engineering, 42(3).

Striepe, M., Thomson, S., & Sefcik, L. (2021). Understanding Academic Integrity Education: Case Studies From Two Australian Universities. Journal of Academic Ethics, 21(1), 1-17. https://doi.org/10.1007/s10805-021-09429-x

Strmic-Pawl, H. V. (2016). Multiracialism and its discontents: A comparative analysis of Asian-White and Black-White multiracials. Lexington Books.

Strully, K. W. (2014). Racially and Ethnically Diverse Schools and Adolescent Romantic Relationships. American Journal of Sociology, 120(3), 750-797. https://doi.org/10.1086/679190

Stypińska, J. (2022). AI Ageism: A Critical Roadmap for Studying Age Discrimination and Exclusion in Digitalized Societies. AI & SOCIETY, 38(2), 665-677. https://doi.org/10.1007/s00146-022-01553-5

Sukhadeve, A. (2021). Artificial Intelligence For Good: How AI Is Helping Humanity. Forbes. Retrieved 08/07/2023 from https://www.forbes.com/sites/forbesbusinesscouncil/2021/02/09/artificial-intelligence-for-good-how-ai-is-helping-humanity/?sh=15c01f613366

Sukkar, A. W., Fareed, M. W., Yahia, M. W., Abdalla, S. B., Ibrahim, I., & Senjab, K. A. K. (2024). Analytical Evaluation of Midjourney Architectural Virtual Lab: Defining Major Current Limits in AI-Generated Representations of Islamic Architectural Heritage. Buildings, 14(3), 786. https://doi.org/10.3390/buildings14030786

Suzor, N. P. (2019). Lawless: The secret rules that govern our digital lives. Cambridge University Press.

Sverdljuk, J., Joranger, T. M. H., & Jackson, E. K. Nordic Whiteness and Migration to the USA.

Szymanski, D. M., & Lewis, J. A. (2015). Race-related stress and racial identity as predictors of African American activism. Journal of Black Psychology, 41(2), 170-191.

Taeihagh, A. (2021). Governance of Artificial Intelligence. Policy and Society, 40(2), 137-157. https://doi.org/10.1080/14494035.2021.1928377

Tan, C. (2013). 'The Tyranny of Appearance': Chinese Australian Identities and the Politics of Difference. In Locating Asian Australian Cultures (pp. 65-82). Routledge.

Tanksley, T. C. (2024). “We’re changing the system with this one”: Black students using critical race algorithmic literacies to subvert and survive AI-mediated racism in school. English Teaching: Practice & Critique, 23(1), 36-56.

Tate, W. F. (1997). Chapter 4: Critical race theory and education: History, theory, and implications. Review of research in education, 22(1), 195-247.

Tavan, G. (2004). The dismantling of the White Australia policy: elite conspiracy or will of the Australian people? Australian Journal of Political Science, 39(1), 109-125.

Tejani, A. S., Ng, Y. S., Xi, Y., & Rayan, J. C. (2024). Understanding and mitigating bias in imaging artificial intelligence. Radiographics, 44(5), e230067.

Testón, A. M., Moriniello, F., Silva, D., Ferrer, J. C., & Muñoz, A. (2023). EXPLORING BACHELOR'S STUDENTS'PERCEPTION OF ARTIFICIAL INTELLIGENCE TEXT-TO-IMAGE GENERATORS IN FINE ARTS AND DESIGN. ICERI2023 Proceedings,

Tiku, N., Schaul, K., & Chen, S. (2023). This is how AI image generators see the world. The Washington Post, 1.

Tolan, S. (2019). Fair and unbiased algorithmic decision making: Current state and future challenges. arXiv preprint arXiv:1901.04730.

Torricelli, M., Martino, M., Baronchelli, A., & Aiello, L. M. (2024). The role of interface design on prompt-mediated creativity in Generative AI. Proceedings of the 16th ACM Web Science Conference,

Trammel, J. M. (2023). Artificial Intelligence for Social Evil: Exploring How AI and Beauty Filters Perpetuate Colorism—Lessons Learned from a Colorism Giant, Brazil. In Black Communication in the Age of Disinformation: DeepFakes and Synthetic Media (pp. 51-71). Springer.

Tran, D. B., Pham, T. D. N., & Nguyen, T. T. (2021). The Influence of Education on Women’s Well-Being: Evidence From Australia. Plos One, 16(3), e0247765. https://doi.org/10.1371/journal.pone.0247765

Tuck, E., & Yang, K. W. (2021). Decolonization is not a metaphor. Tabula Rasa(38), 61-111.

Tukachinsky, R., Mastro, D., & Yarchi, M. (2015). Documenting Portrayals of Race/Ethnicity on Primetime Television Over a 20-Year Span and Their Association With National-Level Racial/Ethnic Attitudes. Journal of Social Issues, 71(1), 17-38. https://doi.org/10.1111/josi.12094

Udah, H. (2023). Anti-black racism and othering: An exploration of the lived experience of black Africans who live in Australia. Social Identities, 29(2), 185-204.

Ulnicane, I., & Aden, A. S. (2023). Power and Politics in Framing Bias in Artificial Intelligence Policy. Review of Policy Research, 40(5), 665-687. https://doi.org/10.1111/ropr.12567

Ungless, E. L., Ross, B., & Lauscher, A. (2023). Stereotypes and smut: The (mis) representation of non-cisgender identities by text-to-image models. arXiv preprint arXiv:2305.17072.

Valenzuela, S., Somma, N. M., Scherman, A., & Arriagada, A. (2016). Social Media in Latin America: Deepening or Bridging Gaps in Protest Participation? Online Information Review, 40(5), 695-711. https://doi.org/10.1108/oir-11-2015-0347

Vallano, J. P., & Compo, N. S. (2011). A Comfortable Witness Is a Good Witness: Rapport-building and Susceptibility to Misinformation in an Investigative Mock-crime Interview. Applied Cognitive Psychology, 25(6), 960-970. https://doi.org/10.1002/acp.1789

VanderLeest, S. (2004). The Built In Bias Of Technology. 2004 Annual Conference,

Vasiliev, D., Bornmalm, L., & Stevens, R. (2024). AI Applications in Environmental, Cultural, and Geological Heritage Conservation. International Congress on Information and Communication Technology,

Vass, G. (2015). Putting critical race theory to work in Australian education research: ‘we are with the garden hose here’. The Australian Educational Researcher, 42(3), 371-394. https://doi.org/10.1007/s13384-014-0160-1

Vears, D. F., & Gillam, L. (2022). Inductive content analysis: A guide for beginning qualitative researchers. Focus on Health Professional Education: A Multi-disciplinary Journal, 23(1), 111-127.

Verma, P. (2022). These robots were trained on AI. They became racist and sexist. Washington Post. Retrieved 06/04 from https://www.washingtonpost.com/technology/2022/07/16/racist-robots-ai/

Vinuesa, R., Azizpour, H., Leite, I., Balaam, M., Dignum, V., Domisch, S., Felländer, A., Langhans, S. D., Tegmark, M., & Fuso Nerini, F. (2020). The role of artificial intelligence in achieving the Sustainable Development Goals. Nature communications, 11(1), 1-10.

Vizcaíno-Verdú, A. Duets, performances and music decolonization on TikTok.

Vyas, D. A., Eisenstein, L. G., & Jones, D. S. (2020). Hidden in plain sight—reconsidering the use of race correction in clinical algorithms. In (Vol. 383, pp. 874-882): Mass Medical Soc.

Wallace, L., Cockrell, D., & Taylor, J. (2010). The University of Newcastle's first cohort of Bachelor of Oral Health students: a social profile. Australian dental journal, 55(4), 436-440.

Wallace, L. R. (2019). The view from somewhere. In The View from Somewhere. University of Chicago Press.

Walsh, T. (2023a). Faking It: Artificial Intelligence in a Human World. Black Inc.

Walsh, T. (2023b). Will AI end privacy? How do we avoid an Orwellian future. AI & SOCIETY, 38(3), 1239-1240.

Walther, A., Logoz, F., & Eggenberger, L. (2024). The gendered nature of AI: Men and masculinities through the lens of ChatGPT and GPT4. Computers in Human Behavior: Artificial Humans, 100076.

Wang, D., Ding, J., Sun, A., Liu, S., Jiang, D., Li, N., & Yu, J.-K. (2023). Artificial Intelligence Suppression as a Strategy to Mitigate Artificial Intelligence Automation Bias. Journal of the American Medical Informatics Association, 30(10), 1684-1692. https://doi.org/10.1093/jamia/ocad118

Wang, F., Feng, Y., & Wang, Z. (2022). Inspiring Desirability or Ensuring Feasibility: Destination Image and Psychological Distance. International Journal of Tourism Research, 24(5), 667-676. https://doi.org/10.1002/jtr.2529

Wang, K., Gou, C., Duan, Y., Lin, Y., Zheng, X., & Wang, F.-Y. (2017). Generative adversarial networks: introduction and outlook. IEEE/CAA Journal of Automatica Sinica, 4(4), 588-598.

Wang, X., Liang, C., & Yin, M. (2023). The Effects of AI Biases and Explanations on Human Decision Fairness: A Case Study of Bidding in Rental Housing Markets. IJCAI, Waring, C. D. (2023). Appearance, parentage, and paradox: The white privilege of bi/multiracial Americans with white ancestry. Sociology of Race and Ethnicity, 9(1), 56-71.

Waring, C. D. (2024). “We’re going to be the new white [people]:” multiracial Americans envision the future. Ethnic and Racial studies, 47(1), 145-166.

Wellner, G., & Rothman, T. (2020). Feminist AI: Can we expect our AI systems to become feminist? Philosophy & Technology, 33(2), 191-205.

Wellner, G. P. (2020). When AI Is Gender-biased. HUMANA. MENTE Journal of Philosophical Studies, 13(37), 127-150.

Werner, J., Arner, T., & White-Taylor, J. (2024). Student Perceptions of Artificial Intelligence. Society for Information Technology & Teacher Education International Conference,

Whittaker, M., Alper, M., Bennett, C. L., Hendren, S., Kaziunas, L., Mills, M., Morris, M. R., Rankin, J., Rogers, E., & Salas, M. (2019). Disability, bias, and AI. AI Now Institute, 8.

Whittaker, M., Crawford, K., Dobbe, R., Fried, G., Kaziunas, E., Mathur, V., West, S. M., Richardson, R., Schultz, J., & Schwartz, O. (2018). AI now report 2018. AI Now Institute at New York University New York.

Whittlestone, J., Nyrup, R., Alexandrova, A., Dihal, K., & Cave, S. (2019). Ethical and societal implications of algorithms, data, and artificial intelligence: a roadmap for research. London: Nuffield Foundation.

Wildman, S. M. (1996). Privilege revealed: How invisible preference undermines America (Vol. 48). NYU press.

Wilkerson, I. (2020). Caste: The International Bestseller. Penguin UK.

Williams-Ceci, S., Jakesch, M., Bhat, A., Kadoma, K., Zalmanson, L., & Naaman, M. (2024). Bias in AI Autocomplete Suggestions Leads to Attitude Shift on Societal Issues. https://doi.org/10.31234/osf.io/mhjn6

Wilson, K., & Caliskan, A. (2024). Gender, race, and intersectional bias in resume screening via language model retrieval. Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society,

Wilton, J. (2017). Identity, racism, and multiculturalism: Chinese-Australian responses. In Migration and identity (pp. 85-100). Routledge.

Wolfe, P. (2006). Settler Colonialism and the Elimination of the Native. Journal of genocide research, 8(4), 387-409.

Wollebæk, D., Karlsen, R., Steen-Johnsen, K., & Enjolras, B. (2019). Anger, Fear, and Echo Chambers: The Emotional Basis for Online Behavior. Social Media + Society, 5(2). https://doi.org/10.1177/2056305119829859

Woo, H., & Heo, N. (2013). A Content Analysis of Qualitative Research in Select ACA Journals (2005–2010). Counseling Outcome Research and Evaluation, 4(1), 13-25. https://doi.org/10.1177/2150137812472195

Wynter, S. (2003). Unsettling the coloniality of being/power/truth/freedom: Towards the human, after man, its overrepresentation—An argument. CR: The new centennial review, 3(3), 257-337.

Xu, Y., Liu, X., Cao, X., Huang, C., Liu, E., Qian, S., Liu, X., Wu, Y., Dong, F., & Qiu, C.-W. (2021). Artificial intelligence: A powerful paradigm for scientific research. The Innovation, 2(4).

Yan, Y., & Bissell, K. (2014). The globalization of beauty: How is ideal beauty influenced by globally published fashion and beauty magazines? Journal of Intercultural Communication Research, 43(3), 194-214.

Yan, Z., Yao, T., Chen, S., Zhao, Y., Fu, X., Zhu, J., Luo, D., Wang, C., Ding, S., & Wu, Y. (2024). Df40: Toward next-generation deepfake detection. arXiv preprint arXiv:2406.13495.

Yang, Q., Steinfeld, A., Rosé, C., & Zimmerman, J. (2020). Re-examining whether, why, and how human-AI interaction is uniquely difficult to design. Proceedings of the 2020 chi conference on human factors in computing systems,

Yared, H., Grové, C., & Chapman, D. (2020). How does race play out in schools? A scoping review and thematic analysis of racial issues in Australian schools. Social Psychology of Education, 23, 1505-1538.

Young, S. G., Hugenberg, K., Bernstein, M. J., & Sacco, D. F. (2012). Perception and motivation in face recognition: A critical review of theories of the cross-race effect. Personality and Social Psychology Review, 16(2), 116-142.

Zainordin, N. M. S. N. M., Basha, N. K., Ann, H. J., & Imm, N. S. (2021). Gender and Age Differences in the Sharing Behavior of Online Video Advertisements: The Case of Selangor, Malaysia. Asian Social Science, 17(12), 49.

Zaugg, I. A., Hossain, A., & Molloy, B. (2022). Digitally-disadvantaged languages. Internet Policy Review, 11(2).

Zhang, Y., & Zavlanos, M. M. (2020). Transfer Reinforcement Learning Under Unobserved Contextual Information. https://doi.org/10.48550/arxiv.2003.04427

Zhao, J., Wang, T., Yatskar, M., Ordonez, V., & Chang, K.-W. (2017). Men also like shopping: Reducing gender bias amplification using corpus-level constraints. arXiv preprint arXiv:1707.09457.

Zhou, W., Deng, Z., Liu, Y., Shen, H., Deng, H. W., & Xiao, H. M. (2022). Global Research Trends of Artificial Intelligence on Histopathological Images: A 20-Year Bibliometric Analysis. International Journal of Environmental Research and Public Health, 19(18), 11597. https://doi.org/10.3390/ijerph191811597

Zion, D. (2023). Class Notes. In.

Zong, R., Zhang, Y., Stinar, F., Shang, L., Zeng, H., Bosch, N., & Wang, D. (2023). A Crowd–AI Collaborative Approach to Address Demographic Bias for Student Performance Prediction in Online Education. Proceedings of the Aaai Conference on Human Computation and Crowdsourcing, 11(1), 198-210. https://doi.org/10.1609/hcomp.v11i1.27560

Zou, J., & Schiebinger, L. (2018a). AI can be sexist and racist—it’s time to make it fair. In: Nature Publishing Group UK London.

Zou, J., & Schiebinger, L. (2018b). Design AI so that it's fair. Nature, 559(7714), 324-326. https://doi.org/https://doi.org/10.1038/d41586-018-05707-8