

**Artificial Impartiality: The Insidious Role of AI in Reinforcing White Supremacy in
Canada**

By
Jessie Thind

A Practicum submitted to the Faculty of Graduate Studies of
The University of Manitoba

In partial fulfillment of the requirements of the degree
MASTER OF HUMAN RIGHTS

Faculty of Law
University of Manitoba
Winnipeg
Copyright © 2023 by Jessie Thind

ABSTRACT

Using a critical race lens, this paper explores how artificial intelligence (AI) oppresses racialized individuals while reinforcing white supremacy. This issue is important given Canada's investment and use of AI technologies in multiple sectors, including law enforcement, immigration, and employment — the use of which has led to systemic discrimination against racialized groups and violations of their fundamental human rights. By exploring how AI is characterized as White, the use of AI in eliminating the fuller attributes of racialized bodies, and the racially homogenous nature of the tech sector, this paper demonstrates that AI advances white supremacist ideology by creating a system where White individuals overwhelmingly control power and privilege. In addition, magnifying the lived experiences of racialized individuals reveals how AI systemically discriminates against racialized groups, further oppressing racialized individuals. This discrimination also infringes on fundamental human rights, as this paper examines in international and domestic law. It is clear that AI governance is required to combat the discriminatory impacts of AI. To do so, this paper advocates for AI governance based on a human rights approach and has provided recommendations to mitigate the discriminatory impact of AI.

Keywords: artificial intelligence, human rights, white supremacy, racial discrimination

TABLE OF CONTENTS

I.	INTRODUCTION	4
II.	ARTIFICIAL INTELLIGENCE AS A WHITE BODY	8
III.	ARTIFICIAL INTELLIGENCE CREATED BY WHITE BODIES	12
IV.	THE SYSTEMIC DISCRIMINATION OF RACIALIZED BODIES	16
	<i>A. FACIAL RECOGNITION SOFTWARE</i>	<i>17</i>
	<i>B. RISK ASSESSMENT TOOLS</i>	<i>20</i>
V.	A MOVE TOWARDS RESPONSIBLE ARTIFICIAL INTELLIGENCE	22
	<i>A. LEX LATA</i>	<i>22</i>
	<i>B. LEX FERENDA</i>	<i>27</i>
VI.	CONCLUSION	29
VII.	BIBLIOGRAPHY	31

I. INTRODUCTION

Since the turn of the millennium, Canada has become an innovation hub for disruptive technologies such as artificial intelligence (AI).¹ Not only has the Canadian State created the world's first national AI strategy, but it also produces the most AI patents per million people in the world among G7 nations and China and continues to maintain the highest density of AI start-ups in the world.² Canada's interest in investing in and expanding upon the use of AI is not just about bolstering national pride; rather, it is a lucrative investment substantiated by the notion that AI drives economic growth, productivity, and prosperity.³ In fact, since 2010, the private Canadian AI sector has drawn \$3 billion in new investments and created 50,000 jobs.⁴ Moreover, private companies and governments alike seize on the promise of AI to provide impartial, data-driven decision-making methods at scale.⁵ However, in the last decade, the use of AI has raised socio-ethical concerns.⁶

The current use of AI presents a multitude of issues regarding the intersectionality of race and technology. For example, police departments in Vancouver, Edmonton and Saskatoon have implemented predictive policing — automated decision-making that predicts when and who will commit a crime.⁷ A practice which has led to the creation of a list that features a

¹ Chris Reynolds, "Canada needs laws to tackle life-and-death problems involving AI, experts warn", *CTV News* (19 May 2019), online: *CTV News* <<https://www.ctvnews.ca/sci-tech/canada-needs-laws-to-tackle-life-and-death-problems-involving-ai-experts-warn-1.4428959?cache=yes>>; Michael Pizzi, Mila Romanoff & Tim Engelhardt, "AI for humanitarian action: Human rights and ethics" (2020) 102:913 *Int Rev Red Cross* 145 at 150.

² Global Advantage Consulting Group, "Canada's AI Ecosystem: Government Investment Propels Private Sector Growth" (2021) at 2, online: *University of Toronto* <https://gro.utoronto.ca/wp-content/uploads/2022/04/Canadas_AI_Ecosystem-_Government_Investment_Propels_Private_Sector_Growth.pdf>.

³ *Ibid* at 12; Olga Akselrod, "How Artificial Intelligence Can Deepen Racial and Economic Inequities" (13 July 2021) at para 1, online: *American Civil Liberties Union* <<https://www.aclu.org/news/privacy-technology/how-artificial-intelligence-can-deepen-racial-and-economic-inequities>>.

⁴ Global Advantage Consulting Group, *supra* note 2 at 2.

⁵ Timnit Gebru, "Race and Gender" in Markus D Dubber, Frank Pasquale & Sunit Das, eds, *The Oxford Handbook of Ethics of AI* (Oxford University Press, 2020) 251 at 252; Evgeni Aizenberg & Jeroen van den Hoven, "Designing for human rights in AI" (2020) 7:2 *Big Data & Society* 1 at 1.

⁶ Aizenberg, *supra* note 5 at 1.

⁷ Reynolds, *supra* note 1 at para 7.

disproportionally high number of Black males with no prior arrest record as potential perpetrators.⁸ It is also used to eliminate characteristics of racialized groups,⁹ such as the accents possessed by racialized call center operators, in order to sound ‘accent-neutral’ and thus more palatable to prospective callers.¹⁰ These examples of racial discrimination perpetuated through AI have been highlighted by both the popular press and academic literature.¹¹ Yet, while racist AI is increasingly well documented, discourse on how AI synchronously advances white supremacy remains scarce. White supremacist ideology should not be conceptualized in narrow terms as only a feature of alt-right movements.¹² Rather, white supremacy is “a political, economic, and cultural system in which Whites overwhelmingly control power and material resources, conscious and unconscious ideas of White superiority and entitlement.”¹³ This paper aims to contribute to the body of literature on race and technology by examining how the tech sector reinforces a racial hierarchy of power and privilege.

Examining how AI advances white supremacist ideology requires the use of an appropriate lens to understand that the ideas of white superiority and entitlement are deeply embedded in the tech sector.¹⁴ Critical race theory (CRT) provides a fitting lens for this examination. CRT is a cross-disciplinary analysis and framework used to explore and explain the

⁸ *Ibid* at para 11.

⁹ For the context of this paper, racialized groups will refer to individuals who are Black, Indigenous or persons of colour.

¹⁰ Chance Townsend, “Startup develops AI that makes call center employees sound like white, American robots” *Mashable* (27 August 2022), online: *Mashable* <<https://mashable.com/article/ai-startup-makes-call-center-foreigners-sound-white>>; Taylor C. Moran, “Racial technological bias and the white, feminine voice of AI VAs” (2021) 18:1 *Communication and Critical/Cultural Studies* 19 at 19; Stephen Cave & Kanta Dihal, “The Whiteness of AI” (2020) 33:4 *Philos Technol* 685 at 685.

¹¹ See Reynolds, *supra* note 1 at para 2.

¹² Stefan Lawrence & Kevin Hylton, “Critical Race Theory, Methodology, and Semiotics: The Analytical Utility of a ‘Race’ Conscious Approach for Visual Qualitative Research” (2022) 22:3 *Cultural Studies ↔ Critical Methodologies* 255 at 260.

¹³ *Ibid*.

¹⁴ Gideon Christian, “Artificial Intelligence, Algorithmic Racism and the Canadian Criminal Justice System” *SLAW* (26 October 2020), online: *Slaw* <http://www.slaw.ca/2020/10/26/artificial-intelligence-algorithmic-racism-and-the-canadian-criminal-justice-system/>.

pervasiveness of racism across society and within social institutions that may otherwise present themselves as objective.¹⁵ By way of Richard Delgado and Jean Stefancic, early scholars of CRT, this framework consists of five central themes, three of which are focalized on and pertinent to this paper.¹⁶ This includes: (1) the understanding that racism is pervasive, not aberrational; (2) the idea that white supremacy operates to benefit White individuals while oppressing racialized individuals; and (3) the belief that racialized voices possess knowledge through their lived experiences.¹⁷ CRT scholars are increasingly utilizing research and data to expose the implicit racism embedded in laws, societal intuitions, and social practices that result in biased outcomes against racialized individuals.¹⁸ Using CRT to examine AI and how it operates will reveal how AI is a scourge to racialized individuals while reinforcing a racial hierarchy of power and privilege that benefits White individuals.

In this paper, I argue that the current use of AI oppresses racialized individuals while reinforcing the notion of white supremacy. To advance this argument, I draw on official and secondary sources. With regard to secondary sources, I rely on peer-reviewed journal articles and books available through online databases such as *SAGE journals*, *Springer*, and *HeinOnline* to situate the characterization of AI as White while erasing racialized bodies from the tech sector (in which AI is created). I then look to official sources, such as international law, Canadian case law, and UN documents to discuss how this discrimination infringes on fundamental human rights. In addition, I rely on online sources and newspaper articles, including those published in *Mashable*, *CTV News*, and *CNN*. These online sources and articles provide the narratives of

¹⁵ Lawrence, *supra* note 12 at 256.

¹⁶ *Ibid*; Sonja Martin Poole et al, “Operationalizing Critical Race Theory in the Marketplace” (2021) 40:2 Journal of Public Policy & Marketing 126 at 131.

¹⁷ Lawrence, *supra* note 12 at 256.

¹⁸ See Amar Khoday, “Black Voices Matter Too: Counter-Narrating *Smithers v The Queen*” (2021) 58:3 Osgoode Hall LJ 567; Christian, *supra* note 14 at para 6.

racialized individuals and their experiences with racial discrimination and AI. These online sources also reveal the homogenous nature of the tech sector and provide research on the issue of algorithmic bias by examining how the biases of tech workers (whether unconscious or otherwise) are coded into the algorithms that create AI.

This paper is divided into four main parts and a conclusion. Part I of this paper will explore the characterization of AI as being White while eliminating the fuller attributes of racialized bodies. Part II will examine how White males predominately develop the algorithms and subsequent AI and how this exacerbates algorithmic bias. Part III will explore the lived experience of racialized individuals through two case studies to magnify how AI systemically discriminates against racialized groups. Lastly, part IV of this paper will discuss how this institutionalized discrimination infringes on fundamental human rights before shifting to a *lex ferenda* analysis examining the need for an AI regulatory framework. In doing so, this paper seeks to illustrate what AI governance in Canada should entail to provide effective human rights protection and advance responsible AI globally.

Finally, it is essential to recognize that experts and researchers have provided many definitions of AI.¹⁹ While there is no single universally accepted definition of AI, there is a consensus that AI is a broad term that refers to systems or machines that seek to mimic aspects of human intelligence to perform tasks and automate decisions.²⁰ Due to its extensive nature, AI is split into subfields such as Robotics and Machine Learning (ML).²¹ The latter refers to a component of AI systems that can detect patterns, learn from those patterns, and apply that new learning to different situations to solve problems and make decisions (e.g., Spotify's

¹⁹ Sheshadri Chatterjee, Sreenivasulu NS & Zahid Hussain, "Evolution of artificial intelligence and its impact on human rights: from sociolegal perspective" (2021) 64:2 International Journal of Law and Management 184 at 186.

²⁰ *Ibid*; Pizzi, *supra* note 1 at 150.

²¹ Chatterjee, *supra* note 19 at 186.

personalized recommendations via their Discover Weekly playlist).²² However, it is essential to note that the ML model may still require human intervention to correct any errors in its decision-making. Many innovations in AI boil down to ML and Deep Learning (DL).²³ DL is a more potent subset of ML that structures algorithms — a set of rules — as layers to create an artificial neural network.²⁴ Modelled after neurons in the human brain, DL is used to learn and make intelligent decisions on its own without any human input (e.g., self-driving cars).²⁵ For this paper, AI will primarily refer to the ML and DL models unless otherwise specified.

II. ARTIFICIAL INTELLIGENCE AS A WHITE BODY

Over the past several decades, both embodied and disembodied depictions of AI have been anthropomorphized.²⁶ For example, search engine image results for ‘artificial intelligence’ yield a preponderance of robots with human characteristics — including a head with facial features that replicate the human head — also referred to as humanoid robots.²⁷ Films and television also anthropomorphize AI, as do virtual assistants available through smartphones and home devices.²⁸ These depictions of AI expand from possessing physical human characteristics to having human traits, such as names, human-like voices, and personalities.²⁹ The intent of designing AI systems and machines to possess human characteristics and traits is to aid in

²² Pizzi, *supra* note 1 at 150.

²³ *Ibid.*

²⁴ *Ibid.*

²⁵ *Ibid.*

²⁶ Cave, *supra* note 10 at 688; Maria Noriega, “The application of artificial intelligence in police interrogations: An analysis addressing the proposed effect AI has on racial and gender bias, cooperation, and false confessions” (2020) 117 *Futures* 102510 at 102515.

²⁷ Cave, *supra* note 10 at 686.

²⁸ *Ibid.* at 690-693; Moran, *supra* note 10 at 20.

²⁹ Cave, *supra* note 10 at 690-693; Moran, *supra* note 10 at 20.

human-machine interaction.³⁰ In doing so, human interaction with AI is thought to be more accessible, thus, increasing the end user's interaction with and use of AI systems and machines.³¹

In addition to possessing human characteristics, research has illustrated that society tends to racialize these AI systems and machines.³² For example, in a 2018 study led by Christoph Bartneck, researchers undertook a range of experiments consisting of sharing multiple images of a humanoid robot whose physical colour was adjusted to match the skin tone of a stock image of a Black individual and a White individual.³³ The study then asked participants to define the robot's racial identity.³⁴ It provided several options, including "does not apply".³⁵ Approximately 53% of participants perceived the colour-adjusted black robot as racially Black, and roughly 70% identified the colour-adjusted white robot as racially White.³⁶ Based on these results, the researchers concluded that individuals are strongly inclined to assign racial identity to AI machines and systems.³⁷ These results are not surprising, as racial identity is an integral part of being human.³⁸ Moreover, numerous studies have shown that a person's connection and trust in an AI system is higher when the AI shares the same racial identity as the individual.³⁹ Based on these studies, one would deduce that any AI designed to possess human characteristics comprises multiple versions of different racial identities based on the numerous racial identities that exist

³⁰ Cave, *supra* note 10 at 688; Noriega, *supra* note 26 at 102515.

³¹ Cave, *supra* note 10 at 688; Noriega, *supra* note 26 at 102515.

³² Christoph Bartneck et al, *Robots and Racism* (New York, NY, USA: Association for Computing Machinery, 2018) at 196-197; Cave, *supra* note 10 at 688, online <<http://hdl.handle.net/10092/15024>>.

³³ Bartneck, *supra* note 32 at 196-197; Cave, *supra* note 10 at 688.

³⁴ Bartneck, *supra* note 32 at 198; Cave, *supra* note 10 at 688.

³⁵ Bartneck, *supra* note 32 at 198; Cave, *supra* note 10 at 688.

³⁶ Bartneck, *supra* note 32 at 198; Cave, *supra* note 10 at 688.

³⁷ Bartneck, *supra* note 32 at 201.

³⁸ Cave, *supra* note 10 at 688.

³⁹ *Ibid* at 689; Noriega, *supra* note 26 at 102525.

globally. However, AI is consistently portrayed as White — as evidenced through search engine image searches, portrayals in film and television, and characterizations of virtual assistants.⁴⁰

As previously mentioned, using image search for the term ‘artificial intelligence’ in any search engine will yield stock images of humanoid robots.⁴¹ These stock images of humanoids are overwhelmingly white in colour and White in racial identity.⁴² In fact, the more realistic these humanoids are, the more Caucasian their features become.⁴³ This portrayal of AI as White is expanded upon in films and television. For example, science fiction films produced in the past several decades imagine AI as a White body — this includes *The Terminator*, the replicants in *Blade Runner*, and Ava in *Ex Machina* — all played by White actors and appearing as physically White on screen.⁴⁴ Furthermore, disembodied AI in films is also portrayed as White.⁴⁵ This includes Sam, the operating system in *Her*; HAL 9000, the AI portrayed in *2001: A Space Odyssey*; and J.A.R.V.I.S., the AI designed by Tony Stark in the Marvel Cinematic Universe (MCU) film franchise — all voiced by White actors.⁴⁶ Lastly, this portrayal of AI as White is seen in virtual assistants such as Apple’s Siri and Amazon’s Alexa.⁴⁷ While these virtual assistants do not exhibit any visual racial cues regarding their racial identity, they possess sociolinguistic markers such as diction, vocal speed, articulation and elocution that signify that these voices come from a predominately Western mode of speech.⁴⁸ One that uses ‘proper English,’ is considered accent-free, and is commonly associated with Whiteness.⁴⁹ The

⁴⁰ Cave, *supra* note 10 at 687.

⁴¹ *Ibid* at 691.

⁴² *Ibid*.

⁴³ *Ibid*.

⁴⁴ Cave, *supra* note 10 at 693; Moran, *supra* note 10 at 20.

⁴⁵ Cave, *supra* note 10 at 693.

⁴⁶ *Ibid*; Moran, *supra* note 10 at 19; I am grateful to Dr. Amar Khoday for drawing my attention to the example of J.A.R.V.I.S. from the MCU and sharing his insights and suggestions throughout this paper.

⁴⁷ Cave, *supra* note 10 at 690; Moran, *supra* note 10 at 20.

⁴⁸ Cave, *supra* note 10 at 690; Moran, *supra* note 10 at 21-23.

⁴⁹ Cave, *supra* note 10 at 690; Moran, *supra* note 10 at 21-23; While a more in-depth discussion of the notion of the ‘White voice’ is beyond the scope of this paper, it is essential to note that the topic of the ‘White voice’ is complex.

preference to communicate with a White voice is so widespread that AI that erases specific racialized characteristics — such as accents — to sound Whiter has begun to emerge.⁵⁰

In 2022, a tech start-up based in California introduced AI to make racialized call center operators sound ‘accent-neutral’ and thus more palatable to White callers.⁵¹ The AI offers real-time accent translation by removing the operator’s accent and replacing the voice with an ‘American English’ voice.⁵² The company argues that its service permits call center operators to be more easily understood, allowing racialized call center employees to “take back the power of their own voice.”⁵³ The company has already received over \$30 million in venture capital funding.⁵⁴

The predominant depiction of AI as White further entrenches the notion that Whiteness is normative.⁵⁵ This is a common notion expressed in CRT scholarship.⁵⁶ Treating Whiteness as a norm is problematic as it leads to social, political, and cultural outputs that are designed only for White individuals.⁵⁷ It also leads to the erasure of racialized voices. More substantively, the depiction of Whiteness as normative seeks to normalize the dehumanization and erasure of racialized individuals and their characteristics from a global population that primarily consists of racialized groups in favour of a White utopia.⁵⁸ By applying a CRT lens, it is evident that the

Indeed, there is no single type of White voice but rather a multiplicity of such voices, including, among other things, regional variations and educational backgrounds. Some may be viewed as ‘less desirable’ than the ‘mainstream,’ accent-free White voices referred to in this context. There is also the lived reality of many racialized individuals who code-switch, whether subconsciously or not, to assimilate to the Western majority’s vernacular in order to be considered more affable.

⁵⁰ Townsend, *supra* note 10 at para 1.

⁵¹ *Ibid.*

⁵² *Ibid* at para 1-4.

⁵³ *Ibid* at para 4.

⁵⁴ *Ibid* at para 2.

⁵⁵ Moran, *supra* note 10 at 19.

⁵⁶ Khoday, *supra* note 18 at 578.

⁵⁷ Rashida Richardson, “Racial Segregation and the Data-Driven Society: How Our Failure to Reckon with Root Causes Perpetuates Separate and Unequal Realities” (2022) 36:3 Berkeley Tech LJ 1051 at 1066.

⁵⁸ Cave, *supra* note 10 at 698.

erasure of racialized bodies and their characteristics from AI in favour of Whiteness further reinforces the idea that societal advancement and intelligence are White — boosting white supremacist ideology.⁵⁹ As previously examined under the CRT framework, white supremacy operates to profess the superiority of White individuals in order to advance benefits of power and privilege to White individuals while oppressing racialized individuals.⁶⁰ This ideology is not new. The supremacy of Whiteness is also observed in the false rationale that White individuals are mentally superior to racialized individuals, which served to justify colonization across many parts of the world, including Canada.⁶¹ Given that the association of intelligence and Whiteness has been purported for centuries, it is unsurprising that beliefs about racial identity, intelligence, and power persist widely today.⁶²

By examining the depiction of AI as White, via search engine image searches, film and television portrayals, and characteristics of virtual assistants, it is evident that the current representation of AI oppresses racialized individuals while purporting the superiority of White individuals. By applying a CRT lens, this depiction further reinforces an invisible racial hierarchy of power and privilege, where White individuals are considered more intelligent and better equipped to advance society. These biased representations of AI have also influenced the tech sector by shaping who is considered suitable to work in the field — leading to a racially homogenous workforce.⁶³

III. ARTIFICIAL INTELLIGENCE CREATED BY WHITE BODIES

⁵⁹ *Ibid* at 686; Moran, *supra* note 10 at 19.

⁶⁰ Lawrence, *supra* note 12 at 256.

⁶¹ Cave, *supra* note 10 at 696.

⁶² *Ibid*; Poole, *supra* note 16 at 127.

⁶³ Cave, *supra* note 10 at 699.

The public perception of technology is that it is objective and neutral — a perception that continues to be salient for computer scientists, developers, and the overall creators and producers of technological systems and machines.⁶⁴ But technology is never neutral — it reflects the ideals and values of those who create it and is shaped by the same inequalities deeply embedded in society.⁶⁵ In the case of the tech industry, the ideals and values of its creators come from predominately White individuals.⁶⁶ Like many industries, the tech industry is notoriously known for having racial homogeneity, and Canada’s tech sector is no different.⁶⁷ Approximately 70% of Canada’s tech workforce is White.⁶⁸ Furthermore, 80% of Canada’s tech workforce identifies as male.⁶⁹ This homogeneity of White males in the tech sector is further concentrated when looking at leadership positions.⁷⁰ The lack of diversity and concentration of primarily White males in leadership positions has led to many criticizing the tech sector as having a “White guy problem.”⁷¹

The lack of diversity is an issue for several reasons. Since White men dominate the industry, this group is most notably associated with the algorithmic design and deployment of AI.⁷² The lack of diversity associated with those primarily responsible for creating AI is problematic as it hinders the ability of these largely homogenous teams to foresee the myriad ways in which their AI may impact communities that differ from their racial identity and

⁶⁴ E. Tendayi Achiume, Racial discrimination and emerging digital technologies: a human rights analysis, HRC 45th Sess, UN Doc A/HRC/44/57 at para 12 (2020), online: <<https://documents-dds-ny.un.org/doc/UNDOC/GEN/G20/151/06/PDF/G2015106.pdf?OpenElement>>.

⁶⁵ *Ibid.*

⁶⁶ *Ibid* at para 17.

⁶⁷ Akselrod, *supra* note 3 at para 2; Richardson, *supra* note 56 at 1052.

⁶⁸ Viet Vu, Creig Lamb & Asher Zafar, “Who Are Canada’s Tech Workers?” (Toronto, ON, CA: Brookfield Institute for Innovation Entrepreneurship, 2019) at 27.

⁶⁹ *Ibid* at 18.

⁷⁰ Richardson, *supra* note 57 at 1052.

⁷¹ *Ibid* at 1053.

⁷² *Ibid.*

gender.⁷³ Rather, these White men come from privileged backgrounds, often unaware of how racial discrimination and bias influence, intersect and shape data sources, data mining, and human biases in the AI development process.⁷⁴ A part of this lack of understanding stems from the fact that computer science is not an interdisciplinary field.⁷⁵ In addition, ethics courses have only recently begun to be included in some computer science curriculums at post-secondary institutions in Canada.⁷⁶

Furthermore, research has shown that these primarily White developers evoke colour-blind ideologies.⁷⁷ Through a critical race lens, colour-blindness is an attitude of not seeing racial identity or assuming people in modern Western society are no longer disadvantaged due to their racial identity.⁷⁸ It is an ideology that perpetuates White hegemony while further marginalizing racialized individuals.⁷⁹ Central to this ideology is the inappropriateness of “seeing race.”⁸⁰ However, for racialized individuals, race and Whiteness is never invisible.⁸¹ This lack of understanding by the primarily White male developers in this sector is problematic when they must make assumptions about what data is valuable and relevant and decide when the AI is correct in its predictions — assumptions which lead to algorithmic bias.⁸² Algorithmic bias is a bias that the developer introduces, sometimes unconsciously, during the development of AI.⁸³ A

⁷³ Morgan Livingston, “Preventing Racial Bias in Federal AI” (2020) 16:02 JSPG at 2, online: <https://www.sciencepolicyjournal.org/article_1038126_jspg160205.html>.

⁷⁴ Richardson, *supra* note 57 at 1071.

⁷⁵ Mike Zajko, “Conservative AI and social inequality: conceptualizing alternatives to bias through social theory” (2021) 36:3 AI & Soc 1047 at 1048, online: <<https://doi.org/10.1007/s00146-021-01153-9>>.

⁷⁶ *Ibid.*

⁷⁷ Cave, *supra* note 10 at 695; Poole, *supra* note 16 at 131; Richardson, *supra* note 57 at 1068.

⁷⁸ Cave, *supra* note 10 at 695; Poole, *supra* note 16 at 131; Richardson, *supra* note 57 at 1068; Khoday, *supra* note 18 at 577.

⁷⁹ Cave, *supra* note 10 at 695; Poole, *supra* note 16 at 131; Richardson, *supra* note 57 at 1068; Khoday, *supra* note 18 at 577.

⁸⁰ Cave, *supra* note 10 at 695.

⁸¹ *Ibid.*

⁸² Bradley Henderson, Colleen M. Flood, and Teresa Scassa, “Artificial Intelligence in Canadian Healthcare: Will the Law Protect Us from Algorithmic Bias Resulting in Discrimination?” (2022) 19:2 CJLT 475.

⁸³ *Ibid.*

notable example of algorithmic bias that can arise when creating AI stems from developers inappropriately selecting proxy variables linked to socioeconomic characteristics.⁸⁴ Proxy variables are those that an AI uses as a substitute to predict its ‘goal.’⁸⁵ Under a colour-blind approach, using proxy variables linked to socioeconomic characteristics ignores how racialized individuals are systemically disadvantaged in areas like housing and income. Additionally, categorizing the discriminatory design of AI as a purely technical problem for developers to solve by producing bias-free data ignores that technology is a product of society, including its values and inequalities.⁸⁶ Using a CRT lens, it is essential to note that these racial inequalities are systemic, extensive, and deeply embedded across all societal institutions, as racism is pervasive, rather than aberrational.⁸⁷ Therefore, even without creating algorithmic rules that explicitly express bias, AI algorithms reproduce the discriminatory bias entrenched in the data it learns from, as these inequalities and biases are deeply embedded in society.⁸⁸

AI relies on training data to learn how to operate.⁸⁹ Training data comprises of historical data or relevant examples that teach the AI how to perform particular tasks.⁹⁰ However, since training data consists of historical data, it reflects the discrimination, biases, and inaccurate profiling deeply embedded in society’s social structures.⁹¹ By being unaware that institutionalized racism is embedded in all areas of society, such as immigration, policing, housing, and health care, developers exacerbate these inequalities.⁹² For example, some predictive policing algorithms use data from past locations of police responses to predict future

⁸⁴ *Ibid*; Pizzi, *supra* note 1 at 154.

⁸⁵ Henderson, *supra* note 82 at 481.

⁸⁶ Achiume, *supra* note 64 at para 13.

⁸⁷ Lawrence, *supra* note 12 at 256; Khoday, *supra* note 18 at 577-578.

⁸⁸ Achiume, *supra* note 64 at para 7.

⁸⁹ Richardson, *supra* note 57 at 1071.

⁹⁰ *Ibid*.

⁹¹ *Ibid*; Achiume, *supra* note 64 at para 8.

⁹² Richardson, *supra* note 57 at 1071.

locations of crime.⁹³ However, this data overrepresents racialized communities.⁹⁴ The outcome is AI that flags already over-policed areas of racialized communities for additional police monitoring, further entrenching racially discriminatory policing.⁹⁵

Also, there is often a disconnect between the design and deployment of AI — an issue often exacerbated when a racially homogenous workforce creates AI in a Western society.⁹⁶ For example, an AI system or machine designed in the Global North but deployed in the Global South may fail to account for that country's unique social, political, and cultural differences.⁹⁷ The developer may be unaware if certain stigmatized groups are underrepresented in their training data and fail to account for that bias in the data.⁹⁸

The racial homogeneity in the tech sector has led to a workforce unaware of the racial discrimination embedded in AI and its insidious impacts. It is also important to note that the tech sector is lucrative.⁹⁹ Workers in this sector earn \$70-\$90K per year, on average.¹⁰⁰ In comparison, workers in the non-tech sector earn, on average, \$45K per year.¹⁰¹ Through a critical race lens, it is evident that having an industry characterized by high pay and the socioeconomic advancement of society but having that industry consist of primarily White individuals reinforces a racial hierarchy of power and privilege. This privilege can be further observed by examining how AI has discriminated against racialized individuals.

IV. THE SYSTEMIC DISCRIMINATION OF RACIALIZED BODIES

⁹³ Livingston, *supra* note 73 at 2.

⁹⁴ *Ibid.*

⁹⁵ *Ibid.*

⁹⁶ Pizzi, *supra* note 1 at 155.

⁹⁷ *Ibid.*

⁹⁸ *Ibid.*

⁹⁹ Vu, *supra* note 68 at 1.

¹⁰⁰ *Ibid* at 9.

¹⁰¹ *Ibid.*

Before examining how AI discriminates against racialized individuals, it is important to understand that AI aims to bring societal advantages.¹⁰² In fact, AI solves pressing problems in contemporary society.¹⁰³ For example, in the Canadian healthcare system, AI has been used to reduce hospital waiting times, increase the speed of diagnoses and provide specialized care through robotic surgery in rural and remote communities.¹⁰⁴ Moreover, AI is used to predict extreme weather events, allowing for arrangements to be made to prepare for the impact of natural disasters.¹⁰⁵ However, believing that technology and AI can solve all societal problems can cloud AI's insidious impact on racialized individuals.¹⁰⁶ Using a critical race lens in advancing the lived experiences of racialized individuals, this section examines AI's discriminatory impact in two key areas: facial recognition and risk-assessment tools.

A. FACIAL RECOGNITION SOFTWARE

Used by police departments across North America, facial recognition software identifies individuals by matching images, such as faces in surveillance videos, to faces found in a database.¹⁰⁷ However, research has shown that facial recognition tools produce higher false positives for darker-skinned individuals.¹⁰⁸ In a study by the U.S. National Institute of Standards and Technology, when comparing various facial recognition AI, they found it falsely identified Black and Asian faces 10 to 100 times more frequently than White faces.¹⁰⁹ A study by Buolamwini and Gebru that analyzed facial recognition software from Microsoft, Face ++, and

¹⁰² Chatterjee, *supra* note 19 at 191.

¹⁰³ *Ibid.*

¹⁰⁴ *Ibid*; Henderson, *supra* note 82 at 475.

¹⁰⁵ Chatterjee, *supra* note 19 at 192.

¹⁰⁶ Achiume, *supra* note 64 at para 13.

¹⁰⁷ Faith Karimi, "People of color have a new enemy: techno-racism", *CNN* (9 May 2021), online: *CNN* <<https://www.cnn.com/2021/05/09/us/techno-racism-explainer-trnd/index.html>>.

¹⁰⁸ Livingston, *supra* note 73 at 2.

¹⁰⁹ Achiume, *supra* note 64 at para 12; Karimi, *supra* note 107 at para 19.

IBM also observed this disparity.¹¹⁰ The researchers found that facial recognition software had a near-perfect classification rate (error rates of 0-0.8%) for lighter-skinned men, while error rates for darker-skinned women were as high as 35.5%.¹¹¹

The increased probability of false positives in identifying racialized individuals can lead to law enforcement incorrectly matching suspects with their criminal database.¹¹² An issue that has occurred several times in the U.S.¹¹³ For example, in 2020, Detroit police arrested Robert Williams — a Black man — outside his home based on a facial recognition match.¹¹⁴ Police forced Williams to spend 30 hours in jail before being cleared of wrongdoing.¹¹⁵ Regarding this experience, he states, “I never thought I’d have to explain to my daughters why Daddy got arrested... How does one explain to two little girls that a computer got it wrong, but the police listened to it anyway?”¹¹⁶ Although Williams was released, he was discriminated against due to the use of AI, an experience a White individual would be less likely to endure based on the rates of false positives of lighter-skinned individuals in facial recognition software.

In Canada, the situation concerning the discriminatory impact of facial recognition is not much different. The facial recognition software tool, Clearview AI, was used by numerous organizations at the federal and provincial levels, including the Royal Canadian Mounted Police (RCMP).¹¹⁷ Clearview AI is a tech start-up that offers facial recognition software created by

¹¹⁰ Gebru, *supra* note 5 at 258.

¹¹¹ *Ibid.*

¹¹² Livingston, *supra* note 73 at 2.

¹¹³ Karimi, *supra* note 107 at para 16-19.

¹¹⁴ *Ibid* at para 17.

¹¹⁵ *Ibid.*

¹¹⁶ *Ibid* at para 18.

¹¹⁷ Brenda McPhail, “Facial recognition explained: How is FRT used in Canada?” (6 December 2022) at paras 4–15, online: CCLA <<https://ccla.org/privacy/facial-recognition-explained-how-is-frt-used-in-canada/>>.

scraping 3 billion images of people from the internet and social media.¹¹⁸ It offers its tool to public and private organizations, like law enforcement agencies.¹¹⁹ However, in July 2020, Clearview AI stopped offering its services in Canada, and shortly after a federal-provincial privacy investigation into this tool, Canada's federal privacy watchdog declared that the use of Clearview AI amounts to illegal mass surveillance.¹²⁰

Other facial recognition software continues to be used by the RCMP, as well as other federal and provincial entities, with discriminatory effects.¹²¹ For example, in the 2022 *Barre v. Canada*¹²² case, two Somali women argued that their refugee status was vacated due to the use of facial recognition software by the Canadian Border Services Agency (CBSA).¹²³ The two women contended that the facial recognition software wrongly identified them as Kenyan, therefore lying about their refugee claim.¹²⁴ Ms. Asha Ali Barre and Ms. Alia Musa Hosh claim to be Sunni-Sufi Muslims who fled gender-based violence from Al-Shabaab.¹²⁵ The women were accepted as refugees by Canada in 2017 and 2018, respectively.¹²⁶ However, in 2020, the Minister of Public Safety and Emergency Preparedness applied to have their refugee status vacated before the Refugee Protection Division (RPD) of the Immigration and Refugee Protection Board because they had misrepresented themselves.¹²⁷ The Minister provided the RPD with photos that compared the two women to two Kenyan citizens who had arrived in Canada on

¹¹⁸ Maura Forrest, "RCMP's use of facial recognition extends well beyond Clearview AI" *Politico* (30 September 2022) online: *POLITICO* <<https://www.politico.com/news/2022/09/30/rcmps-facial-recognition-clearview-ai-00059639>>.

¹¹⁹ *Ibid.*

¹²⁰ *Ibid.*

¹²¹ *Ibid.* at para 7; McPhail, *supra* note 117 at para 4-15.

¹²² 2022 FC 1078

¹²³ *Ibid.* at para 1.

¹²⁴ *Ibid.*

¹²⁵ *Ibid.* at para 2.

¹²⁶ *Ibid.* at para 3.

¹²⁷ *Ibid.* at para 6.

student visas, thereby lying about their identity and allegations of persecution.¹²⁸ The two women lost their refugee status as a result.¹²⁹ Barre and Hosh maintained that they were not the two Kenyan students and appealed the decision by the RPD.¹³⁰ Justice Go of the Federal Court granted their application for judicial review and ruled that the RPD's decision was unreasonable as the RPD did not provide adequate reasoning to justify their findings concerning the facial similarities of the two women and the Kenyan students.¹³¹ This case highlights the use of facial recognition with discriminatory impacts.

The discriminatory impact of facial recognition software ties back to its training data.¹³² When a facial recognition tool is trained on primarily White faces, the subsequent AI will be better at identifying White faces and worse at identifying darker-skinned racialized faces.¹³³ The discriminatory impact of AI can further be examined through risk assessment tools.

B. RISK ASSESSMENT TOOLS

Risk assessment tools provide algorithmically driven decisions intended to deliver objective, data-centred predictions to assist or replace the judgement of human decision-makers.¹³⁴ However, these tools are trained on flawed data with deeply engrained biases that disproportionately impact racialized individuals.¹³⁵ As previously mentioned, one notable risk assessment tool used by Canadian police departments is predictive policing.¹³⁶ However,

¹²⁸ *Ibid* at para 7.

¹²⁹ *Ibid* at para 8.

¹³⁰ *Ibid* at para 10.

¹³¹ *Ibid* at para 11.

¹³² Livingston, *supra* note 73 at 2.

¹³³ *Ibid*.

¹³⁴ Achiume, *supra* note 64 at para 43; Petra Molnar and Lex Gill. "Bots at the Gate: A Human Rights Analysis of Automated Decision-Making in Canada's Immigration and Refugee System" (2018), online: *Citizen Lab and International Human Rights Program (Faculty of Law, University of Toronto)* <<https://citizenlab.ca/wp-content/uploads/2018/09/IHRP-Automated-Systems-Report-Web-V2.pdf>> at 3.

¹³⁵ Achiume, *supra* note 64 at para 43; Molnar, *supra* note 134 at 3.

¹³⁶ Reynolds, *supra* note 1 at para 7.

research suggests that predictive policing leads to the over-policing of racialized communities and is more predictive of a disadvantage than crime, as previously discussed.¹³⁷ Risk assessment tools are also used in healthcare and Canada's correctional system, resulting in discriminatory practices against racialized individuals.¹³⁸ For example, a US study found that a risk assessment tool used in healthcare prevented approximately 30% of Black patients from receiving additional care by giving them lower risk scores than White patients with equal diagnoses.¹³⁹ The study identified biases in this outcome resulting from the algorithm's inability to account for the unequal access to healthcare for Black patients.¹⁴⁰ As a result, Black patients were less likely to be recommended to healthcare programmes to improve their health.¹⁴¹ In addition to the discrimination racialized individuals experience from the use of AI, using these algorithmic risk assessment tools can infringe on a racialized individual's freedom.¹⁴² For example, the Correctional Service of Canada (CSC) utilizes risk assessment tools to evaluate a prisoner's risk of recidivism.¹⁴³ In one case, this led to the Supreme Court of Canada adjudicating that the use of these tools was not proven to be accurate when applied to Indigenous offenders.¹⁴⁴

In the 2018 case of *Ewert v. Canada*¹⁴⁵, Jeffery Ewert — a Métis man — argued that the CSC's risk assessment tools used to determine his recidivism risk are primarily developed for and tested on non-Indigenous offenders.¹⁴⁶ As such, no research confirms that these tools led to non-biased outcomes when applied to Indigenous offenders.¹⁴⁷ After a lengthy judicial process,

¹³⁷ Achiume, *supra* note 64 at para 43; Livingston, *supra* note 73 at 2.

¹³⁸ Henderson, *supra* note 82 at 484; Livingston, *supra* note 73 at 2; Reynolds, *supra* note 1 at para 8.

¹³⁹ Henderson, *supra* note 82 at 484; Livingston, *supra* note 73 at 2.

¹⁴⁰ Henderson, *supra* note 82 at 484.

¹⁴¹ Achiume, *supra* note 64 at para 30.

¹⁴² Molnar, *supra* note 134 at 45.

¹⁴³ *Ibid.*

¹⁴⁴ *Ewert v Canada* 2018 SCC 30 at para 6.

¹⁴⁵ *Ibid.*

¹⁴⁶ *Ibid* at para 12.

¹⁴⁷ *Ibid.*

the Supreme Court of Canada ruled that the CSC's use of these risk assessment tools breached their obligation under s. 24 (1) of the *Corrections and Conditional Release Act* (CCRA) — to ensure any information the CSC uses about an offender is as accurate as possible.¹⁴⁸ This case highlights that using algorithmic tools, such as risk assessment AI is unlikely to generate an accurate, individualized assessment.¹⁴⁹ Using AI created with training data that primarily involves one racial group cannot determine outcomes for another racial group that was not adequately represented (if at all) in the training data.¹⁵⁰ Rather, as observed in *Ewert*, using these risk assessment tools with biased data will result in a tool that is likely affected by cultural bias, resulting in a flawed risk assessment.¹⁵¹

It is crucial to note that the brief examples highlighted here of how AI has impacted racialized individuals are in no way exhaustive. These examples are meant to magnify how AI systemically discriminates against racialized groups while advancing the primacy of White individuals. Furthermore, the use of AI oppresses racialized individuals and infringes on their fundamental human rights at the international and domestic levels.

V. A MOVE TOWARD RESPONSIBLE ARTIFICIAL INTELLIGENCE

A. LEX LATA

In Canada, discrimination based on race is prohibited under provincial and territorial human rights codes and federal human rights legislation, such as the *Canadian Human Rights Act*.¹⁵² Racial discrimination is also prohibited under section 15 of the *Canadian Charter of*

¹⁴⁸ *Ibid* at para 80; *Corrections and Conditional Release Act*, SC 1992, c 20.

¹⁴⁹ Christian, *supra* note 14 at para 18.

¹⁵⁰ *Ibid*.

¹⁵¹ *Ibid*; *Ewert v Canada*, *supra* note 144 at para 13.

¹⁵² *Canadian Human Rights Act*, RSC 1985, c H-6; Vivek Krishnamurthy, “AI and Human Rights Law” in Florian Martin-Bariteau & Teresa Scassa, eds, *Artificial Intelligence and the Law in Canada* (Toronto: LexisNexis Canada, 2021) at 7.

Rights and Freedoms.¹⁵³ Additionally, the prohibition of racial discrimination is a peremptory norm under international law and an obligation *erga omnes*.¹⁵⁴ The prohibition of racial discrimination is further codified in several international human rights treaties ratified by Canada, such as: the *International Covenant on Economic, Social and Cultural Rights (ICESR)*,¹⁵⁵ the *International Covenant on Civil and Political Rights (ICCPR)*,¹⁵⁶ and the *International Convention on the Elimination of All Forms of Racial Discrimination (CERD)*.¹⁵⁷ Specifically, Article 26 of the *ICCPR* states:

All persons are equal before the law and are entitled without any discrimination to the equal protection of the law. In this respect, the law shall prohibit any discrimination and guarantee to all persons equal and effective protection against discrimination on any ground such as race, colour, sex, language, religion, political or other opinion, national or social origin, property, birth or other status.¹⁵⁸

The prohibition of racial discrimination is also examined in Article 2 (2) of the *ICESCR*¹⁵⁹ and is the basis of *CERD*¹⁶⁰. For example, Article 2 (d) of *CERD* states: “Each State Party shall prohibit and bring to an end, by all appropriate means, including legislation as required by circumstances, racial discrimination by any persons, group or organization.”¹⁶¹ The *CERD* treaty committee

¹⁵³ *Canadian Charter of Rights and Freedoms*, Part I of the *Constitution Act, 1982*, being Schedule B to the *Canada Act 1982 (UK)*, 1982, c 11.

¹⁵⁴ Achiume, *supra* note 64 at para 46.

¹⁵⁵ *International Covenant on Economic, Social and Cultural Rights*, 19 December 1966, 993 UNTS 3, Can TS 1976 No 46 (entered into force 3 January 1976) [ICESCR].

¹⁵⁶ *International Covenant on Civil and Political Rights*, 19 December 1966, 999 UNTS 171, Can TS 1976 No 47 (entered into force 23 March 1976) [ICCPR].

¹⁵⁷ *International Convention on the Elimination of All Forms of Racial Discrimination*, 21 December 1965, 660 UNTS 195 (entered into force 4 January 1969) [CERD].

¹⁵⁸ *International Covenant on Civil and Political Rights*, *supra* note 156.

¹⁵⁹ *International Covenant on Economic, Social and Cultural Rights*, *supra* note 155.

¹⁶⁰ *International Convention on the Elimination of All Forms of Racial Discrimination*, *supra* note 157.

¹⁶¹ *Ibid.*

(Committee on the Elimination of Racial Discrimination) has also observed that the use of AI can reinforce racial discrimination.¹⁶²

However, these ratified international treaties are only binding on States, and the *Charter* only applies to governments and government policies/programs.¹⁶³ This is an issue since private companies primarily develop AI.¹⁶⁴ Nonetheless, States do have an obligation to prevent human rights violations from private actors, such as private companies. This sentiment is further stated by the Committee on the Elimination of Racial Discrimination, which recommends that States adopt regulations to ensure AI that is created and used by private and public companies comply with human rights standards.¹⁶⁵ By not doing so, the Committee considers the State in violation of *CERD*.¹⁶⁶

Legislation such as the *Canadian Human Rights Act* protects individuals from discrimination by private companies regulated by the federal government, such as banks and telecommunication companies. Moreover, provincial and territorial human rights codes protect individuals from discrimination in areas of provincial and territorial jurisdiction, such as housing. Therefore, a person subject to discrimination by AI developed by a private company can launch a complaint to the federal or provincial human rights commissions or tribunals. If successful, their complaint may be referred to the respective federal or provincial human rights tribunal. Nevertheless, this avenue of redress is riddled with limitations.¹⁶⁷ For example, human rights tribunals are notoriously backlogged due to insufficient resources and have low success rates in

¹⁶² CERD, General recommendation No. 36 (2020) on preventing and combating racial profiling by law enforcement officials, UN Doc CERD/C/GC/36 at para 63 (2020), online: <https://tbinternet.ohchr.org/_layouts/15/treatybodyexternal/Download.aspx?symbolno=CERD%2FC%2FGC%2F36&Lang=en>.

¹⁶³ Krishnamurthy, *supra* note 152 at 6; Achiume, *supra* note 64 at para 47.

¹⁶⁴ Pizzi, *supra* note 1 at 165.

¹⁶⁵ CERD, *supra* note 162 at para 63.

¹⁶⁶ *Ibid.*

¹⁶⁷ Henderson, *supra* note 82 at 491.

determining discrimination in complaints.¹⁶⁸ Furthermore, trying to prove this discrimination legally poses a challenge.¹⁶⁹ For example, contract and intellectual property laws shield the code, data, and algorithms used to create AI.¹⁷⁰ In many cases, the developer may be unable to explain how their code functions due to how algorithms operate in a black box (i.e., the inability to peer into an AI system to understand its processes due to its opaque nature).¹⁷¹ This lack of transparency makes it difficult for impacted groups to prove discrimination as required by evidentiary burdens of proof.¹⁷²

As this paper has examined, the use of AI in both the private and public sphere discriminates against racialized individuals, infringing upon their fundamental human rights codified in national and international laws. The private companies primarily responsible for the development of AI operate in a position of power where they cannot be held accountable through the adjudicative process for racial discrimination as a result of their AI. However, they do have a responsibility to prevent discrimination from occurring in the first place.

As articulated in the *UN Guiding Principles on Business and Human Rights*, private companies have a responsibility to respect human rights.¹⁷³ Respecting human rights demands that businesses avoid contributing to human rights violations through their activities and operations.¹⁷⁴ To do so, the *UN Guiding Principles on Business and Human Rights* calls on businesses to conduct human rights due diligence which requires assessing actual and potential

¹⁶⁸ *Ibid* at 492.

¹⁶⁹ Achiume, *supra* note 64 at para 44.

¹⁷⁰ *Ibid*.

¹⁷¹ Achiume, *supra* note 64 at para 34; Henderson, *supra* note 82 at 491; Pizzi, *supra* note 1 at 147.

¹⁷² Achiume, *supra* note 64 at para 34; Henderson, *supra* note 82 at 491; Pizzi, *supra* note 1 at 147.

¹⁷³ *Guiding Principles on Business and Human Rights*, HRC Res 17/4, UN Doc HR/PUB/11/04 (16 June 2011) at 1; Achiume, *supra* note 64 at para 60; Krishnamurthy, *supra* note 152 at 6.

¹⁷⁴ Krishnamurthy, *supra* note 152 at 6.

human rights impacts, acting on the findings, and stating how these impacts are addressed.¹⁷⁵ As stated in Article 11 of the *UN Guiding Principles on Business and Human Rights*, “[b]usiness enterprises *should* [emphasis added] respect human rights.”¹⁷⁶ It is important to recognize how this language differs from what is required by States. For example, article 1 of the *Guiding Principles* asserts that “States *must* [emphasis added] protect against human rights abuse within their territory and/or jurisdiction by third parties, including business enterprises.”¹⁷⁷ While the *UN Guiding Principles on Business and Human Rights* is not a mandatory framework for private companies, these *Guiding Principles*, along with the international treaties previously mentioned, place the responsibility of prevention on States. However, States should work with the various private companies that design and deploy AI to further develop special measures in mitigating the racially discriminatory impacts of AI in both the public and private spheres.¹⁷⁸ This collaborative approach has been observed through Canada’s Algorithmic Impact Assessment.

The Treasury Board of Canada Secretariat recently released an Algorithmic Impact Assessment tool for public sector departments to use when considering AI.¹⁷⁹ The Algorithmic Impact Assessment tool was created by working closely with the public, experts in the AI industry, and academics across Canada.¹⁸⁰ However, this tool only applies to a limited number of departments in the public sector, is modelled after an environmental assessment, and lacks any human-rights basis.¹⁸¹ While the current use of Canada’s Algorithmic Impact Assessment is

¹⁷⁵ *UN Guiding Principles on Business and Human Rights*, *supra* note 173 at 17; Achiume, *supra* note 64 at para 60; Pizzi, *supra* note 1 at 165.

¹⁷⁶ *UN Guiding Principles on Business and Human Rights*, *supra* note 173 at 13.

¹⁷⁷ *Ibid* at 3.

¹⁷⁸ Achiume, *supra* note 64 at para 53; Pizzi, *supra* note 1 at 165.

¹⁷⁹ Fenwick McKelvey & Margaret MacDonald, “Artificial Intelligence Policy Innovations at the Canadian Federal Government” (2019) 44:2 *Canadian Journal of Communication* 43 at 46, online: <<https://cjc.utpjournals.press/doi/10.22230/cjc.2019v44n2a3509>>.

¹⁸⁰ *Ibid* at 45-46.

¹⁸¹ *Ibid*.

admirable, it remains weak. In order to provide effective human rights protection, address systemic racial discrimination, and advance responsible AI globally, AI governance should be extensive, applicable to both the private and public spheres, and based on a human rights approach.

B. LEX FERENDA

Responsible AI is not an abstract concept. Many organizations at the public and private level have developed their own assessments to mitigate the inequalities that AI advances.¹⁸² However, there is a lack of standardization as these assessments vary from organization to organization, and without enforcement, there is little guarantee of adherence.¹⁸³ A practice often referred to as ‘ethics-washing’.¹⁸⁴ The phenomenon where private companies self-regulate the design and deployment of their AI through vague, opaque, and unenforceable codes of ethics.¹⁸⁵ Furthermore, these frameworks that promote compliance are also often done without the input of those adversely impacted by AI.¹⁸⁶ Also, research on the discriminatory impacts of AI fails to provide answers as to what AI governance should entail besides stressing the importance of transparency, accountability and fairness.¹⁸⁷

To advance responsible AI, Canada must address the lack of representation in the tech sector.¹⁸⁸ To do so, this paper recommends establishing an independent body, similar to the Office of the Privacy Commissioner, to regulate the development of AI technologies in both the

¹⁸² Livingston, *supra* note 73 at 3.

¹⁸³ *Ibid.*

¹⁸⁴ Pizzi, *supra* note 1 at 167.

¹⁸⁵ *Ibid.*

¹⁸⁶ *Ibid* at 148.

¹⁸⁷ Aizenberg, *supra* note 5 at 1.

¹⁸⁸ Achiume, *supra* note 64 at para 55; Karimi, *supra* note 107 at para 36.

private and public sectors.¹⁸⁹ This independent body must work with private companies to address the equity gap in the tech sector by ensuring racialized groups are meaningfully represented in the decision-making process, design, and deployment of emerging AI technologies. In addition, this newly established body must work with those adversely affected by AI to create and administer a mandatory human rights assessment. This assessment will serve as standardized due diligence for all companies developing AI. Doing so will address the criticisms that researchers have brought forth concerning the lack of standards and compliance with similar assessments that the private companies create.¹⁹⁰

Lastly, this paper recommends the creation of a regulatory framework for high-risk AI systems. The European Union has proposed such legislation in the AI Act.¹⁹¹ The AI Act creates a framework that assesses and categorizes AI into high or low-risk categories, with legislation in place for AI deemed high-risk and unacceptable based on human rights laws.¹⁹² The draft regulation was proposed by the European Commission in April 2021, and on December 6, 2022, the European Council adopted its common position.¹⁹³ The European Council's adoption of the general approach will now allow the Council to enter negotiations with the European Parliament regarding this proposed regulation.¹⁹⁴ To combat the discriminatory impact of AI, it is recommended that Canada follow in the European Union's footsteps with the AI Act in

¹⁸⁹ Christelle Tessono & Sonja Solomun, "How to Fix Canada's Proposed Artificial Intelligence Act" *Tech Policy Press* (6 December 2022), online: *Tech Policy Press* <<https://techpolicy.press/how-to-fix-canadas-proposed-artificial-intelligence-act/>>; Andrew Clement, "The Artificial Intelligence and Data Act needs a reset" *The Hill Times* (23 November 2022), online: *The Hill Times* <<https://www.hilltimes.com/story/2022/11/23/the-ai-and-data-act-needs-a-reset/356482/>>.

¹⁹⁰ Livingston, *supra* note 73 at 3.

¹⁹¹ Clement, *supra* note 189 at para 6; Future of Life Institute, "The Artificial Intelligence Act" (7 September 2021), online: <<https://artificialintelligenceact.eu/>>; European Council, "Artificial Intelligence Act: Council calls for promoting safe AI that respects fundamental rights" (6 December 2022), online: <<https://www.consilium.europa.eu/en/press/press-releases/2022/12/06/artificial-intelligence-act-council-calls-for-promoting-safe-ai-that-respects-fundamental-rights/>>.

¹⁹² Clement, *supra* note 189 at para 6; Future of Life Institute, *supra* note 191; European Council, *supra* note 191.

¹⁹³ European Council, *supra* note 191.

¹⁹⁴ *Ibid* at para 24.

establishing a regulatory framework for all AI created in its jurisdiction. In putting forth these recommendations for AI governance that provides effective human rights protections, Canada can be a leader in advancing responsible AI globally, in addition to its reputation as an AI innovation hub.¹⁹⁵

VI. CONCLUSION

While one may hope AI would enhance objectivity by counteracting human biases, this is not the case.¹⁹⁶ Humans build AI with their own biases, and data that reflects deeply engrained societal inequalities.¹⁹⁷ As this paper has proven, this has led to AI perpetuating, and in some cases, worsening discrimination against racialized individuals.¹⁹⁸ Examined through a critical race lens, the exploration of how AI is characterized as White, the use of AI in eliminating the fuller attributes of racialized bodies, and the racially homogenous nature of the tech sector, demonstrates that AI advances white supremacist ideology by creating a system where White individuals overwhelmingly control power and privilege. In addition, magnifying the lived experiences of racialized individuals reveals how AI systemically discriminates against racialized groups, further oppressing racialized individuals. This discrimination also infringes on fundamental human rights, as examined in international and domestic law. It is clear that AI governance is required to combat the discriminatory impacts of AI. To do so, this paper advocates for AI governance based on a human rights approach and has provided recommendations to mitigate the discriminatory impact of AI.

Of note, discourse on AI governance was initially framed in terms of ethics, which some advocates find to offer more flexibility than the rigidity of the law and more adaptability to

¹⁹⁵ Reynolds, *supra* note 1 at para 2; Pizzi, *supra* note 1 at 150.

¹⁹⁶ Henderson, *supra* note 82 at 475.

¹⁹⁷ *Ibid.*

¹⁹⁸ *Ibid.*

technological evolution.¹⁹⁹ However, there is increasing recognition that the impacts of AI must be addressed from a human rights perspective.²⁰⁰ It is worth observing that human rights and ethics essentially have the same goals, and any ‘unethical’ use of AI will also likely violate human rights (and vice versa).²⁰¹ Nonetheless, the recommendations suggested in this paper are based on a human rights approach to curtail private companies from ‘ethics-washing.’²⁰²

There can no longer be any doubt that AI can reproduce, reinforce and exacerbate racial inequalities within society.²⁰³ As a result of growing concern about the role of AI in bolstering societal inequities, the federal government of Canada has recently drafted the Artificial Intelligence and Data Act.²⁰⁴ Legislation aimed at prohibiting specific outcomes of AI that result in harm to individuals.²⁰⁵ However, critics have noted that this legislation is far too narrow in scope to address the widely examined harms posed by AI.²⁰⁶ For example, the proposed legislation does not apply to certain government institutions such as the Canadian Security Intelligence Service (CSIS), does not address any human rights risks associated with AI, and does not prohibit the use of any AI systems which may be deemed high-risk.²⁰⁷ It also lacks provisions for independent oversight of the AI industry and has significant inconsistencies with regard to what it defines as an AI system.²⁰⁸ While it remains to be seen whether this drafted legislation will press ahead, it is evident that Canada needs to act quickly. Only through proper AI regulation and governance will all Canadians be able to relish in the benefits of AI.

¹⁹⁹ Krishnamurthy, *supra* note 152 at 3; Pizzi, *supra* note 1 at 166.

²⁰⁰ Krishnamurthy, *supra* note 152 at 3.

²⁰¹ Pizzi, *supra* note 1 at 167.

²⁰² *Ibid.*

²⁰³ Achiume, *supra* note 64 at para 12.

²⁰⁴ Clement, *supra* note 189 at para 2.

²⁰⁵ *Ibid.*

²⁰⁶ *Ibid* at para 3.

²⁰⁷ Tessono, *supra* note 189.

²⁰⁸ *Ibid.*

Bibliography

Legislation

Corrections and Conditional Release Act, SC 1992, c 20

Canadian Human Rights Act, RSC 1985, c H-6

Canadian Charter of Rights and Freedoms, Part I of the *Constitution Act, 1982*, being Schedule B to the *Canada Act 1982 (UK)*, 1982, c 11.

Jurisprudence

Barre v Canada 2022 FC 1078

Ewert v Canada 2018 SCC 30

International Law Sources

International Covenant on Civil and Political Rights, 19 December 1966, 999 UNTS 171, Can TS 1976 No 47 (entered into force 23 March 1976) [ICCPR].

International Covenant on Economic, Social and Cultural Rights, 19 December 1966, 993 UNTS 3, Can TS 1976 No 46 (entered into force 3 January 1976) [ICESCR].

International Convention on the Elimination of All Forms of Racial Discrimination, 21 December 1965, 660 UNTS 195 (entered into force 4 January 1969) [CERD].

International Documents

E. Tendayi Achiume, Racial discrimination and emerging digital technologies: a human rights analysis, HRC 44th Sess, UN Doc A/HRC/44/57 (2020), online: <<https://documents-dds-ny.un.org/doc/UNDOC/GEN/G20/151/06/PDF/G2015106.pdf?OpenElement>>.

Guiding Principles on Business and Human Rights, HRC Res 17/4, UN Doc HR/PUB/11/04 (16 June 2011).

CERD, General recommendation No. 36 (2020) on preventing and combating racial profiling by law enforcement officials, UN Doc CERD/C/GC/36 (2020), online: <https://tbinternet.ohchr.org/_layouts/15/treatybodyexternal/Download.aspx?symbolno=CERD%2FC%2FGC%2F36&Lang=en>.

Secondary Materials: Journal Articles

Aizenberg, Evgeni & Jeroen van den Hoven, “Designing for human rights in AI” (2020) 7:2 Big Data & Society 1.

Bartneck, Christoph et al, Robots and Racism (New York, NY, USA: Association for Computing Machinery, 2018), online <<http://hdl.handle.net/10092/15024>>.

Cave, Stephen & Kanta Dihal, “The Whiteness of AI” (2020) 33:4 Philos Technol 685.

Chatterjee, Sheshadri, Sreenivasulu NS & Zahid Hussain, “Evolution of artificial intelligence and its impact on human rights: from sociolegal perspective” (2021) 64:2 International Journal of Law and Management 184.

Henderson, Bradley, Colleen M. Flood, and Teresa Scassa, “Artificial Intelligence in Canadian Healthcare: Will the Law Protect Us from Algorithmic Bias Resulting in Discrimination?” (2022) 19:2 CJLT 475.

Khoday, Amar, “Black Voices Matter Too: Counter-Narrating *Smithers v The Queen*” (2021) 58:3 Osgoode Hall LJ 567.

Lawrence, Stefan & Kevin Hylton, “Critical Race Theory, Methodology, and Semiotics: The Analytical Utility of a ‘Race’ Conscious Approach for Visual Qualitative Research” (2022) 22:3 Cultural Studies ↔ Critical Methodologies 255.

Livingston, Morgan, “Preventing Racial Bias in Federal AI” (2020) 16:02 JSPG, online: <https://www.sciencepolicyjournal.org/article_1038126_jspg160205.html>.

McKelvey, Fenwick & Margaret MacDonald, “Artificial Intelligence Policy Innovations at the Canadian Federal Government” (2019) 44:2 Canadian Journal of Communication 43, online: <<https://cjc.utpjournals.press/doi/10.22230/cjc.2019v44n2a3509>>.

Moran, Taylor C, “Racial technological bias and the white, feminine voice of AI VAs” (2021) 18:1 Communication and Critical/Cultural Studies 19.

Noriega, Maria, “The application of artificial intelligence in police interrogations: An analysis addressing the proposed effect AI has on racial and gender bias, cooperation, and false confessions” (2020) 117 *Futures* 102510.

Pizzi, Michael, Mila Romanoff & Tim Engelhardt, “AI for humanitarian action: Human rights and ethics” (2020) 102:913 *Int rev Red Cross* 145.

Poole, Sonja Martin et al, “Operationalizing Critical Race Theory in the Marketplace” (2021) 40:2 *Journal of Public Policy & Marketing* 126.

Richardson, Rashida, “Racial Segregation and the Data-Driven Society: How Our Failure to Reckon with Root Causes Perpetuates Separate and Unequal Realities” (2022) 36:3 *Berkeley Tech LJ* 1051.

Zajko, Mike, “Conservative AI and social inequality: conceptualizing alternatives to bias through social theory” (2021) 36:3 *AI & Soc* 1047, online: <<https://doi.org/10.1007/s00146-021-01153-9>>.

Secondary Materials: Online Resources

Akselrod, Olga, “How Artificial Intelligence Can Deepen Racial and Economic Inequities” *ACLU* (13 July 2021), online: *American Civil Liberties Union* <<https://www.aclu.org/news/privacy-technology/how-artificial-intelligence-can-deepen-racial-and-economic-inequities>>.

Christian, Gideon, “Artificial Intelligence, Algorithmic Racism and the Canadian Criminal Justice System” *SLAW* (26 October 2020), online: *SLAW* <<http://www.slaw.ca/2020/10/26/artificial-intelligence-algorithmic-racism-and-the-canadian-criminal-justice-system/>>.

European Council, “Artificial Intelligence Act: Council calls for promoting safe AI that respects fundamental rights” (6 December 2022), online: <<https://www.consilium.europa.eu/en/press/press-releases/2022/12/06/artificial-intelligence-act-council-calls-for-promoting-safe-ai-that-respects-fundamental-rights/>>.

Forrest, Maura, “RCMP’s use of facial recognition extends well beyond Clearview AI” *Politico* (30 September 2022), online: *POLITICO* <<https://www.politico.com/news/2022/09/30/rcmps-facial-recognition-clearview-ai-00059639>>.

Global Advantage Consulting Group, “Canada’s AI Ecosystem: Government Investment Propels Private Sector Growth” (2021), online: *University of Toronto* <https://gro.utoronto.ca/wp-content/uploads/2022/04/Canadas_AI_Ecosystem-_Government_Investment_Propels_Private_Sector_Growth.pdf>

McPhail, Brenda, “Facial recognition explained: How is FRT used in Canada?” (6 December 2022), online: *CCLA* <<https://ccla.org/privacy/facial-recognition-explained-how-is-frt-used-in-canada/>>.

Molnar, Petra & Lex Gill. “Bots at the Gate: A Human Rights Analysis of Automated Decision-Making in Canada’s Immigration and Refugee System” (2018), online: *Citizen Lab and International Human Rights Program (Faculty of Law, University of Toronto)* <<https://citizenlab.ca/wp-content/uploads/2018/09/IHRP-Automated-Systems-Report-Web-V2.pdf>>.

Tessono, Christelle & Sonja Solomun, “How to Fix Canada’s Proposed Artificial Intelligence Act” *Tech Policy Press* (6 December 2022), online: *Tech Policy Press* <<https://techpolicy.press/how-to-fix-canadas-proposed-artificial-intelligence-act/>>.

Townsend, Chance, “Startup develops AI that makes call center employees sound like white, American robots” (27 August 2022), online: *Mashable* <<https://mashable.com/article/ai-startup-makes-call-center-foreigners-sound-white>>.

Vu, Viet, Creig Lamb & Asher Zafar, “Who Are Canada’s Tech Workers?” (Toronto, ON, CA: Brookfield Institute for Innovation Entrepreneurship, 2019).

Future of Life Institute, “The Artificial Intelligence Act” (7 September 2021), online: <<https://artificialintelligenceact.eu/>>.

Secondary Materials: Newspapers

Clement, Andrew, “The Artificial Intelligence and Data Act needs a reset” *The Hill Times* (23 November 2022), online: *The Hill Times* <<https://www.hilltimes.com/story/2022/11/23/the-ai-and-data-act-needs-a-reset/356482/>>.

Karimi, Faith, “People of color have a new enemy: techno-racism” *CNN* (9 May 2021), online: *CNN* <<https://www.cnn.com/2021/05/09/us/techno-racism-explainer-trnd/index.html>>.

Reynolds, Chris, “Canada needs laws to tackle life-and-death problems involving AI, experts warn” *CTV News* (19 May 2019), online: *CTV News* <<https://www.ctvnews.ca/sci-tech/canada-needs-laws-to-tackle-life-and-death-problems-involving-ai-experts-warn-1.4428959?cache=yes>>.

Secondary Materials: Books

Gebu, Timnit, “Race and Gender” in Markus D Dubber, Frank Pasquale & Sunit Das, eds, *The Oxford Handbook of Ethics of AI* (Oxford University Press, 2020) 251.

Krishnamurthy, Vivek, “AI and Human Rights Law” in Florian Martin-Bariteau & Teresa Scassa, eds, *Artificial Intelligence and the Law in Canada* (Toronto: LexisNexis Canada, 2021).