The Impacts and Governance of Artificial Intelligence: A Transfeminist Perspective

by

Blair Attard-Frost

A thesis submitted in conformity with the requirements

for the degree of Doctor of Philosophy

Faculty of Information

University of Toronto

The Impacts and Governance of Artificial Intelligence:
A Transfeminist Perspective

Blair Attard-Frost
Doctor of Philosophy
Faculty of Information
University of Toronto
2025

Abstract

This thesis investigates the impacts and governance of artificial intelligence (AI) systems through a transfeminist lens, focusing analysis upon challenges of power, exclusion, and injustice alongside opportunities for advancing equity, community-based resistance, and transformative change. AI governance is a field of research and practice seeking to maximize benefits and minimize harms caused by AI systems. However, AI governance is frequently ineffective at preventing AI systems from causing harm to society and the environment, with historically marginalized groups being particularly vulnerable to harm. Applying a framework of theories drawn from service science, feminist studies, and trans studies, I analyze relationships between AI governance and harm prevention in three separate co-authored, peer-reviewed articles. The first article develops a theory linking beneficial and harmful impacts caused by AI systems to the value chains through which various actors integrate resources and co-create value throughout the AI system lifecycle. This theory is applied to an integrative review of ethical concerns implicated in AI systems and to discuss future directions for intervening in the impacts caused by AI systems. The second article presents a semi-systematic review and content analysis of 84 AI governance initiatives launched by federal and provincial governments in Canada from 2017 to 2022. AI governance initiatives are used to organize many types of interventions, and Canada's initiatives favor intervention in the impacts of AI on Canadian industry, innovation, and technology production and adoption over more direct intervention in societal and environmental impact. The third article applies thematic analysis methods to data from interviews with 20 leaders of Canadian AI governance initiatives and subject matter experts. AI governance systems in Canada function at multiple levels of scale, and Canada's national-scale AI governance system consists of a diverse range of actors, resources, networks, logics, functional bounds, rules, and perceptions of benefit and harm. The thesis concludes by synthesizing findings from the three articles into a set of reflections with a unifying theme: the effectiveness of AI governance at preventing harm is limited by power imbalances, and a transfeminist approach to AI governance can support future AI governance research, practices, and systems in addressing those limitations.

Acknowledgements

A cornerstone of many trans and feminist relational ecologies is that there is no such thing as individual action. Instead, there are networks of interdependent human and nonhuman beings who think together, feel together, act together, and support each other in various modalities and intensities as we transition through our lives together. Though the title page bears my name, this dissertation was written by a network. Many people have powerfully contributed to my life, my research, and my perspective as a scholar over the course of my doctoral studies.

When others ask who is serving on my supervisory committee, a common reaction they have after I tell them is something along the lines of “that’s a unique and excellent committee!” As my supervisor, Kelly Lyons has guided me with extraordinary care and patience through all sorts of academic obstacle courses, labyrinthian research challenges, and strange new projects that I felt were important for me to pursue (even when I couldn’t quite articulate a clear explanation of why the project was so important to me). When I venture into new areas of thought or new approaches to my research, Kelly always lends a keen eye and a steady hand, giving me freedom and encouragement to go forth and explore while also knowing when to help anchor me if I stray too far from a gainful path. Kelly understands how systems work: I am frequently amazed by how easily she can help me navigate extremely complex problem spaces and pick apart the structures and mechanisms we find in those spaces. First during my master’s and now throughout my PhD, Kelly has been a constant source of confidence, intellect, and growth in my academic life. Chun Wei Choo and Leslie Shade, my other supervisory committee members, have also had a vital part in helping me flourish as a scholar. Chun Wei has brought kindness and gentleness, tremendous insight into organizational behavior, and a vast depth of knowledge to both my master’s and my doctoral studies. Chun Wei has an unparalleled ability to point me toward the exact theoretical and methodological resources I need to help me understand how the organizations I'm studying manage their technologies and create value from their knowledge. Leslie mixes cheerfulness with rigor and diligence, showering praise upon my work while challenging me to go even further. Leslie has a profound and intuitive understanding of how policy and power intersect–when I am trying to hold a system, a policy, or a process up to the light of scrutiny, Leslie always points out several ways I could keep sharpening my critique. I owe much to all three of you. Thank you for all of your patience, guidance, and wisdom. Thank you also to Christoph Becker and Joanna Redden for all of your insightful comments as internal-external and external examiner of my thesis.

I've had the immense privilege of collaborating, exchanging ideas with, and learning from many co-authors, colleagues, and scholarly friends. Ana Brandusescu, David Widder, Christelle Tessono, and Em Cariglino: the systems we study are too often machines of injustice, and you have my deepest gratitude for always helping me remain grounded, motivated, and lighthearted while we slogged through darkness together. Helen Hayes, Andrés De los Ríos, Deneille Walters, and Meghan Voll: Thank you for all the trust you placed in me to explore new territory in AI policy research, AI ethics, and mobile dating studies alongside you. Réka Gál, Kai-Hsin Hung, and Julian Posada: At critical junctures throughout my research, you've shared wondrous ideas and provided advice that this dissertation greatly benefits from. Thank you for helping to steer me in the right direction.

It has been an honor and a delight teaching so many bright and talented students. To my students: thank you for teaching me. Thank you for reminding me of the courage it takes to find your voice and to engage in conversations about new, unfamiliar topics. Thank you for sharing all of your hopes and fears about AI, all of your rich perspectives, and all of your brilliant insights into how AI impacts your own work and your own lives. Thank you for showing me new windows into my own research that I didn't even know existed.

In addition to my students and my supervisory committee members, I am also fortunate to have learned much from many other educators over these years. Christoph Becker, Colin Furness, Harald Bathelt, and Costis Dallas: Thank you for not only teaching me about information systems and research design, but also for teaching me how to teach through the example you set in your own classrooms. Fenwick McKelvey, Pamela Lirio, Phoebe Moore, Rafael Grohmann, Seamus Ross, Renee Sieber, and Luke Stark: Thank you for helping me learn how to show leadership in research and in education through our collaborations and conversations together.

A very special thanks to Christine Chan and all the other iSchool administrative staff: thank you for always being so friendly and supportive in helping me navigate all kinds of academic processes and procedures. Thank you also to all the anonymous reviewers of my research papers and all the participants in my research: this dissertation has been deeply shaped by all the thoughtful input you've generously provided.

Beyond academia, my research has been inspired by many storytellers, advocates for change, and public servants with whom I've had the privilege of discussing our concerns about AI or collaborating as part of various projects, articles, and events. Tanya Kan and Bianca Wylie: Thank you for showing such integrity, conviction, and compassion in consistently holding power to account and speaking out against injustice in tech, even when your voice may not be well-received by some. My own advocacy owes much to the example you set. Nafouel Testaouni, Andy Saldaña, Natalie Bravo, Layah Glassman, and all the other lovely folks at QueerTech, Start Proud, and Canada Learning Code: Thank you for your enduring efforts to build queer community and equip 2SLGBTQI+ people with the knowledge, skills, and lenses we need to prosper and protect ourselves in a world that is too often unkind to us. Mél Hogan, Justin Hendrix, Jess Reia, Jackson Weaver, Sarah Samuel, Titus Blome, Andrey Kurenkov, Murad Hemmadi, Anita Balakrishnan, Mani Moksha, Daniel Sarah Karasik, Ruth Bankey, Michael Anton Dila, Neha Vatish, Liva Zemitis, Ella Snyder, Elisha Lim, and Christine Tran: Thank you for sharing your platforms and your spaces with me. I am so grateful for the opportunities you provided for me to dig deeper into my research and share my perspective with new audiences.

Outside of work, I've been blessed with friends who have supported me as a human being in so many ways. Thank you all for your friendship. It was certainly not easy to begin writing a dissertation in the midst of a pandemic, a family fracture, and a gender transition, but the kindness and care you showed me gave me the strength to keep pushing ahead. Thanks especially to my best friend Huxley, a quirky and sweet little dog who helps keep my spirits high and my thoughts grounded amidst so many abstractions.

Finally, infinite thanks and infinite love to my family, and especially to my Mom. She has unconditionally supported me through everything, through every day of my studies and every day of my life. In my most desperate times, I kept writing this thing for her and because of her.

Funding Statement

This research was supported by a scholarship from the Social Sciences and Humanities Research Council of Canada (SSHRC), by scholarships from the Government of Ontario, by awards from the University of Toronto Faculty of Information, and by a research stipend generously provided by my supervisor Kelly Lyons through her Discovery Grant from the Natural Sciences and Engineering Research Council of Canada (NSERC).

Table of Contents

Chapter 1: Introduction 1

1.1. Background & Motivation 2

1.2. Research Scope & Objectives 4

1.3. Structure & Content of Dissertation 6

1.4. Meta-theoretical Framework 10

1.4.1. Overview of Meta-theoretical Framework 10

1.4.2. Researcher Positionality 12

1.4.3. Service Realism 13

1.4.4. Transfeminist AI Ethics 18

1.5. Contributions of Dissertation 23

Chapter 2: The Impacts of AI: A Theory and Analysis of AI Value Chains 25

2.1. Introduction 27

2.2. Theory 29

2.2.1. Value Chains 29

2.2.2. Supply Chains vs. Value Chains 31

2.2.3. AI Value Chains 34

2.3. Methodology 37

2.4. Ethical Implications of AI Value Chains 41

2.4.1. AI Value Chains & Benefits of AI 41

2.4.2. AI Value Chains & Issues Arising from Machine Learning 42

2.4.3. AI Value Chains & Ethics of Living in a Digital World 44

2.4.4. AI Value Chains & Metaphysical Issues 48

2.5. Limitations 49

2.6. Future Directions for Research, Practice, & Policy 49

2.7. Conclusion 52

Chapter 3: AI Governance Initiatives in Canada 54

3.1. Introduction 56

3.2. Theoretical Background 58

3.2.1. Artificial Intelligence & AI Governance 58

3.2.2. Initiatives & AI Governance Initiatives 59

3.3. Research Design & Methodology 60

3.4. Findings 65
3.4.1. Overview of Findings 65
3.4.2. Intervention Types 68
3.4.3. Governance Areas 69
3.4.4. Program Initiatives 71
3.4.5. Policy Initiatives 74
3.4.6. Strategic Plan Initiatives 76
3.4.7. Ethics Statement Initiatives 77
3.4.8. Standards Initiatives 78
3.5. Opportunities for Strengthening AI Governance in Canada 79
3.5.1. Opportunities for Researchers 79
3.5.2. Opportunities for Practitioners 85
3.6. Limitations 92
3.7. Conclusions 94
Chapter 4: AI Governance Systems in Canada 96
4.1. Introduction 98
4.2. Theoretical Background 100
4.2.1. Artificial Intelligence 100
4.2.2. AI Governance Systems at Micro and Macro Scales 101
4.2.3. Service Systems 102
4.3. Research Context: AI Governance in Canada 103
4.4. Research Design & Methodology 106
4.5. Findings 110
4.5.1. Work Contexts 112
4.5.2. Actors 113
4.5.3. Benefits, Risks, & Harms 114
4.5.4. Resources 115
4.5.5. Networks 116
4.5.6. Interactions 117
4.5.7. Evaluations 118
4.5.8. Logics 118
4.5.9. Functional Bounds 121
4.5.10. Rules 122
4.5.11. Ecosystem 125
4.5.12. Opportunities for Improvement 127

4.6. Limitations 129
4.7. Implications for Future Research, Practice, & Policy 131
4.7.1. Implications for Research 131
4.7.2. Implications for Practice & Policy 134
4.8. Conclusion 135
Chapter 5: Conclusion: A Future for Transfeminist AI Governance 138
5.1. Summary of Findings & Recommendations 139
5.2. Reflections on Research: Further Discussion of Findings & Future Pathways 146
5.2.1. Reflection 1: AI governance is an exercise of power 146
5.2.2. Reflection 2: Top-down AI governance is limited by power imbalances 153
5.2.3. Reflection 3: AI governance must be transformed 160
References 172
References for Chapter 1 173
References for Chapter 2 186
References for Chapter 3 205
References for Chapter 4 226
References for Chapter 5 237
Appendices 245
Appendix 2A: Integrated Inventory 246
Appendix 3A: Inventory of Relevant Initiatives 252
Appendix 3B: Breakdown of Initiative Data 271
Appendix 4A: Dataset 272
Appendix 4B: Service System Analysis Framework & AIG System Process Model 297
Appendix 4C: Questionnaire for Initial Interviews 299
Appendix 4D: Questionnaire for Follow-up Interviews 302
Appendix 4E: Inclusion Criteria for Participation in Study 303
Appendix 4F: Strategic Objectives & Potential Interventions 304

List of Figures

Figure 1.1: AI systems, AI governance systems, and socio-material context	3
Figure 2.1: Supply chain ontology & value chain ontology	33
Figure 2.2: Process model of AI value chains	37
Figure 2.3: Research process for integrative literature review	40
Figure 3.1: Research process for semi-systematic literature review	61
Figure 3.2: Content analysis process, codes, themes, & opportunities	65
Figure 3.3: Key federal organizations & initiatives as of 2023	68
Figure 3.4: Bar chart of initiative intervention types & years of origin	69
Figure 3.5: Bar chart of initiative governance areas & years of origin	71
Figure 4.1: Key federal organizations & initiatives as of 2024	103
Figure 4.2: Research process for AI governance systems analysis	107
Figure 4.3: Typology of actor types, sub-types, & instances	114
Figure 4.4: Ecosystem-level entities and relations	127
Figure 4.5: Relationships between opportunities, gaps, & strategic objectives	135
Figure 4B1: Service systems analysis framework	297
Figure 4B2: AI governance process model	298

List of Tables

Table 1.1: Summary of meta-theoretical framework 11

Table 1.2: Summary of contributions 24

Table 3.1: Intervention types identified during content analysis 66

Table 3.2: Governance areas identified during content analysis 66

Table 4.1: Summary of key findings from AI governance systems analysis 110

Table 5.1: Research objectives & key findings of dissertation 139

Table 5.2: Lessons learned from four cases of AI countergovernance 167

Chapter 1

Introduction

1.1. Background & Motivation

Artificial intelligence (AI) has no universally agreed upon definition, but AI is often characterized as a technological system that processes data and other resources into predictions, recommendations, decisions, or content within a socio-material context and with some degree of autonomy from human actors (Dwivedi et al., 2021; Madan & Ashok, 2023; Nitzberg & Zysman, 2022; OECD, 2024). Because AI systems are capable of causing many different types of beneficial and harmful impacts to society and the environment, *AI governance* has emerged as a field of research and practice. Like AI itself, AI governance has no universally agreed upon definition, but scholars often characterize AI governance as a system of practices intended to maximize benefits and minimize harms caused by AI systems (Camilleri, 2023; Dwivedi et al., 2021; Kuziemski & Misuraca, 2020; Nitzberg & Zysman, 2022; Wirtz, Weyerer, & Kehl, 2022). AI governance practices are intended to intervene in the impacts of AI systems through the design and implementation of several types of initiatives, including (but not limited to) strategies, policies, government programs, standards, and ethics codes (Birkstedt et al., 2023).

In this dissertation, I position AI governance as a field of study and practice, and as a type of social system. AI governance can be conceptualized as a "a *system* whose constituent elements should be interlinked to form a functional entity" (Mäntymäki et al., 2022, p. 604, emphasis added). In this understanding of AI governance as a systemic phenomenon, *AI governance systems* include "rules, practices, processes, and technological tools" (Mäntymäki et al., p. 605) in addition to many other structural and functional components that collectively enable the AI governance system to socially and materially intervene in benefits and harms caused by AI systems. These characteristics position AI governance systems as second-order systems that enable social actors to intervene in interactions between AI systems and their socio-material contexts (see Figure 1.1). AI governance systems have been studied across various contexts and at various scales of activity, including international AI governance activities (Erdélyi & Goldsmith, 2022; Roberts et al., 2023; Tallberg et al., 2023), national AI governance activities (Djeffal, Siewert, & Wurster, 2022; Radu, 2021; Wilson, 2022), local and municipal AI governance activities (Kinder et al., 2023; Wan & Sieber, 2023), sectoral AI governance activities (Kuziemski & Misuraca, 2020; Zuiderwijk, Chen, & Salem, 2021), and organizational AI governance activities (Cihon, Schuett, & Baum, 2021; Mäntymäki et al., 2022).

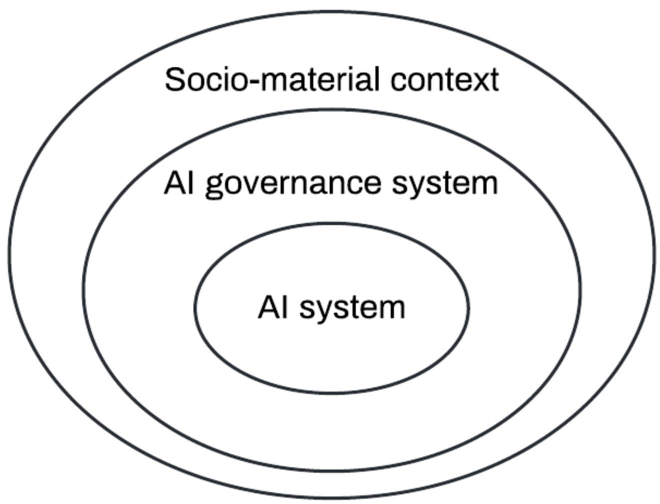

Figure 1.1: Model depicting the intermediary position of AI governance systems between AI systems and their socio-material contexts.

Despite many practices of AI governance across these scales and contexts, AI systems frequently cause harm to society and to the environment. Several datasets and databases have been created to take inventory of actual harms caused by real-world AI systems (AIAAIC, 2025; AI Incident Database, 2025a). As of January 2025, the AIAAIC Repository contains records of 1854 incidents of AI systems causing harm to humans, while the AI Incident Database contains records of 897 incidents. The AI Incident Database categorizes AI harms into many different levels of severity and types, including harm to social or political systems, psychological harm,

harm to physical health and safety, harm to civil liberties, financial harm, harm to physical property, harm to intangible property, reputational harm, and harm to privacy (AI Incident Database, 2025b). The harms caused by AI systems have become so significant in scale and scope that many researchers have categorized and taxonomized AI harms and their ethical implications based on real-world cases of AI harm (Abercrombie et al., 2024; Hernández et al., 2024; Shelby et al., 2023; Solaiman et al., 2023; Stahl et al., 2022; Weidinger et al., 2021), while other researchers have conducted reviews and meta-analyses of real-world incidents in which AI systems caused different types of harm to humans (Bender et al., 2021; DeVrio, Eslami, & Holstein, 2024; Diberardino, Baleshta, & Stark, 2024; Turri & Dzombak, 2023). In addition to directly causing harms to humans, environmental harms caused by AI systems are also well-documented: the hardware sourcing and training processes required to develop large-scale machine learning models and operate their computing infrastructure is extremely energy-intensive and water-intensive (Li et al., 2023; Luccioni, Jernite, & Strubell, 2024). As a result, the environmental impacts and sustainability of AI systems have recently become an area of significant concern to researchers, practitioners, policymakers, and communities that had their local environments directly impacted by AI development projects (Adarlo, 2023; GPAI, 2021; Lehuedé, 2024; OECD, 2022; Ren & Wierman, 2024; Tessono, 2024). Although the existence of societal and environmental harms caused by AI systems is clear, the reasons why AI governance practices are so frequently ineffective at preventing those harms are relatively unclear.

1.2. Research Scope & Objectives

In this dissertation, I investigate harms caused by AI systems and AI governance practices that are frequently ineffective at preventing those harms. I apply principles and practices of transfeminist ethics to guide my analysis and to identify opportunities for changing future AI governance practices to provide greater benefit and lesser harm, especially for historically marginalized groups. The precise elements of this transfeminist approach to studying the impacts and governance of AI are discussed in more detail in Section 1.4.

This dissertation addresses three overarching inter-related research objectives:

Research Objective 1 – AI Impacts: Determine what types of benefits and harms AI systems are capable of causing, the actors responsible for those benefits and harms, the actors impacted by those benefits and harms, and the activities through which those benefits and harms are caused. This objective is addressed in Chapter 2 (*The Impacts of AI: A Theory and Analysis of AI Value Chains*).

Research Objective 2 – AI Governance Initiatives: Determine what types of AI governance initiatives have been created to intervene in those impacts, and the degree to which those initiatives are effective or ineffective at intervening in those impacts. This objective is addressed in Chapter 3 (*AI Governance Initiatives in Canada*).

Research Objective 3 – AI Governance Systems: Determine how those initiatives function as part of larger AI governance systems that exist across multiple contexts and levels of scale. This objective is addressed in Chapter 4 (*AI Governance Systems in Canada*).

Because AI governance is a complex, dynamic, and systemic phenomenon that exists across many contexts and scales of human activity, much of this dissertation focuses on AI governance activities in Canada from 2017 to 2023 as a means of scoping clear research contexts for data collection and analysis. With the launch of the Pan-Canadian AI Strategy in 2017, Canada was the first country to begin implementing a national AI strategy. Canada's early entry into nationwide AI governance–along with many other political, economic, and cultural factors–

has resulted in Canada developing a relatively more mature and information-rich national AI governance system than most other nations. In Chapters 3 and 4, I describe in greater detail the unique features that make Canada an ideal context for AI governance research, and the limitations to validity and generalizability that arise from focusing on the Canadian context within the timeframe of 2017 to 2023. In addition to contextual limitations, the timeliness of the research presented in this dissertation is limited by the rapid pace at which the field of AI governance changes in response to new data and technological advancements. I mitigate this limitation to the greatest extent possible by specifying timeframes in which data collection and analysis processes occurred, as well as by re-assessing findings presented across each chapter in light of more recent developments in AI technology, policy, and discourse in Chapter 5.

1.3. Structure & Content of Dissertation

This dissertation is structured into five chapters. Chapter 1 (*Introduction*) describes the background and motivation for the research, the scope and objectives of the research, the structure, content, and contributions of the dissertation, and the meta-theoretical framework that was applied to carry out the research presented in this dissertation. Chapters 2, 3, and 4 each present a co-authored article that addresses one of the dissertation’s three research objectives (Objective 1: “Impacts of AI,” Objective 2: “AI Governance Initiatives,” and Objective 3: “AI Governance Systems”). Chapter 5 (*Conclusion: A Future for Transfeminist AI Governance*) synthesizes the findings of each of the three co-authored articles into a set of reflections on the future of AI governance. To enhance the readability of this dissertation, reference lists and appendices from each chapter are provided in two separate sections at the end of the dissertation (*References* and *Appendices*). The remainder of this section introduces each of the three co-authored articles presented in Chapters 2-4 and explains how each article makes an important contribution to researchers, practitioners, and policymakers.

Chapter 2 (*The Impacts of AI: A Theory and Analysis of AI Value Chains*) presents an article that I began co-authoring with David Gray Widder in May 2022. As lead author, I led and contributed to all phases of research ideation and design, data collection, data analysis, and writing and editing of every section of the article. The article was accepted for publication in a future issue of *Big Data & Society* in April 2025. The article was originally titled *The Ethics of AI Value Chains*, but the title has been changed in this dissertation to more clearly reflect the contribution of the article in relation to the three overarching research objectives of this dissertation. Chapter 2 contributes a novel theoretical framework for holistically understanding the types of impacts caused by AI systems and for understanding how those impacts are caused. In Chapter 2, we present an integrative review of research literature and grey literature from the fields of service science, strategic management, economic geography, AI ethics, and AI governance. Based on our review, we argue that *AI value chains* can serve as an integrative concept for theorizing and analyzing the many different types of beneficial and harmful impacts that AI systems are capable of causing to society and the environment. We define AI value chains as “co-creation structures that exist within a network of actors and enable actors to pattern the resource inputs they provide to and the resource outputs they receive from AI systems.” We discuss different types of actors that co-create value with one another by interacting within AI value chains, different types of resources that those actors input into and receive from AI systems, and different types of beneficial and harmful impacts that those actors cause to society and the environment as a result of their resourcing activities. Based on the ethical and practical implications of our integrative review, we identify three directions for future research and practice: conduct more research into the ethical implications of AI value chains, continue developing and applying new theories and methods to analyze AI value chains, and implement a variety of governance mechanisms–including legislation and regulations, industry standards, certification programs, guidance documents, and codes of conduct–to ensure that the resources that flow through AI value chains are ethically sourced.

With the impacts of AI systems and their implications for practice and policy established in Chapter 2, Chapter 3 (*AI Governance Initiatives in Canada*) presents an article that I began co-authoring with Ana Brandusescu and Kelly Lyons in January 2022. As lead author, I led and contributed to all phases of research ideation and design, data collection, data analysis, and writing and editing of every section of the article. The article was published in *Government Information Quarterly* in April 2024. The article was originally titled *The Governance of Artificial Intelligence in Canada: Findings and Opportunities from a Review of 84 AI Governance Initiatives*, but the title has been changed in this dissertation to more clearly reflect the contribution of the article in relation to the three overarching research objectives of this dissertation. Chapter 3 makes theoretical, empirical, and practical contributions to researchers, practitioners, and policymakers with an interest in AI governance. In Chapter 3, we contribute a theoretical framework for understanding how AI governance practices are organized into different types of *initiatives*–including strategic plans, programs, policies, standards, and ethics statements–with the goal of intervening in different types of AI impacts. We then apply this framework to conduct a semi-systematic review and thematic analysis of 84 AI governance initiatives launched by Canada's federal government and the provincial governments of Ontario, Québec, and Alberta from 2017-2022. Based on our review, we identify seven opportunities that researchers and practitioners can act upon to strengthen Canada's AI governance initiatives.

These seven opportunities address issues such as stronger monitoring and transparency of initiative outcomes, securing public trust in AI governance initiatives, ensuring a greater diversity of different types of AI impacts are represented in initiatives, and fostering more inclusive collaboration between organizations, governments, sectors, and civil society.

With the characteristics and opportunities for strengthening Canada's AI governance initiatives established in Chapter 3, Chapter 4 (*AI Governance Systems in Canada*) presents an article that I began co-authoring with Kelly Lyons in January 2022. The article was published in *AI and Ethics* in September 2024. As lead author, I led and contributed to all phases of research ideation and design, data collection, data analysis, and writing and editing of every section of the article. The article was originally titled *AI Governance Systems: A Multi-scale Analysis Framework, Empirical Findings, and Future Directions*, but the title has been changed in this dissertation to more clearly reflect the contribution of the article in relation to the three overarching research objectives of this dissertation. Chapter 4 contributes a novel theoretical framework for conceptualizing and analyzing *AI governance systems* across multiple contexts and levels of scale, such as international, national, subnational, sectoral, and organizational scales. We build upon theoretical perspectives from the fields of AI governance, service science, and organizational studies to argue that an AI governance system is "an interdependent set of components that are (1) situated in a context, (2) structured by the perceptions of "AI" that exist within that context, and (3) intended to maximize benefits and minimize harms that actors within that context perceive as being caused by "AI." We identify 12 main components of AI governance systems, and we apply our theoretical framework to analyze data from interviews with 20 Canadian AI governance initiative leaders and subject matter experts that were conducted throughout 2023. Our analysis of the interviews contributes empirical findings and practical implications regarding the characteristics, strengths, and gaps in Canada's national AI governance system. Based on the findings from our interviews, we recommend three directions for future research and five strategic priorities for strengthening Canada's national AI governance system. Our recommendations address specific topics and methods for extending our research into additional studies and research contexts, as well as practical issues such as guidance for strengthening multistakeholder collaboration and participation throughout Canada's national AI governance system, expanding access to key resources needed for effective AI governance practices, and advancing diversity, equity, and inclusion of marginalized groups throughout Canada's AI governance activities.

In Chapter 5 (*Conclusion: A Future for Transfeminist AI Governance*), I summarize the findings and recommendations presented through Chapters 2-4 and their significance for AI governance research and practice. I conclude the dissertation by reflecting on my experience of conducting this research and on the implications of my findings for the future of AI governance: AI governance is too often built upon harmful power relations, and must be re-imagined through a transfeminist lens that affords a sharper focus on power imbalances and structural inequities, openness and inclusion, collective resistance and contestation, and community empowerment in AI governance systems.

1.4. Meta-theoretical Framework

1.4.1. Overview of Meta-theoretical Framework

In each of the three articles presented in Chapters 2-4 of this dissertation, we develop and apply a different framework of theories and methods to address the dissertation's overarching Research Objectives 1-3. The selection of theories and methods in each of the three articles was influenced by a higher-level framework of meta-theoretical assumptions that I apply across all of my research. These assumptions are described in this section. Table 1.1 contains a summary of the researcher positionality, meta-theoretical assumptions, and research implications that are described in this section.

Table 1.1: Summary of elements of my meta-theoretical framework, related experiences and assumptions, and implications for my research approach.

Elements of framework	Related experiences/assumptions	Implications for my research approach
Researcher positionality	• Nonbinary trans femme researcher and professional who has witnessed and experienced harm and marginalization in AI spaces. • White settler in Canada with material and institutional privileges.	• Transfeminine ways of thinking, feeling, and doing shape my research approach. • Indigenous and racialized perspectives on AI are cited and discussed in my research.
Service realism	• *Service-dominant logic*: Acts of *service* (i.e., applying one's resources for the benefit of another) are foundational to social reality. • Services are co-created through *service systems* consisting of actors, resources, networks, activities, institutional arrangements, and ecosystems.	• AI systems are developed, used, and governed through acts of service. • AI governance is understood as a service system intended to maximize the benefits and minimize the harms of AI development and use.
Transfeminist AI ethics	• *Feminist AI*: AI systems are developed under oppressive social conditions that should be corrected through the application of feminist ethical principles. • *Feminist standpoint epistemology*: Lived experiences of marginalized groups are valid grounds for the production of scientific knowledge and epistemic power. • *Transfeminist ethics*: Lived experiences of trans people provide ethical principles and community-based practices for intervening in harmful systems.	• Transfeminist principles of justice, anti-normativity, fluidity, agency, security, community, solidarity, and resistance should be applied to AI governance. • AI governance should involve trans practices of community-based organizing, contestation, resource-sharing, and expanding political inclusion into stronger arrangements of accountability and justice.

1.4.2. Researcher Positionality

Many philosophies of science encourage researchers to be as self-reflexive as possible in accounting for how our lived experiences, ontological assumptions, epistemological assumptions, and ethical assumptions influence our research approach (for a detailed review of such philosophies, see Holmes, 2020). Feminist theorists in particular have long held that self-reflexivity, researcher positioning, and recognition of researcher partiality are especially important practices for ensuring the rigor, trustworthiness, and validity of research (Haraway, 1988; Harding, 1992, 1995). Therefore, in this sub-section, I describe the lived experiences that most significantly influence my research approach.

My research is influenced by my lived experience as a nonbinary trans woman who has worked as a research and management professional on AI development, digital transformation, policy research, and information management projects in public sector, private sector, and post-secondary organizations. I have also worked on academic research projects at the interdisciplinary intersection of "information studies," where perspectives from service sciences, cognitive sciences, information systems design, knowledge and information management, policy studies, queer and trans studies, and feminist science and technology studies have all influenced my thoughts, feelings, intuitions, and actions in approaching my research. Over the course of working and studying in cis-hetero-normative and male-dominated social spaces as a trans femme, I have become greatly concerned with patterns of harm and marginalization that I have witnessed and personally experienced within academic and professional environments. Though I have been subjected to some forms of anti-trans harm and marginalization in my professional and personal life, my experience living in Canadian society as a white settler has afforded me material and institutional advantages over the course of my life that many Indigenous and racialized persons do not benefit from. Knowing that Canadian society and Canada's governance systems have historically been structured to uphold those relations of advantage and disadvantage, I endeavor to always remain mindful of how those relations might apply to the AI systems and governance systems I study. Knowing that Indigenous perspectives have historically been marginalized in Canadian education systems and in the governance systems of the Canadian state, I also endeavor to educate myself on Indigenous perspectives on AI ethics and AI governance as part of my research process, and include and cite Indigenous perspectives in my analysis of AI ethics and governance issues.

As a consequence of my lived experience, I consciously and unconsciously adopt meta-theoretical assumptions about ontology, epistemology, and ethics when selecting research topics to investigate and when selecting theories and methods to apply to my research. My research is influenced by two main sets of meta-theoretical assumptions that I refer to as *service realism* and *transfeminist AI ethics*, described in the following sub-sections.

1.4.3. Service Realism

What I refer to as *service realist* assumptions about ontology and epistemology are derived from the interdisciplinary field of service science, engineering, management, and design (SSMED) (Spohrer & Kwan, 2010). Service realism provides a lens for perceiving social structures and social activities, such as political and economic actors, technologies, exchanges of value, organizational practices, and institutions. The ontological and epistemological assumptions of service realism are based on a foundational concept in the SSMED literature known as *service-dominant (S-D) logic*. Originally coined by Vargo and Lusch (2004), S-D logic refers to a set of ontological and epistemological assumptions that regard *service* as the ontological bedrock of all economic activity, with service defined in an influential paper by

Maglio et al. (2009) in extremely general terms as “the application of the resources of one or more systems for the benefit of another system in economic exchange” (p. 405). This service realist “worldview”–as Maglio et al. characterize the ontological and epistemological assumptions of S-D logic–stands in contrast to the *goods-dominant logic* of mainstream economics, which assumes that economic reality consists primarily of tangible goods that are produced or consumed by atomized, individual actors in response to supply and demand signals.

Earlier theories and applications of S-D logic tended to focus on the role of economic and organizational activities in the design and delivery of services, such as processes of value co-creation between economic actors, integration of tangible and intangible resources from across service systems and networks, and measurement and evaluation of service outcomes (see for example Alter, 2008; Katzan, 2009; Lyons & Tracy, 2013; Maglio et al., 2009; Maglio & Spohrer, 2007). More recently, theories and applications of S-D logic have evolved to recognize the importance of a greater range of social, political, and ecological activities in the design and delivery of services (see for example Akaka, Koskela, & Vargo, 2019; Frost, Cheng, & Lyons, 2019; Ng et al., 2018; Siltaloppi & Wieland, 2018; Vargo & Lusch, 2017; Vargo & Lusch, 2016). Through these more recent developments in the SSMED literature, S-D logic has subsumed a variety of social, political, and economic theories into its axiomatic assumptions, such as the actor-network theory of Latour (2005), institutionalist theories from sociology, political science, and economics (reviewed in detail and integrated into S-D logic by Vargo & Lusch, 2016), as well as foundational principles of ecological economics such as environmental interdependence and sustainability (Common & Stagl, 2012). These theoretical developments have resulted in S-D logic evolving into a more generalizable worldview: a broad set of ontological and epistemological assumptions not just about economic reality, but about all of social reality, including the interaction of social, economic, political, organizational, cultural, technological, material, and ecological phenomena in the pursuit of beneficial service outcomes between various social actors. The analysis framework developed by Frost, Cheng, and Lyons (2019) is particularly notable for building upon two systematic reviews of the SSMED literature (Frost & Lyons, 2017; Lyons & Tracy, 2013) in order to describe a richly integrated set of social, political, economic, and ecological service system components (see Figure 1.1).

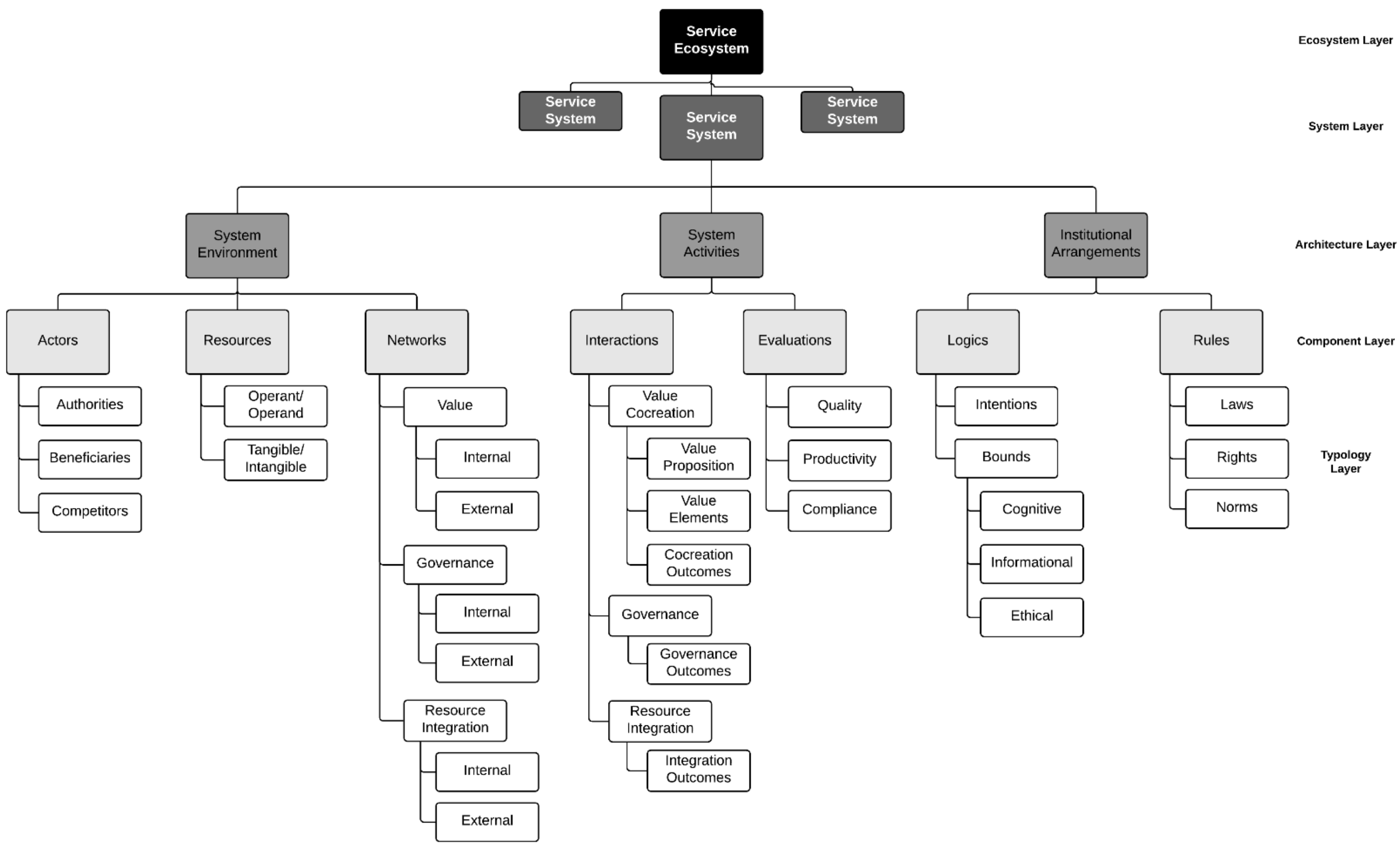

Figure 1.1: The multilayer service system analysis framework developed and illustrated by Frost, Cheng, and Lyons (2019).

Despite the evolution of S-D logic toward a more generalizable worldview, there remain some gaps in S-D logic's ability to account for social reality. Three theoretical gaps are of particular note in applying service realist assumptions to study the impacts of AI, AI governance initiatives, and AI governance systems: gaps in accounting for *harm*, *power*, and *history*.

(1) *Harm:* The capacity for AI systems to cause harm–such as social, political, economic, material, psychological, and environmental harm–is a matter of great concern to AI governance researchers and practitioners. However, the SSMED literature has so far shown little interest in studying the potential for service systems to generate harmful outcomes, and consequently, S-D logic lacks a rigorous theoretical account of harm. Applications of S-D logic focus greatly on possibilities of benefit, beneficial activity, beneficial or win-win outcomes, and the beneficiaries of services (Maglio et al., 2009; Vargo & Lusch, 2016; Vargo & Lusch, 2004), but afford little attention to possibilities of harm, harmful activity, harmful outcomes or loss situations, or the actors harmed by service systems. Some notable exceptions include Spohrer's (2015) typology of value co-creation outcomes, loss-lead outcomes, coercion outcomes, and value co-destruction outcomes, as well as Frost, Cheng, and Lyons (2019) identifying "competitors" as "the actors who are disadvantaged by a given value cocreation interaction, though they are not directly involved in the interaction themselves" (p. 292). Overall though, strong theoretical accounts of harm are greatly lacking in the SSMED literature.

(2) *Power:* Because the SSMED literature lacks strong theoretical accounts of the causes and effects of harm, S-D logic is limited in its ability to analyze the impact that power structures and power relations have on service processes and outcomes. Siltaloppi and Wieland (2018) observe this gap in S-D logic, commenting that "more work is required to fully define power and explain its effect on resource integration, service exchange and value determination in service ecosystems" (p. 12). Siltaloppi and Wieland and many other researchers have identified institutional structures and institutional changes as sources of power dynamics in service systems (Frost, Cheng, & Lyons, 2019; Jaakkola, Aarikka-Stenroos, & Ritala, 2019; Vargo & Lusch, 2016). However, S-D logic still lacks a strong theory of how power is distributed and

redistributed through institutions and institutional change. This poses an obstacle for applying S-D logic to studies of AI governance, as AI systems are well-known to cause changes in and be affected by many dynamics of political power and economic power throughout their lifecycles.

(3) *History:* Applications of S-D logic typically approach the historical context of service through the lens of what Katzan (2009) calls "service temporality," an ontological view of time as a measurable object that can be divided into discrete units known as "service events," with each event having manageable qualities, durations, and triggers. Rarely in the SSMED literature is historical context treated as a matter of *history* rather than as a matter of mere temporality. The distinction between history and temporality is ontologically and epistemologically significant: rather than the managerialist assumptions of service temporality, the *history* of a service system assumes the existence of a continuously unfolding temporal process in which the system's components–including its actors, resources, networks, value co-creation activities, and institutions–all constantly change through iterative activity. Akaka and Perry (2019) explain that historical context "draws attention to how value emerges through the development and evolution of a service system" (p. 467). Their view of historical context affords history with phenomenological, ethical, and political dimensions that are not shared by mere temporality: to describe a history requires a historiographical approach to identifying and recounting events, an approach that is shaped by the positionalities, perspectives, and perceptions of value assumed by those who describe the history. The historical context of a service system is therefore linked to issues of harm and power, and is extremely important for studies and practices of AI governance, as the harms that AI systems cause to historically marginalized groups must be accounted for.

The meta-theoretical assumptions of transfeminist AI ethics that I describe in the following sub-section are capable of filling service realism's gaps in accounting for harm, power, and history. Additionally, when applying service realist theories to study AI systems and AI governance systems in Chapters 2 and 4 of this dissertation, other bodies of literature are reviewed alongside the SSMED literature in order to identify context-specific ontological and epistemological gaps in S-D logic, as well as supplementary theories and methods that can fill those gaps.

1.4.4. Transfeminist AI Ethics

Feminist ethics describes a diverse set of theories of applied and normative ethics that share a common orientation toward resisting the devaluation of femininity and embracing historically feminized values such as care, relationality, interdependence, and maintenance (de la Bellacasa, 2017; Keller & Kittay, 2017; Mattern, 2018; Moriggi et al., 2020). Shade (2023) observes that a wide range of feminist ethical principles—such as principles of care, relationality, pluralism, privacy, consent, embodiment, access, participation, non-discrimination, and equity—have shaped the evolution of early 2000s media reform movements into the intersectional digital and data justice movements of the late 2010s and 2020s. As part of that evolution, *feminist AI* has emerged as an approach to AI research and practice that applies feminist ethical principles to the development and use of AI systems. Proponents of feminist AI recognize that AI systems are often developed and used under oppressive social conditions. Proponents therefore advocate for enacting feminist principles in AI systems to prevent further harms and promote more beneficial outcomes for oppressed and vulnerable groups (A+ Alliance, 2021; Toupin, 2024; Varon & Peña, 2021; Wellner & Rothman, 2020). Feminist AI shares much in common with other recent applications of feminist ethics to socio-technical practices, particularly within the area of feminist data ethics and data justice (D'Ignazio & Klein, 2020; Gray & Witt, 2021; Garcia et al., 2020; Marčetić & Nolin, 2022). The ethics of "data feminism" as described by D'Ignazio and Klein is especially notable for holistically applying feminist ethical principles to scrutinize power asymmetries in data science, in AI and data technologies, and in their attendant digital infrastructures. Data feminism advances an approach to AI and data ethics that is grounded in examination of and resistance to inequitable power structures, in knowledges that are pluralistic, embodied, and contextual, and in socio-technical practices that call attention to labor and challenge oppressive AI and data systems.

Feminist approaches to ethics entail ontological, epistemological, and political assumptions about how and why we should acquire scientific knowledge of reality. As imagined by feminist scholars such as Haraway (1988), Harding (1992, 1995), and Suchman (2007), *feminist standpoint epistemologies* understand scientific knowledge as situated in social and historical contexts, embodied in social agents with partial perspectives of their world, and produced pluralistically across many communities with differing norms, values, beliefs, abilities, and lived experiences. Feminist standpoint epistemologies contrast with traditional empiricist and positivist epistemologies that understand scientific knowledge as disembodied, disembedded from social life and historical context, value-neutral and politically impartial, and produced by individual scientists in accordance with a universally shared standard of rigor and objectivity. Harding (1992) explains that feminist scientific standpoints have the potential to offer greater rigor and objectivity than traditional positivist standpoints by "critically identifying all of those broad, historical social desires, interests, and values that have shaped the agendas, contents, and results of the sciences much as they shape the rest of human affairs" (p. 359). A feminist standpoint can thereby privilege historically marginalized perspectives to produce knowledge that "can be *for* marginalized people . . . rather than *for* the use only of dominant groups in their projects of administering and managing the lives of marginalized people" (p. 445). For many

Black and intersectional feminists, standpoint epistemologies provide ground for producing collective knowledge and building epistemic power from marginalized experiences that cross intersections of race, gender, and class (Collins, 2000; Crenshaw, 1991; Dotson, 2014).

Researchers have recently applied feminist epistemologies to study AI systems that cause harm to marginalized groups, as well as to recommend actions for empowering those groups with greater agency in the design and governance of those systems (Benjamin, 2019; Birhane, 2021; Hancox-Li & Kumar, 2021; Kong, 2022; McQuillan, 2022; Ricaurte, 2022; Schelenz, 2022; Widder, 2024). These studies demonstrate that feminist standpoints are effective for analyzing the social and historical contexts of AI governance activities, and for identifying how power relations in AI governance activities can be changed to prevent harms to marginalized groups. This makes feminist approaches to ontology, epistemology, and ethics ideal for filling service realism's theoretical gaps in accounting for issues of harm, power, and history.

Trans researchers and activists extend feminist theories of ontology, epistemology, and ethics into *transfeminist ethics*, theories of applied ethics that are ontologically and epistemologically grounded in the lived experiences of trans women, trans femmes, and other transgender and gender-diverse people who self-identify with the umbrella term "trans." Transfeminist ethics therefore centers ethical principles that are of particular importance to trans experiences of existing within cis-hetero-normative social systems, such as the importance of transformative change, anti-normativity, fluidity, agency, security, community, care, solidarity, and resistance against governance systems that perpetuate anti-trans violence (Barad, 2015; Galpin et al., 2023; Malatino, 2020; Marvin, 2019; McFadden et al., 2024; Spade, 2015; van der Drift & Raha, 2020).

In their analysis of the political implications of transfeminist principles, van der Drift and Raha (2020) argue for adopting transfeminist practices of ethics and governance that are “active and anti-normative, rather than defined in a stable form,” recognizing that “the dynamism of the term ‘trans’ indicates that we must attend to questions of agency and structures of action” (p. 13). Transfeminist ethics therefore demands not just theoretical commitments, but also practical commitments to particular actions and political goals. Transfeminist ethics seeks to protect the fluidity and vitality of trans lives against various forms of anti-trans violence, including physical, psychological, economic, and administrative violence. Transfeminist ethics seeks to cultivate collective resistance against dominant social structures and norms that cause harm to trans communities, to radically change those harmful structures and norms, and to secure greater agency, solidarity, and power for trans communities and other marginalized communities. Accordingly, trans practices of governance are often grounded in community-led design and collective organizing of direct action against governance systems that cause harm to trans people and other marginalized groups (Costanza-Chock, 2020; Nownes, 2019; Verloo & van der Vleuten, 2020). Trans governance practices involve polycentric capacity-building, resource-sharing, and mutual care arrangements within localized networks of trans people, as well as co-creating policy within communities and with sympathetic actors in powerful institutions to transform narrow relations of trans political inclusion into stronger relations of security, accountability, and justice for trans people (Davidson, 2007; Malatino, 2020, 2022; McFadden et al., 2004; van der Drift & Raha, 2020; Verloo & van der Vleuten, 2020).

In my research, I apply transfeminist ethical principles and practices through an approach I refer to as *transfeminist AI ethics.* I define transfeminist AI ethics as *the application of transfeminist ethical principles to the study and practice of AI governance*. Researchers have found that AI systems put trans people at risk of experiencing many types of physical, psychological, social, and economic harms, such as harms caused by trans-exclusionary healthcare automation, recruitment and hiring automation, facial and gender recognition, security and law enforcement applications, and generative AI applications (Costanza-Chock, 2018; Keyes, 2018; Scheuerman, Pape, & Hanna, 2021; Scheuerman, Paul, & Brubaker, 2019; Ungless, Ross, & Lauscher, 2023). However, rather than studying the ethical issues arising from those and other specific anti-trans use cases of AI systems, I apply principles and practices of transfeminist ethics to study AI systems and their governance through a transfeminist lens more generally. My approach to transfeminist AI ethics scaffolds upon my assumptions of service realism to center a broad scope of ethical issues and governance interventions, all of which are of particular importance to the protection and flourishing of trans communities and other marginalized communities impacted by AI systems. There are six main ethical issues and governance interventions that I center in my research:

(1) *Socio-political & socio-economic conditions*: Conditions of political and economic inclusion, exclusion, injustice, and accountability in AI systems and their governance.

(2) *Inequitable value co-creation*: The co-creation of inequitable AI systems and governance systems through networks of dominant actors, marginalized actors, resource dependencies, and institutional arrangements.

(3) *Norms & values of dominant actors*: Exclusionary norms and values underlying state-led and industry-led AI governance systems.

(4) *Harms of dominant actors*: Harmful outcomes of AI governance actions and inactions taken by dominant state actors and industry actors.

(5) *Opportunities for top-down change*: Practical opportunities for preventing harm and securing greater power for marginalized communities by changing state-led and industry-led AI governance systems.

(6) *Opportunities for bottom-up change*: Practical opportunities for preventing harm and securing greater power for marginalized communities by developing alternative AI governance systems led by communities, workers, and civil society organizations.

As we will see in the following chapter, AI systems have a multitude of impacts on society and the environment. Because of the global scale, high frequency, and multi-contextual complexity of those impacts, the emerging field of AI governance covers such a vast space of ethical issues and governance interventions that it is impossible to comprehensively cover every conceivable issue and intervention within a scope of a single study (or even within the scope of a single dissertation). Transfeminist AI ethics therefore provides a meta-theoretical heuristic for orienting my analysis of large-scale phenomena–such as global AI value chains (Chapter 2), federal and provincial AI governance initiatives (Chapter 3), and a national AI governance system (Chapter 4)–toward the ethical issues and practical opportunities I outline above.

1.5. Contributions of Dissertation

Each chapter of this dissertation makes novel contributions to researchers, practitioners and policymakers. The contributions of Chapters 2, 3, and 4 are described in detail in the concluding sections of each of those chapters. Table 1.2 provides a summary of each chapter's contributions to research, practice, and policy.

Table 1.2: *Summary of this dissertation's contributions to research, practice, and policy*.

Chapter	**Contributions to research**	**Contributions to practice & policy**
Chapter 1 (*Introduction*)	• A meta-theoretical framework that can be applied to theorize and analyze AI governance activities in any context.	• A meta-theoretical framework that can be applied to guide AI governance practices and policymaking in any context.
Chapter 2 (*The Impacts of AI*)	• A novel theoretical framework of AI value chains. • Framework is applied to a literature review that (1) integrates a diverse range of AI impacts and ethical concerns, and (2) clarifies the ontological, ethical, practical, and policy implications of AI value chains.	• Guidance for (1) developing ethical sourcing standards to prevent harmful impacts of AI systems across their lifecycles and value chains, and (2) enforcing those standards through international and national policy.
Chapter 3 (*AI Governance Initiatives in Canada*)	• A framework of theories and methods for national-scale analysis of AI governance initiatives. • Novel empirical data about Canadian AI governance initiatives; analysis of that data. • Data, findings, and framework are transferable to other studies of AI governance initiatives.	• Guidance for stronger monitoring of AI governance outcomes, public trust in AI, intervention across a diversity of AI impacts, and coordination between governments, sectors, and • Recommendations are transferable to other AI governance contexts facing similar challenges.
Chapter 4 (*AI Governance Systems in Canada*)	• A framework of theories and methods for macro-scale analysis of AI governance systems. • Novel empirical data about components of Canada's national AI governance system; analysis of that data. • Data, findings, and framework are transferable to other studies of AI governance systems.	• Strategic objectives, practices, and policy interventions for addressing challenges of collaboration, coordination, participation, access to resources, diversity, equity, and inclusion in a national AI governance system. • Recommendations are transferable to other AI governance contexts facing similar challenges.
Chapter 5 (*Conclusion*)	• The application of a transfeminist lens to provide a novel synthesis of AI governance research and practice.	• Guidance for enacting AI governance based on transfeminist ethical principles to researchers, practitioners, public servants, journalists, creatives, and community organizers.

Chapter 2

The Impacts of AI:
A Theory and Analysis of AI Value Chains

About this Article: This article was co-authored with David Gray Widder beginning in May 2022 and was originally titled *The Ethics of AI Value Chains*. The article was accepted for publication in a future issue of *Big Data & Society* in April 2025.

Abstract

Researchers, practitioners, and policymakers with an interest in AI ethics need more integrative approaches for studying and intervening in AI systems across many contexts and scales of activity. This paper presents *AI value chains* as an integrative concept that satisfies that need. To more clearly theorize AI value chains and conceptually distinguish them from supply chains, we review theories of value chains and AI value chains from the strategic management, service science, economic geography, industry, government, and applied research literature. We then conduct an integrative review of a sample of 67 sources that cover the ethical concerns implicated in AI value chains. Building upon the findings of our integrative review, we recommend three future directions that researchers, practitioners, and policymakers can take to advance more ethical practices across AI value chains. We urge AI ethics researchers and practitioners to move toward value chain perspectives that situate actors in context, account for the many types of resources involved in co-creating AI systems, and integrate a wider range of ethical concerns across contexts and scales.

Keywords: Artificial intelligence, AI ethics, value chains, supply chains, governance, policy

2.1. Introduction

AI ethics is a field of study and practice seeking values, principles, and methods for guiding the implementation of artificial intelligence (AI) systems. Principles and practices of ethical AI often fail to prevent many societal and environmental harms (Attard-Frost, De los Ríos, & Walters, 2023; Greene, Hoffman, & Stark, 2019; Hagendorff, 2020; Lauer, 2021; Morley et al., 2023; Rességuier & Rodrigues, 2020). In response, many researchers have called for AI ethics to be re-centered around new principles or conceptual focal points such as participatory design practices (Birhane et al., 2022a; Bondi et al., 2021), organizational practices (Attard-Frost, De los Ríos, & Walters, 2023; Mäntymäki et al., 2022; Schneider et al., 2023), or relational structures (Birhane, 2021; Crawford, 2021; Crawford & Joler, 2018). Many have also called for new principles and practices of ethical AI based on Ubuntu and Indigenous value systems (Gwagwa, Kazim, & Hilliard, 2022; Lewis et al., 2018; Lewis et al., 2020; Mhlambi, 2020).

In parallel with these developments in AI ethics, policymakers are taking a strong interest in the value chains required to provide resource inputs into and receive resource outputs from AI systems. Emerging regulatory frameworks in the European Union (European Commission, 2018, 2021; European Parliament, 2023) and Canada (Minister of Innovation, Science and Industry, 2023; Parliament of Canada, 2022) aim to set obligations on actors within the “supply chain” or “value chain” of AI systems. Meanwhile, research literature has emerged that analyzes the policy implications of “AI supply chains” or “AI value chains” (Brown, 2023; Cobbe, Veale, & Singh, 2023; Engler & Renda, 2022; Kak & West, 2023; Lee, Cooper, & Grimmelmann, 2023; Widder & Nafus, 2022, 2023; Widder & Wong, 2023). However, the emerging regulatory frameworks and policy research literature on AI supply chains/value chains lack a strong theory of AI supply chains/value chains. AI regulatory frameworks and AI researchers often use the terms "supply chain" and "value chain" as though they are interchangeable, when in fact, supply chains and value chains are different types of structures with different ontological, ethical, practical, and policy implications.

In this paper, we present an integrative approach to AI ethics that foregrounds the value chains involved in providing resource inputs to and receiving resource outputs from AI systems. Our study of the *ethics of AI value chains* aims to accomplish two objectives:

Research Objective 1 – *Integration of ethical concerns*: We aim to overcome the limitations of many current approaches to AI ethics by integrating a wide range of ethical concerns across many actors, resources, contexts, and scales of activity.

Research Objective 2 – *Clarification of value chain implications*: We aim to better theorize and clarify the ontological, ethical, practical, and policy implications of AI value chains.

To accomplish those two objectives, we first review theories of value chains and AI value chains in Section 2.2. In Section 2.3, we describe our methodology for an integrative review of literature on the ethical implications of AI value chains. In Section 2.4, we present the findings of our integrative review. In Section 2.5, we acknowledge the limitations of our review. In Section 2.6, we recommend future directions for researchers and practitioners with an interest in the ethics of AI value chains. We conclude in Section 2.7 by highlighting the contributions of our review.

2.2. Theory

2.2.1. Value Chains

In the strategic management literature, the first in-depth theorization of value chains was Porter's (1985) "value chain model." Porter's value chain model specifies five "primary activities" (inbound logistics, outbound logistics, operations, marketing & sales, and service) and four "support activities" (firm infrastructure, human resource management, technology development, and procurement), with each activity transforming resource inputs into valuable outputs and gradually moving resources further downstream in a linear chain-like structure. Later theories in the economic geography literature apply the value chain concept to contexts beyond Porter's predefined "primary" and "support" activities, accounting for the role of value chains in more complex organizational systems and economic networks such as *global value chains* (Gereffi, Humphrey, & Sturgeon, 2006; Humphrey & Schmitz, 2000; Kano, Tsang, & Yeung, 2020) and *global production networks* (Coe, Dicken, & Hess, 2008; Coe & Yeung, 2019; Henderson et al., 2001). More recently, researchers have further extended those theories to critically study the political economies and economic geographies of digital platforms and artificial neural network production (Butollo et al., 2022; Butollo & Schneidemesser, 2022; Ferrari, 2023; Howson et al., 2022a, 2022b).

Researchers in the field of service science, management, engineering, and design (SSMED) have also developed theories of value chains. SSMED researchers conceptualize value chains as linear structures through which value is co-created and progressively added to the chain by a series of actors who exist in diverse service contexts. In these value chains, *value* is conceptualized as phenomenologically co-created preferences for action (Frost, Cheng, & Lyons, 2019), rather than as a positivistic, quantifiable, priceable, and objectively measurable phenomenon as value is generally conceptualized in mainstream economics (Spash 2012). In addition to value chains, SSMED researchers theorize *value networks* as interactive structures that enable value to be co-created between many interdependent actors who are situated across contexts, spaces, times, positionalities, and scales of activity (Edvardsson, Skålén, & Tronvoll, 2015; Frost, Cheng, & Lyons, 2019; Lusch, Vargo, & Tanniru, 2010; Vargo & Lusch, 2016). Foundational to value network ontologies are *resourcing activities*, the activities through which multiple actors across the network assemble and integrate their resources with the goal of co-creating value.

While some regard value *network* ontologies as a conceptually stronger successor to value *chain* ontologies (Basole, 2019; Buhman et al., 2005; Dyer, 2000), others see them as highly compatible. Compatibilist theories view value chains as value network sub-structures through which a set of dyadic actor-actor pairings integrate some of their resourcing activities spatially, temporally, as well as vertically (within a particular industry) and horizontally (across multiple industries) (Alter, 2008; Chen & Chiu, 2015; Lim et al., 2018). For example, Alter's "service value chain framework" assumes that value chains enable linear sets of resourcing activities to be "continuously or repeatedly" (p. 76) performed within pre-negotiated service delivery workflows. Similarly, the "data-value chain" model of Lim et al. characterizes data as a resource from which many networked actors can gradually co-create value through multiple linear chains of data collection, data analysis, and information use activities that are situated across many service contexts. These compatibilist theories of value chains and networks reflect an earlier understanding in the economic geography literature: within network configurations such as global production networks, "there is inevitably an element of linearity or verticality in the structure of its nodes and links" (Coe, Dicken, & Hess, 2008, p. 274). In this view, value chains are linear co-creation structures embedded within larger, non-linear networks of production, distribution, and consumption.

Building upon compatibilist theories of value chains and value networks, we define *value chains* as *co-creation structures that exist within a network of actors and enable patterned resourcing activities to occur between actors*. The resourcing activities that occur within value chains have three main properties:

(1) *Situatedness*: Resourcing activities are situated within specific contexts.

(2) *Pattern*: Resourcing activities are spatially, temporally, and organizationally patterned, and thus capable of recurring with some degree of regularity.

(3) *Value co-creativity:* Resourcing activities are co-created by, perceived differently by, and valued differently by many interdependent actors.

These three properties make the ontologies of value chains markedly different and relatively less linear than the ontologies of *supply chains*, in which resourcing is typically understood as a series of one-way, downstream movements from producers to consumers.

2.2.2. Supply Chains vs. Value Chains

Value chains have different properties than *supply chains* (Feller, Shunk, & Callarman, 2006). Supply chains are organized according to a "goods-dominant logic," characterized by Vargo and Lusch (2006) as an outdated logic of economic organization in which "tangible output and discrete transactions were central" (p. 4). In contrast, value chains are organized according to a "service-dominant logic" in which "intangibility, exchange processes, and relationships are central" (p. 4). While supply chain ontologies account for a linear set of activities needed to provide tangible resource inputs to production processes (ending in the consumption of those resources), the value chain ontologies of SSMED account for a broader range of intangible and tangible resourcing activities, upstream and downstream relations, and co-creative processes that are simultaneously productive and consumptive.

In the AI ethics literature, the ontological distinction between value chains and supply chains has recently been re-affirmed by Widder and Nafus (2022, 2023). Widder and Nafus call for AI development practices to move away from the task modularity and linearity inherent to supply chain ontologies and toward the broader forms of co-creativity and relationality assumed by value chain ontologies. In the contexts of AI systems, the difference between supply chain and value chain ontologies is crucial: while the supply chains of AI systems scope off a particular set of linear tasks required to make a system usable and make its outputs consumable (thereby ending the supply chain at the "end user" or "consumer"), value chains extend the scope of the system's ethical, practical, and policy considerations into a broader network of co-creative relations. For example, both ontological perspectives can account for the flow of data resources downstream from data subjects to data owners and brokers, model developers, application developers, and end users. However, only value chain ontologies can additionally account for simultaneous upstream flows of financial resources, information resources, and knowledge. A value chain ontology is also more capable of accounting for the production and consumption of material resources, such as the energy and water required to train the model and operate its data infrastructure, or the minerals and fuel required to build and transport the system's hardware components. These materials are omitted from the scope of relatively narrow "AI supply chain", "algorithmic supply chain", and "data supply chain" ontologies that are primarily focused on the downstream flow of data resources (Brown, 2023; Cobbe, Veale, & Singh, 2023; Lee, Cooper, & Grimmelmann, 2023). Figure 2.1 illustrates this difference in perspective between a hypothetical supply chain ontology of AI systems (focused on vertical production-consumption relations as resources move downstream) and a hypothetical value chain ontology of AI systems (focused on co-creation relations as resources move downstream, upstream, and horizontally through a larger value network).

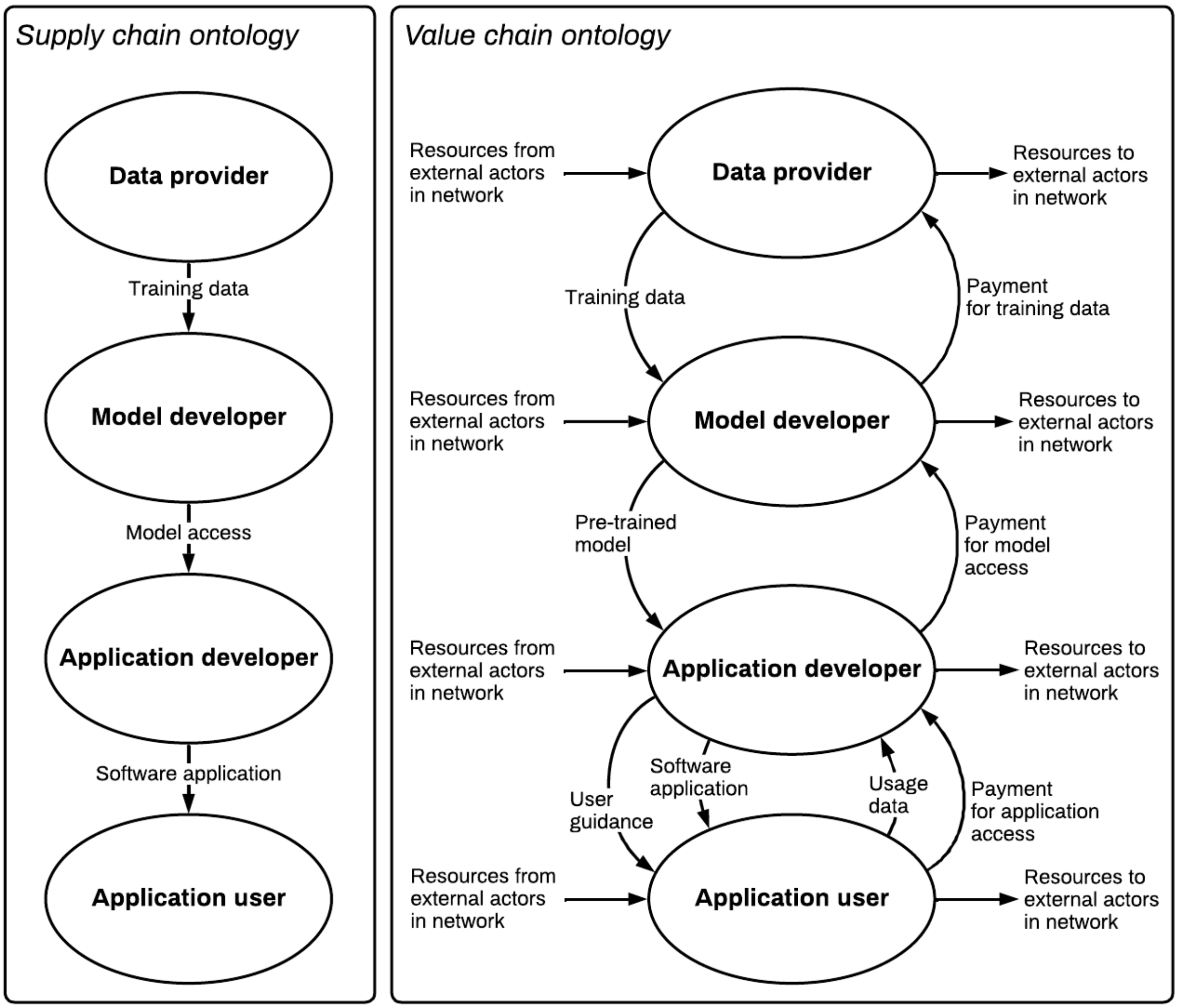

Figure 2.1: A supply chain ontology of AI systems contrasted with a value chain ontology of AI systems. The actors and resources that appear here are intended to illustrate key differences between these two perspectives, not to provide a wholly representative or exhaustive view of the actors and resources involved in AI systems.

2.2.3. AI Value Chains

A growing body of recent industry, government, and applied research literature examines the value chains of AI systems. Applying our theory of value chains from the previous sub-sections, we define *AI value chains* as *co-creation structures that exist within a network of actors and enable actors to pattern the resource inputs they provide to and the resource outputs they receive from AI systems.*

Much of the applied research literature on AI value chains comes from a strategic management or industrial engineering perspective, examining the role of AI systems in adding value or risk to pre-existing industrial value chains (Chan-Olmsted, 2019; Liu, Chen, & Chen, 2022; Oosthuizen et al., 2020; Staubli, 2022). However, some researchers directly study the ethical and policy implications of providing resource inputs to and/or receiving resource outputs from AI systems. Engler and Renda (2022) propose a typology of AI value chains, common resourcing activities involved in AI value chains, and recommendations for EU policymakers seeking to set more specific obligations for AI value chain participants. Other policy researchers have examined how responsibility and accountability is distributed throughout the value chains involved in supplying data resources to AI systems (Brown, 2023; Cobbe, Veale, & Singh, 2023; Kak & West, 2023; Lee, Cooper, & Grimmelmann, 2023). These studies primarily focus on the software resourcing activities involved in AI systems (e.g., the preparation and use of training and testing data, the purchasing and use of compute, the development and use of models, algorithms, code, and APIs) and propose policy interventions that target those software resourcing activities. Widder and Nafus (2022, 2023) combine theories from computer science and feminist science and technology studies to take a more critical approach to the ontologies and ethics of AI value chains. In examining the practices of 27 AI engineers, they describe AI value chains as "heterogenous, cross-cutting, not always linear social interactions and relations that occupy multiple social locations and cultural logics at the same time" (2022, p. 3). Reflecting SSMED theories of value chains, Widder and Nafus emphasize that AI value chains are situated across social, political, and economic contexts with varying patterns of resource distribution and diverse perceptions of developer responsibility.

Alongside the applied research literature, perspectives on AI value chains from industry represent another emerging body of literature. Similarly to the applied research literature, industry perspectives on AI value chains are predominantly interested in how AI systems can add value to pre-existing industrial value chains by increasing efficiency, effectiveness, or productivity (Appen, 2021; Fife, 2022; Härlin et al., 2023; Shaw & Arkan, 2019). When industry perspectives do discuss the ethics of AI value chains, claims about "responsible AI" or "ethical" AI practices center on the software resources required to develop and use AI systems (e.g., datasets, models, compute, APIs) rather than the social, political, economic, and ecological contexts in which the software resources and resourcing activities are situated.

Many governments take a broader perspective on AI value chains than industry, as governments often aim to intervene in a larger set of societal and environmental impacts than industry is typically concerned with. For example, amendments to the EU's *AI Act* adopted by the European Parliament in June 2023 impose new legal obligations on several value chain actors for conducting data resourcing activities, open-source AI development activities, development and use of "general-purpose AI systems" and "generative foundation models", and environmental impact mitigation activities. Seeking alignment with the EU regulatory framework, amendments to Canada's proposed *Artificial Intelligence and Data Act* (Minister of Innovation, Science and Industry, 2023; Parliament of Canada, 2022) also aim to impose legal obligations throughout Canada's "AI value chain." However, the Canadian framework has not yet set requirements on as broad a range of data resourcing, software resourcing, model development, and environmental impact mitigation activities as the EU framework has. Notably, both the EU and Canadian frameworks neglect to set requirements on the hardware resources involved in developing and using AI systems. This omission indicates that both regulatory frameworks are built upon an incomplete theory of AI value chains that privileges the socio-technical contexts, spatial/temporal/organizational patterns, and value co-creation interactions that are involved in AI *software* lifecycles. The contexts, patterns, and value co-created throughout AI *hardware* lifecycles–along with many other AI value chain actors and resourcing activities–are absent from the incomplete theoretical assumptions underlying these regulatory frameworks.

In addition to software resources and hardware resources, many other types of resources are implicated in AI systems, such as financial resources, knowledge resources, labor and human resources, and governance resources. These resources are input to and output from AI systems by a multitude of actors as they perform various interconnected activities throughout the system lifecycle, such as design, development, deployment, and operation. Ultimately, these actors and their resourcing activities cause many different types of beneficial and harmful impacts to themselves and/or to other actors. Figure 2.2 illustrates some examples of the actors, resources, activities, and impacts that exist across the value chains of AI systems. The contents of Figure 2.2 are discussed in more detail as part of our integrative literature review in Section 2.4.

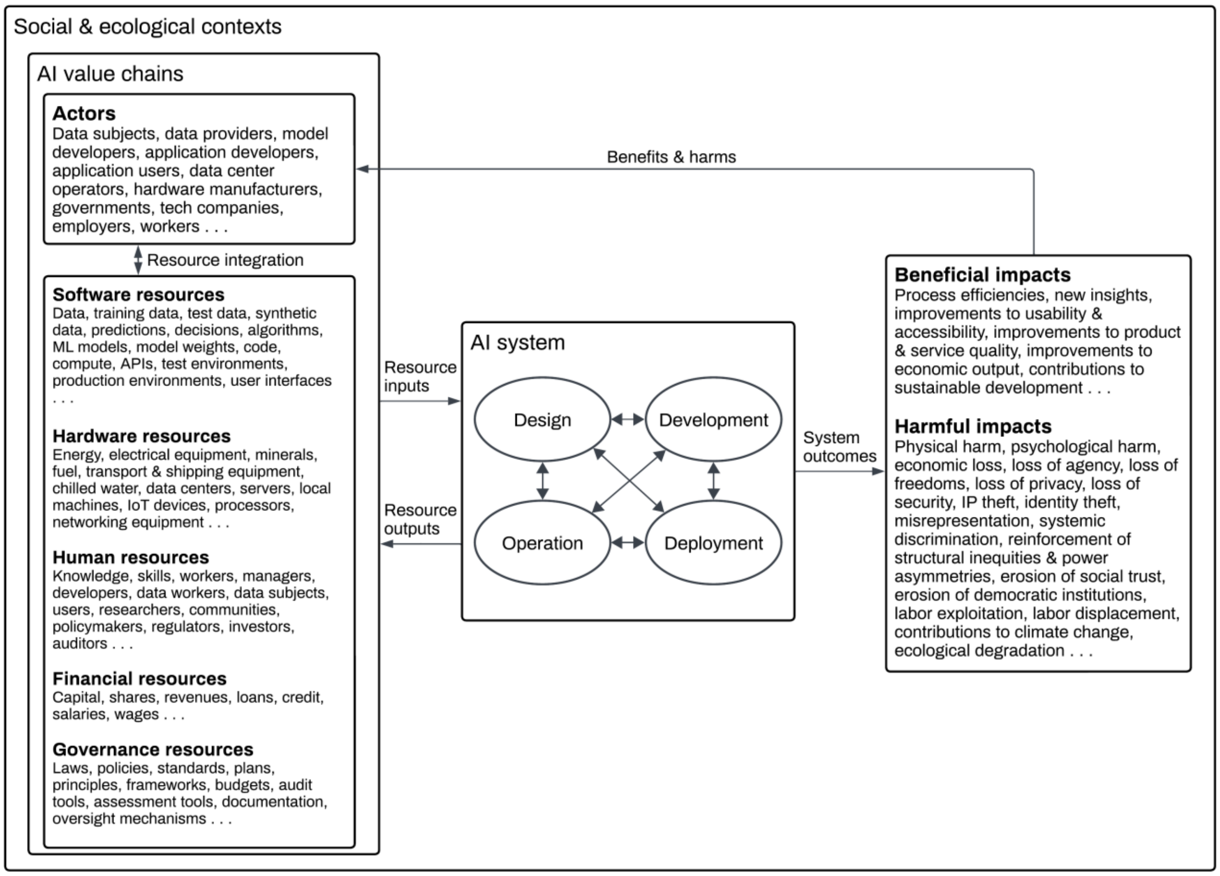

Figure 2.2: The process through which different types of actors situated within AI value chains cause beneficial or harmful impacts by providing resources to or receiving resources from an AI system throughout its lifecycle.

2.3. Methodology

To fulfill our two objectives of integrating AI ethics concerns and clarifying the ontological, ethical, practical, and policy implications of AI value chains, we conducted an integrative review of literature that covers the ethical concerns implicated in our theory of AI value chains. Snyder’s (2019) comparison of literature review methodologies recommends integrative review as an ideal method for investigating newly emerging topics to “create initial or preliminary conceptualizations and theoretical models” (p. 336). As a newly emerging topic in need of more detailed conceptual and theoretical development, the ethical implications of AI value chains are a suitable subject for an integrative review. Snyder notes that integrative review "often requires a more creative collection of data, as the purpose is usually not to cover all articles ever published on the topic but rather to combine perspectives and insights from different fields or research traditions" (p. 336). We therefore combined a high-level structure for data collection and analysis with a relatively unstructured set of methods for selecting and integrating the data (see Figure 2.3). Striving for broad representation and integration of the literature enabled us to more openly explore many newly emerging ideas, relationships, and data sources related to the ethics of AI value chains.

To conduct our integrative review, we first applied the typology of AI ethics concerns developed by Stahl et al. (2022) to create a high-level structure for our data collection and analysis process. This typology covers a uniquely wide breadth of both potential harms and benefits, such as concerns regarding control of data, security and malicious use, concentration of economic power, loss of human autonomy and freedoms, and the development of speculative "superintelligence." Other typologies of AI ethics concerns focus on categorizing harms into fine-grained subsets (Shelby et al., 2023) or on the ethical concerns related to specific types of AI models such as language models or generative AI (Solaiman et al., 2023; Weidinger et al., 2021), while inventories such as the AI Incident Database (2023) and AIAAIC (2023) categorize instances of real-world AI systems causing harm. Applying the typology of Stahl et al. allows us to present our review in a structure that illustrates that our theoretical understanding of AI value chains integrates a wide range of ethical concerns across many actors, resources, contexts, and scales of activity, thereby addressing our first research objective (*Integration of ethical concerns*).

The typology of Stahl et al. includes 4 high-level categories of ethical issues and is further subdivided into 6 types of potential benefits of AI, 8 types and 30 sub-types of potential harms, and 5 “metaphysical issues.” For each of those categories and sub-categories of ethical concerns, we identified several value chain actors and resourcing activities related to each of the concerns and recorded those actors and activities in an integrated inventory (see Figure 2.3 and Appendix 2A). We then applied a purposive sampling procedure to select sources for inclusion in our integrative review. Purposive sampling requires a high degree of researcher judgment and pre-existing domain knowledge, and is widely used in studies that are intended to provide a rich and illustrative account of some particular phenomena of interest, rather than an exhaustive or statistically representative account of all relevant data (Palinkas et al., 2015; Robinson, 2014). To provide a rich and illustrative account of the ethical issues implicated in AI value chains, we conducted an unstructured series of searches on Scopus, ScienceDirect, and Google using keywords drawn from the ethical concerns, value chain actors, and resourcing activities recorded in our inventory (Appendix 2A). We selected a source for inclusion in our sample if it met three criteria:

Criterion 1: The source was published in a scholarly journal, or by a government organization, researcher, or journalist with expertise in AI ethics.

Criterion 2: The source provides a detailed description of one or more of the ethical concerns recorded in our inventory.

Criterion 3: The source provides a detailed description of the actors and resourcing activities implicated in those ethical concerns.

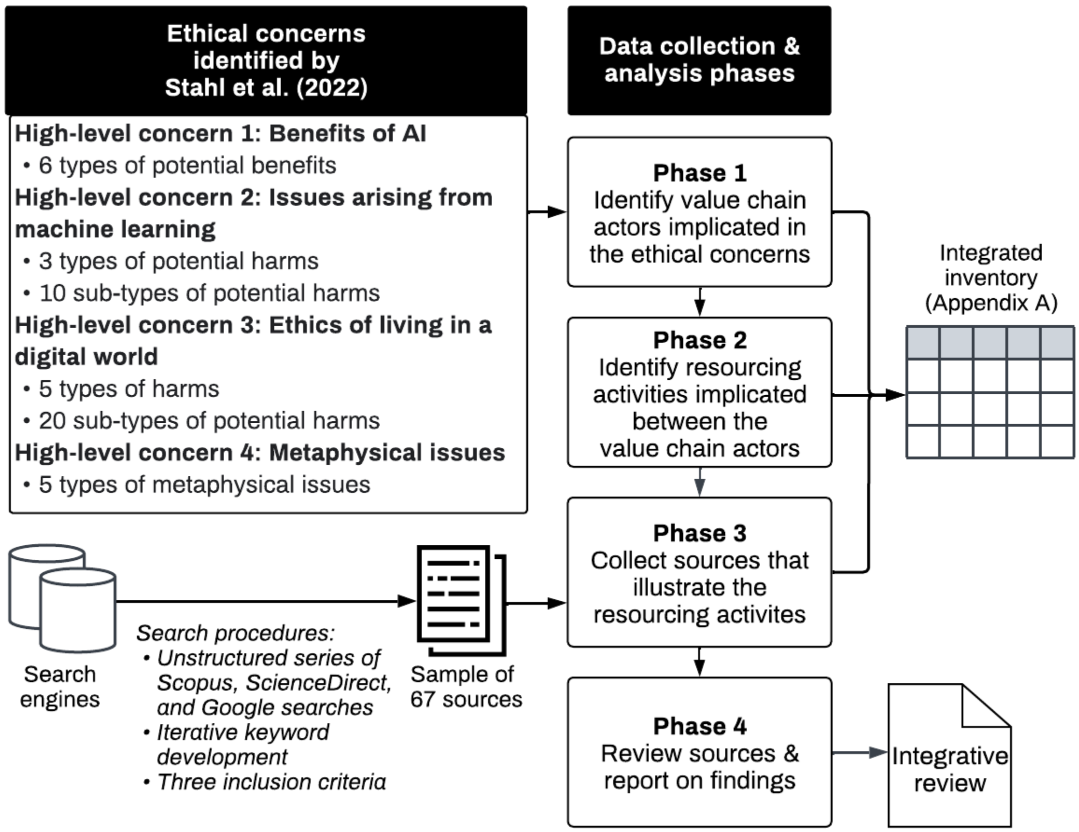

Figure 2.3: Diagram of our data collection and analysis process.

We iteratively developed search keywords using these criteria until we had collected a sample of sources sufficient to illustrate all ethical concerns, actors, and activities in our inventory. We then recorded these sources in our inventory alongside their corresponding issues, actors, and activities (Appendix 2A). This procedure resulted in 67 sources being included in our sample from research literature, grey literature, and news media that collectively illustrate a wide range of ethical concerns implicated in the AI value chain concepts we theorized in Section 2.2 (e.g., AI value chains, value chain actors, resources, resourcing activities). We reviewed each source in our sample and integrated our findings into a written description of the ethical concerns, value chain actors, and resourcing activities represented in the sources, thereby addressing our second research objective (*Clarification of value chain implications*). A full inventory of the ethical concerns, value chain actors, resourcing activities, and sources sampled in our review can be found in Appendix 2A. In writing our review, we used concepts and terminology from the theoretical framework developed in section 2.2 (e.g., AI value chains, value chain actors, resources, resourcing activities) to describe our findings. This enabled us to operationalize the framework and validate its applicability to integration and analysis of a wide range of ethical concerns. Our integrative review of the 67 sampled sources is presented in the following section.

2.4. Ethical Implications of AI Value Chains

2.4.1. AI Value Chains & Benefits of AI

Stahl et al. (2022) note that although AI ethics usually foregrounds the harms of AI systems, AI systems may also present benefits that should be accounted for. The potential benefits include: insights or efficiencies from automating the processing of large volumes of data to make predictions, decisions, or generate synthetic data outputs; improvements in economic output and reductions of environmental damage as a result of more effective and efficient production processes; contribution to United Nations Sustainable Development Goals (SDGs), as well as other international and national pursuits of socially beneficial AI adoption. However, accounting for these potential benefits within a theory of AI value chains enables us to identify many concomitant harms. Novel insights or gains to efficiency in some parts of an AI value chain may raise new risks in other parts of the value chain (Cobbe, Veale, & Singh, 2023; Gansky & McDonald, 2022; Widder & Nafus, 2023). Contributions to SDGs or “AI for good” initiatives may only be successful relative to a narrow set of values and measures (Aula & Bowles, 2023; Madianou, 2021; Moore, 2019). Economic prosperity or environmental benefits may be inequitably distributed across different groups, communities, or geographies. While AI systems may enable some value chain actors to co-create mutually beneficial outcomes, pre-existing structural injustices in the social contexts of AI systems warrant an assumption that the same systems will also result in harmful outcomes for other actors, particularly for those who belong to historically marginalized communities (Birhane, 2021; Hind & Seitz, 2022).

2.4.2. AI Value Chains & Issues Arising from Machine Learning

Stahl et al. (2022) describe many ethical concerns related to the use of machine learning (ML) technologies and methods in AI systems. These concerns include issues related to (1) *control of data*, (2) *reliability*, and (3) *lack of transparency*.

Control of data in AI value chains has been widely studied. The resources and activities required to regulate data–such as public funding, policy development, and enforcement of data protection laws in the training and application of ML models–are a significant concern related to the control of data in AI value chains (European Parliament, 2020; MacKinnon & King, 2022; Veale, Binns, & Edwards, 2018). Many other resourcing activities are implicated within the broader domain of data governance, such as informed consent from data subjects in data collection and data use activities, as well as the sale, purchase, brokerage, and ownership of ML training and testing data (Crain, 2018; Lamdan, 2022). Knowledge and expertise acquired by ML experts is also required to develop ML models, and to identify how vulnerabilities in ML models might be exploited through methods such as inversion attacks or injection attacks (Greshake et al., 2023; Wang et al., 2022). This need for specialized knowledge resources to conduct activities such as ML development and vulnerability testing implicates education and training programs for ML and ML security in AI value chains. Corporate capture of the financial and data resources needed to conduct ML research is also a concern related to control of data (Whittaker, 2021).

Control of data issues can be observed in many real-world cases. For example, the company Clearview AI scraped billions of images from platforms such as Facebook and YouTube to develop facial recognition and surveillance applications of ML that have been used by thousands of law enforcement agencies globally (Hatmaker, 2022; Perrigo, 2022). Clearview's data collection practices raise concerns regarding consent, ownership, regulation, data resourcing, financial resourcing, and other public sector resourcing activities further upstream and downstream from Clearview. Similar ethical concerns are implicated in the value chains of generative AI systems such as ChatGPT, Stable Diffusion, and Midjourney, which are trained on large volumes of data scraped from the open web, typically without explicit consent from the creators or copyright owners of that data (Lee, Cooper, & Grimmelmann, 2023). In generative AI value chains, control and ownership of data is an issue of particular importance to value chain actors such as generative AI developers, AI users, and artists and other creative workers (De Vynck, 2023; GitHub Copilot Litigation, 2023; Stable Diffusion Litigation, 2023; Vincent, 2023).

Activities and ethical concerns related to the *reliability* of ML methods and applications have also been widely studied. Inaccurate predictions, decisions, and other data outputs created through the use of unreliable ML models cause many social, political, economic, physical, and psychological harms (Angwin et al., 2016; Bender et al., 2021; Grote & Berens, 2022; Mökander & Axente, 2023; Rankin et al., 2020). The implementation of effective quality assurance practices for ML model training, testing, and management at multiple points throughout AI value chains is viewed as essential for reliable and ethical ML applications (Burr & Leslie, 2023; Eitel-Porter, 2021). The use of cloudwork platforms and labor outsourcing practices to improve data quality, model reliability, and accuracy can also cause social, political, economic, physical, and psychological harms by subjecting data workers to physically and psychologically unsafe working conditions (Irani, 2015; Miceli & Posada, 2022; Miceli, Posada, & Yang, 2022; Perrigo, 2023).

Lack of transparency in ML methods and applications presents major ethical concerns. Transparency of the funding sources for ML research and development is one such concern (Ahmed, Wahed, & Thompson, 2023; Ochigame, 2019; Whittaker, 2021). Documentation, disclosure, and explanation of the data and computational resources, processes, and outcomes involved in ML development activities and automated decision-making activities is another transparency concern (Miceli et al., 2022; Mitchell et al., 2019; Raji et al., 2020). Also of concern is disclosure of the extent to which diverse stakeholder knowledges were included in ML model design, development, and use–particularly the inclusion of vulnerable data subjects and historically marginalized groups (Birhane et al., 2022a, 2022b; Widder & Nafus, 2023). Distribution of accountability for harms amongst AI value chain actors represents another concern, as accountabilities can only be fairly distributed across actors if the resourcing activities are made transparent to one another and to authorities (Bartneck et al., 2020; Brown, 2023; Cobbe, Veale, & Singh, 2023; European Commission, 2022; Zech, 2021). Additionally, practices of collective organizing and resistance against harmful AI systems are an important ethical issue in cases where an ML application has caused harm due to a lack of fairness, accountability, and transparency in its value chain (e.g., ACLU, 2023; Broderick, 2023).

2.4.3. AI Value Chains & Ethics of Living in a Digital World

Stahl et al. (2022) outline many ethical concerns related to harms caused by AI systems as a consequence of “living in a digital world.” These concerns include: (1) *economic issues*, (2) *justice*, (3) *human freedoms*, (4) *broader societal issues*, and (5) *unknown issues*.

Economic issues are especially significant in AI value chains. The use of automation and biometrics in human resources practices–along with related labor resourcing activities such as the hiring, contracting, dismissal, and surveilling of workers–represents a political-economic issue of ethical concern in AI value chains (Bales & Stone, 2020; Hickok & Maslej, 2023), as does distribution of wealth, capital, and other financial resources and labor exploitation across AI value chains (Dyer-Witheford, Kjøsen, & Steinhoff, 2019; Miceli & Posada, 2022; Miceli, Posada, & Yang, 2022). Open-source development of AI systems and open access to data, code, and other software resources also pose concerns related to asymmetries of political and economic power (Langenkamp & Yue, 2022; Masiello & Slater, 2023; Widder, West, & Whittaker, 2023).

One particular case involving many of these economic issues is OpenAI's outsourcing of data labeling activities to workers employed by Sama AI in Kenya, many of whom were psychologically harmed and undercompensated during their employment (Perrigo, 2023). In another case, Amazon's development, use, and subsequent disuse of a hiring automation tool that discriminated against women represents another issue related to harmful labor resourcing activities (Dastin, 2018). The consolidation of models and datasets in an increasingly small group of private sector actors further demonstrates that political and economic harms can be caused by resource distribution imbalances across AI value chains (Ahmed, Wahed, & Thompson, 2023), as do the summer 2023 labor strikes of the Writers' Guild of America (WGA) and Screen Actors Guild-American Federation of Television and Radio Artists (SAG-AFTRA). WGA and SAG-AFTRA workers demanded that their employers refrain from using their likenesses or union-protected creative materials in generative AI training datasets, as well as from requiring them to use generative AI applications in their work activities without their consent (Broderick, 2023; Webster, 2023).

There are many ethical concerns related to *justice* in AI value chains. For example, public sector procurement, development, and use of automated decision systems in public service and courtroom contexts has caused harm to vulnerable groups (Angwin et al., 2016; Eubanks, 2018; Gans-Combe, 2022; Mulligan & Bamberger, 2019). The inclusion of knowledges and perspectives from historically marginalized groups in AI education, development, and governance processes is another concern in resourcing the knowledge required to ethically develop AI systems (Birhane et al., 2022a; West, Whittaker, & Crawford, 2019). The just distribution of value co-creation outcomes across AI system lifecycles is a concern to many researchers, as is the potential for macro-scale social, political, and economic injustice caused by widespread AI adoption (Dyer-Witheford, Kjøsen, & Steinhoff, 2019; Pasquale, 2020; Solaiman et al., 2023).

Many resourcing activities in AI value chains have impacts on *human freedoms*. Ethical concerns related to human freedoms often overlap with economic issues and justice issues, as these concerns typically stem from a structural lack of freedom to access a particular kind of resource, which in turn, perpetuates a further loss of freedoms. For example, harmful outcomes of exploitative labor outsourcing and algorithmic discrimination often result in marginalized groups experiencing a further loss of access to resources needed to pursue social, political, and economic opportunities (Angwin et al., 2016; Eubanks, 2018; GPAI, 2022; Miceli & Posada, 2022). Restrictive access to public sector and private sector data, information, and computational resources can also result in disproportionate levels of access to and benefit from those resources becoming further reinforced between individuals, groups, sectors, and governments (Ahmed, Wahed, & Thompson, 2023; Whittaker, 2021).

What Stahl et al. (2022) refer to as *broader societal issues* is a category of ethical concerns containing a variety of large-scale impacts on potentials for physical conflict, environmental degradation, and erosion of democratic institutions. For example, military and police procurement of use-of-force and surveillance applications represents a broad societal issue in which many value chains become implicated in potentials for causing physical conflict and violence (Hoijtink & Hardeveld, 2022; Mahoney, 2020; Mulligan & Bamberger, 2019; Taddeo et al., 2021). Social filter bubbles created from the algorithmic profiling and manipulation of social media users is another broad societal issue, implicating many value chains in potentials for causing social, political, physical, and psychological harms (Krönke, 2019; Woolley, 2018). Energy and water are material resources required to train AI models and operate AI systems, and these material resourcing activities also constitute a broad societal issue. Energy and water usage in AI hardware and infrastructure results in substantial carbon emissions, depletion of freshwater reserves, and other global and local environmental harms (GPAI, 2021; Li et al., 2023; Luccioni & Hernandez-Garcia, 2023). In addition to energy and water, mineral extraction and other mining, manufacturing, transportation, and assembly processes involved in the material resourcing of AI systems all represent a broad societal issue, as does the disposal and recycling of environmentally harmful electronic waste at the end of AI hardware lifecycles (Crawford, 2021; Crawford & Joler, 2018).

What Stahl et al. (2022) refer to as *unknown issues* is a category of ethical concerns containing a variety of complex harms that are difficult to predict the potential consequences of. For example, malicious value chain actors might engage in unforeseen misuses of personal data, digital identities, misinformation, or other resources in their malicious development and/or use of AI systems (Brundage et al., 2018). The enforcement of excessively strict or excessively permissive AI regulations may also cause a variety of complex social, political, and economic harms (Ada Lovelace Institute, 2021; Smuha, 2021). Additionally, excessive funding of AI research that prioritizes finding solutions to the wrong problems might result in some AI risks and harms becoming less foreseeable and/or preventable (Tiku, 2023).

2.4.4. AI Value Chains & Metaphysical Issues

Stahl et al. (2022) describe several "metaphysical issues" pertaining to speculative ethical concerns such as machine consciousness, artificial moral agents, artificial "super-intelligence", and changes to "human nature" enabled by new AI technologies. These "metaphysical issues" are purely speculative. However, the ethical implications of these hypothetical activities are comparable to the ethical implications of empirically observable resourcing activities. For example, concerns related to the distribution of benefit/harm through the development of a speculative "autonomous" *artificial moral agent* are comparable to real-world concerns related to *human moral agents* distributing benefit/harm through the development and use of automated systems. Similarly, issues of resource distribution, consolidation, and power asymmetry arising from the development of speculative *superintelligent agents* are comparable to issues of resource distribution, resource consolidation, and power asymmetry that exist between real-world *human agents*.

Some researchers convincingly argue for disregarding these speculative ethical concerns and instead accounting for real, present harms (Gebru & Torres, 2023; Torres, 2023). We advance these arguments by noting that the ethical concerns underlying these speculative "metaphysical issues" are futurological extensions of ethical concerns that can already be observed in real-world, present-day AI value chains. Therefore, greater study can and should be given to the empirically observable AI value chains. We also note that AI value chain theories and methods are flexible enough to account for the benefits and harms of AI systems across multiple spatial, temporal, and organizational scales (including those benefits and harms that exist only in speculative futures not meriting significant study).

2.5. Limitations

We have applied Snyder's (2019) guidance on integrative review methodology to illustrate the ethical implications of AI value chains through an integrative account of the ethical concerns, value chain actors, resources, and activities described in a sample of 67 sources. Our review is not intended to provide a comprehensive account of every source that describes the ethical implications of AI value chains. Because we do not apply systematic review methods, the comprehensiveness of our review is limited: our review provides analysis of high-level ethical concerns that appear across a sample of the AI ethics literature, rather than a comprehensive discussion of every relevant source pertaining to every one of those ethical concerns. Despite this limitation, our review fulfills our two research objectives: our review provides an integrated account of a wide range of ethical concerns (*Objective 1*) and clarifies the ontological, ethical, practical, and policy implications of AI value chains by applying concepts developed in our theory of AI value chains (*Objective 2*).

2.6. Future Directions for Research, Practice, & Policy

Future research, practice, and policy should more comprehensively account for, integrate, and intervene in the range of ethical concerns, value chain actors, and resourcing activities we outlined in Section 2.4. An integrative approach to accounting for and intervening in the ethics of AI value chains requires intervening in many types of resources, activities, and societal and environmental impacts. Software resources and the activities performed to integrate software resources into AI systems–for example, the collection and preparation of training and test data, the development and deployment of models, and operation and monitoring of a model throughout its software lifecycle–raise ethical concerns at many upstream and downstream sites in AI value chains. In addition to software resources, many other types of resources, activities, and impacts raise ethical concerns, including the resourcing of hardware, finances, knowledge, labor, and governance mechanisms throughout the lifecycles of AI systems.

There are three opportunities for researchers and practitioners to further investigate the wide range of ethical concerns implicated in AI value chains:

(1) *Conduct more empirical and action research* into the specific ethical concerns, value chain actors, and resourcing activities we outlined in Section 2.4. Future research agendas could include, for example, empirical and participatory studies of the impacts of generative AI development and use on artists and workers, or studies of the impacts of outsourcing practices on marginalized workers in AI value chains. By collecting and analyzing more quantitative and qualitative data pertaining to a variety of real-world AI value chains and their related actors and activities, researchers can provide a rich evidence base upon which other researchers, practitioners, and policymakers can develop further research, practice, and policy on the basis of. Additionally, by empowering value chain actors to participate in research activities, identifying their concerns and needs, and developing interventions that are designed to satisfy their needs, researchers can generate more detailed insights on stakeholder perspectives, best practices, policy gaps, and policy options.

(2) *Develop and apply theories and methods* for systematically modeling AI value chains, analyzing a diverse range of ethical concerns in those value chains, and enacting interventions in those value chains. Many frameworks for systematically modeling and analyzing value chains and value networks can be applied to studies of AI value chains, such as the service system

analysis framework of Frost, Cheng, and Lyons (2019) and the data value chain framework of Lim et al. (2018). These and similar frameworks can help to ground future research on the ethics of AI value chains in stronger value chain theories and methodologies.

(3) *Design and implement ethical sourcing practices* across all of the value chains that provide resource inputs to or receive resource outputs from AI systems. Ethical sourcing practices have been adopted in with different degrees of effectiveness in many industries, and are intended to support in "managing all processes of supplying the firm with required materials and services from a set of suppliers in an ethical and socially responsible manner" (Kim, Colicchia, & Menachof, 2018, p. 1033). In the context of AI practices, ethical sourcing requires all actors that provide resources inputs to or receive resource outputs from AI systems to account for a diverse range of impacts that their activities have on society and the environment (Widder & Wong, 2023). Many frameworks for guiding ethical sourcing practices in AI systems have been developed by academic researchers, such as documentation and auditing frameworks for data and model resourcing (Mitchell et al., 2019; Miceli et al., 2022; Raji et al., 2020). Fairwork's principles and practices for preventing harms to workers across AI value chains (Fairwork, 2023; GPAI, 2022) and Global Partnership on AI's principles and practices for mitigating the harmful environmental impacts of AI systems (GPAI, 2021) represent two ethical sourcing frameworks that cover a range of ethical concerns related to labor and the environment.

Governance mechanisms such as industry standards, certification programs, guidance documents, and codes of conduct should also be used to support the implementation of ethical sourcing practices across AI value chains. Many existing standards, certification programs, guidance documents, and codes of conduct focus on a narrow socio-technical context, while affording minimal or no attention to concerns implicated in the broader social, political, economic, material, and ecological contexts of AI value chains (Government of Canada, 2023a; ISO, 2023; NIST, 2023; Responsible Artificial Intelligence Institute, 2022). In contrast, the Treasury Board of Canada Secretariat has published a voluntary guide for using generative AI applications in Canada's federal public sector that provides principles and best practices for a wide-ranging set of ethical concerns, such as AI literacy development, professional autonomy, and environmental impact mitigation (Government of Canada, 2023b). Future iterations of these and other AI governance mechanisms–such as legislation, regulations, and other policy instruments–should be used to implement more comprehensive principles and best practices for ethical sourcing across the many actors, activities, and contexts of AI value chains.

2.7. Conclusion

We have reviewed and synthesized theories of value chains and AI value chains. We have also conducted an integrative review of recent literature on the ethical implications of AI value chains. We have therefore accomplished our two objectives of (1) integrating AI ethics concerns into our conceptualization of AI value chains, and (2) clarifying the ontological, ethical, practical, and policy implications of AI value chains. In doing so, we have made a scientific contribution to the theoretical development, scholarly knowledge, and practitioner knowledge of AI value chains. The opportunities for future research and practice we outline also represent a significant practical contribution, as these opportunities provide a preliminary agenda for further advancing studies and practices of ethical AI value chain governance.

AI value chains will remain a focal point of AI ethics and governance initiatives into the foreseeable future. As those initiatives continue to unfold, researchers, practitioners, and policymakers with an interest in the ethics of AI value chains can look to this paper for guidance in conceptualizing and conducting their work. AI ethics must advance beyond decontextualized discussions of ethics and toward value chain perspectives that situate actors in context, account for the many types of resources involved in co-creating AI systems, and integrate a wider range of ethical concerns across contexts and scales.

Chapter 3

AI Governance Initiatives in Canada

About this article: This article was co-authored with Ana Brandusescu and Kelly Lyons beginning in January 2022 and was originally titled *The Governance of Artificial Intelligence in Canada: Findings and Opportunities from a Review of 84 AI Governance Initiatives*. The original article was published in *Government Information Quarterly* in April 2024 and can be found at: https://www.sciencedirect.com/science/article/pii/S0740624X24000212

Abstract

In recent years, the effective governance of artificial intelligence (AI) systems has become a strategic necessity for many nations. Among those nations, Canada is particularly noteworthy: Canada was the first nation to implement a national AI strategy, and more recently, Canada's federal and provincial governments have designed and implemented a wide range of initiatives that attempt to intervene in a variety of potential impacts associated with AI systems. We present a semi-systematic review and synthesis of 84 of those AI governance initiatives. We find that those 84 initiatives predominantly focus on developing programs, policies, and strategic plans to intervene in industry and innovation, technology production and use, AI research, and public administration. Conversely, we find relatively little focus on developing ethics statements or standards, as well as little focus on intervening in social and workforce development services, AI education and training, and digital infrastructure. We suggest three opportunities for researchers and four opportunities for practitioners that, if enacted, would strengthen the overall state of Canadian AI governance. Our study contributes a novel macro-scale synthesis of AI governance initiatives within a national context, as well as practical opportunities for intervening in national AI governance challenges related to evaluation of initiative outcomes, public trust and participation in initiatives, AI impact representation in initiatives, and national unification.

Keywords: Artificial intelligence, governance, policy, strategy, AI ethics, literature review, content analysis

3.1. Introduction

The governance of artificial intelligence (AI) systems has quickly become a strategic imperative for governments around the world, with more than 40 national governments having adopted AI strategies as of 2022 (OECD, 2022b). Among those nations, Canada is especially notable: in 2017, Canada became the first country to adopt a national AI strategy (Radu, 2021), and in the years since then, Canada's system of AI governance has continued to evolve with a distinctive focus on enabling AI research and talent acquisition, leadership in ethical AI, and multistakeholderism in AI policymaking (Kung, Boskovic, & Stix, 2020). Throughout that evolution, many initiatives intended to govern the development and use of AI systems have been launched across Canada's public and private sector, as well as across many levels of government.

Canada's early entry into AI governance in 2017–along with the country's large-scale and continuous advancements in AI governance across multiple sectors and levels of government (Brandusescu, 2021; Frost, 2020)–have made Canada into a uniquely information-rich national AI governance context. The complexity, diversity, and maturity of Canada's system of AI governance makes Canada an ideal subject for an empirical and integrative study of AI governance initiatives and stakeholders within a national context. A study of Canada's system of AI governance would therefore fill a gap that Birkstedt et al. (2023) note in their recent review of the AI governance literature: AI governance research requires more empirical, integrative, and context-sensitive analysis of a diverse range of AI governance initiatives and stakeholders within specific AI governance contexts.

To fill that gap, we conducted a study of a selection of 84 AI governance initiatives that were launched or commissioned by the federal government and three provincial governments in Canada from 2017 to 2022. Many observers in academia, industry, and civil society have already noted that Canadian AI governance initiatives face many challenges, including: public mistrust in AI technologies (Deloitte, 2019; Gaon & Stedman, 2019; Ipsos, 2022; Reeveley, 2021), public engagement on major AI governance initiatives (Attard-Frost, 2022; Brandusescu & Sieber, 2022; Centre for Digital Rights, 2022; Karanicolas, 2019; McKelvey & MacDonald, 2019; Rice, 2021; Tessono et al., 2022), accountability in the public funding and public procurement of AI technologies (Brandusescu, 2021; Lepage-Richer & McKelvey, 2022; McKelvey, McPhail, & Rajabiun, 2022; Stevens & Brandusescu, 2021; Stevens & Solomun, 2021), commercialization, adoption, and retention of Canadian AI innovation and talent (Babashahiashtiani, 2021; Innovation, Science and Economic Development Canada, 2020; Ouimette, 2022), and protection of consumers, workers, and citizens against harmful practices of AI use (Bailey et al., 2021; Centre for Digital Rights, 2022; Munro, 2019; Office of the Privacy Commissioner of Canada, 2021a, 2021b; Rice, 2021; Robertson, Khoo, & Song, 2020). Given these existing observations of the challenges facing Canadian AI governance initiatives, our study is guided by two research questions and two main research objectives that respond to those research questions:

Research Question 1: What AI governance initiatives have been undertaken by Canadian federal and provincial governments?

Research Question 2: What opportunities exist for strengthening Canada’s system of AI governance?

Objective 1: Review a diversity of Canadian AI governance initiatives to fill the research gap in empirical, integrative, and context-sensitive analysis of a diverse range of AI governance initiatives and stakeholders.

Objective 2: Synthesize the findings of our review into a set of opportunities that researchers and public servants can act upon to strengthen Canada's system of AI governance.

In Section 3.2, we discuss the theoretical background for our study. In Section 3.3, we describe the research design and methodology for our study. In Section 3.4, we present our key findings from a semi-systematic review of Canada's AI governance initiatives. In Section 3.5, we outline a set of opportunities for researchers and practitioners to strengthen Canada's system of AI governance based on the findings of our review. In Section 3.6, we describe the limitations of our study. We conclude in Section 3.7 by discussing the contributions of our study to AI governance research and practice.

3.2. Theoretical Background

3.2.1. Artificial Intelligence & AI Governance

Artificial intelligence (AI) has no universally agreed upon definition, but AI is often characterized as a type of technological system that is capable of performing cognitive activity within a specific application context and with some degree of autonomy from human actors (Lewis et al., 2018; Dwivedi et al., 2021; Nitzberg & Zysman, 2022; OECD, 2022a; Madan & Ashok, 2023). Because AI systems are capable of causing a wide range of beneficial and harmful impacts, *AI governance* has emerged as a field of research and practice. Like AI itself, AI governance has no universally agreed upon definition, but scholars tend to characterize AI governance as a practice intended to maximize various types of beneficial impacts and minimize various types of harmful impacts in the development and use of AI systems (Kuziemski & Misuraca, 2020; Dwivedi et al, 2021; Nitzberg & Zysman, 2022; Wirtz, Weyerer, & Kehl, 2022; Camilleri, 2023). While AI governance practices can at times overlap with data governance practices (Janssen et al., 2020; Mäntymäki et al., 2022), distinct practices of AI governance have been observed across various contexts and scales of activity, including practices of global and international AI governance (Erdélyi & Goldsmith, 2022; Roberts et al., 2023; Tallberg et al., 2023), national AI governance (Radu, 2021; Djeffal, Siewert, & Wurster, 2022; Wilson, 2022), local and municipal AI governance (Kinder et al., 2023; Wan & Sieber, 2023), sectoral AI governance (Kuziemski & Misuraca, 2020; Zuiderwijk, Chen, & Salem, 2021), and organizational AI governance (Cihon, Schuett, & Baum, 2021; Mäntymäki et al., 2022). AI governance practices are intended to intervene in the impacts of AI systems through the design and implementation of several types of initiatives, including strategies, policies, government programs, technical standards, and ethics statements (Birkstedt et al., 2023). We describe the goals and activities involved in each of these types of initiatives in greater detail in Section 3.3.

3.2.2. Initiatives & AI Governance Initiatives

In the organizational studies literature, *initiatives* are commonly characterized as planned organizational actions in which resources are integrated and directed toward a long-term outcome that significantly changes organizational structures and operations, such as the implementation of new strategies, policies, or programs (Flamholtz & Randle, 2008; Saunders, Mann, & Smith, 2008). Initiatives are therefore a broader category of organizational activity than projects or programs, and are especially useful in strategic contexts where new rules and norms are emerging, organizational knowledge and structures are fluid, and the operational environment is complex and volatile (Keenan et al., 2013; Saunders, Mann, & Smith, 2008; Vasileva, 2019). AI governance contexts at national scales are often theorized as strategic contexts where new rules and norms are emerging, organizational knowledge and structures are fluid, and the operational environment is complex and volatile (Djeffal, Siewert, & Wurster, 2022; Frost, 2020; Radu, 2021). This makes organizational theory's conceptualization of initiatives especially applicable to studying a national AI governance context such as Canada's system of AI governance. In our study of Canada's system of AI governance, we apply an organizational theory perspective to conceptualize *AI governance initiatives* as organizational activities intended to institute new strategic plans, programs, policies, standards, or ethical guidance for intervening in impacts caused by the development and/or use of AI systems.

3.3. Research Design & Methodology

Snyder's (2019) influential comparative analysis of literature review methodologies finds that semi-systematic review is an ideal method in multidisciplinary and voluminous review contexts (such as a review of AI governance literature), where "to review every single article that could be relevant to the topic is simply not possible" (p. 335). Unlike systematic review methodologies (which use only highly structured search and data collection methods) and integrative review methodologies (which use relatively unstructured search strategies to facilitate greater critical reflection on the literature), a semi-systematic review combines structured and unstructured search strategies to "synthesize the state of knowledge" (p. 335) in complex research contexts.

We therefore applied a semi-systematic methodology to search for and review web content and documentation pertaining to the AI governance initiatives of the Canadian federal government, as well as the AI governance initiatives of the three provincial governments of Ontario, Québec, and Alberta. These three particular provinces were selected for inclusion in the study due to the high concentration of AI-related research, innovation, and investment that has been cultivated in Ontario, Québec, and Alberta through regionally targeted initiatives, such as the three National AI Institutes that are located within these three provinces and administered through the Pan-Canadian AI Strategy. Although our selection of provinces and initiatives does not provide an exhaustive account of all AI governance initiatives undertaken by Canadian governments, our focus on the federal government along with these three particular provincial governments ensured that our study would represent the most strategically significant public sector AI governance initiatives in Canada.

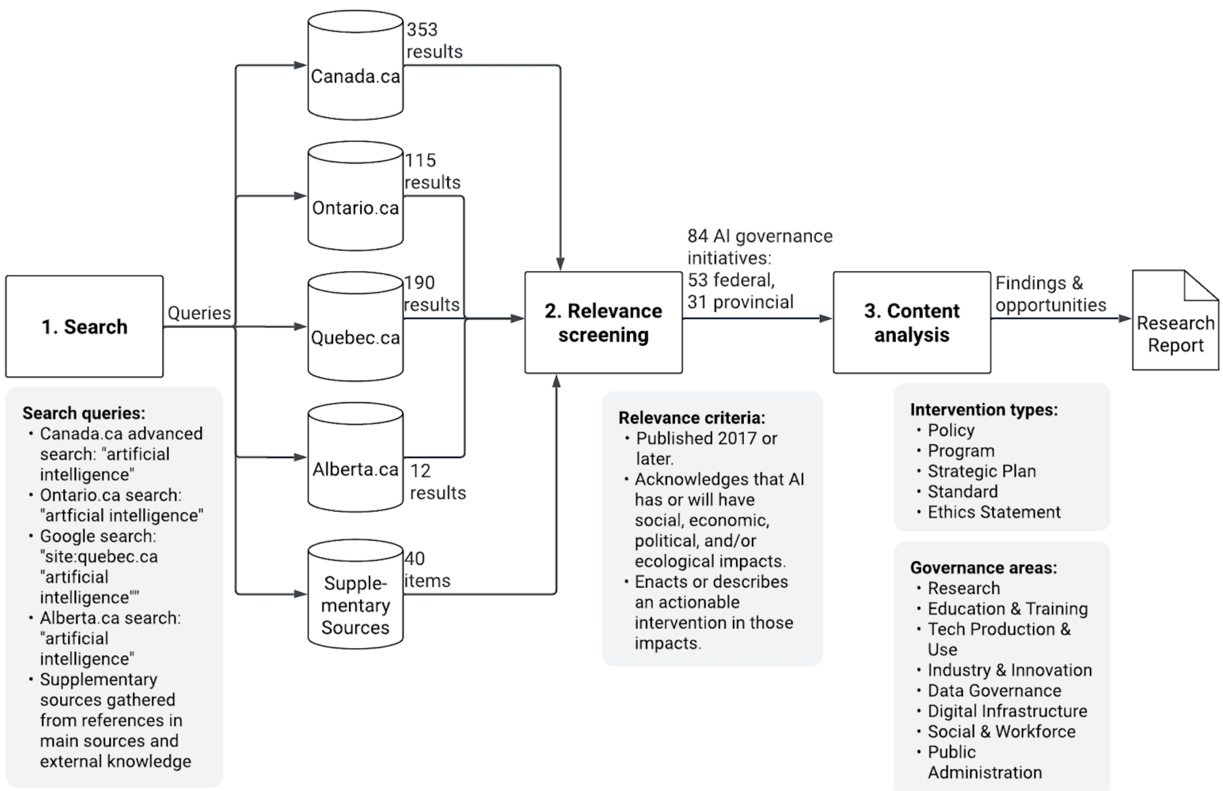

Figure 3.1: Diagram created by the authors of the data collection and analysis process of our study.

As illustrated in Figure 3.1, we conducted our study in three phases: *search*, *relevance screening*, and *content analysis*.

Phase 1: Search. The search process was mainly conducted from January 2022 to March 2022. Within that timeframe, we conducted searches using the search engines of federal and provincial government websites to locate all content on each website containing the term "artificial intelligence." Additionally, we compiled supplementary sources based on our prior knowledge of AI governance initiatives in Canada, our knowledge of new initiatives that were publicly announced during a later phase of the study, as well as sources that were referenced in the content gathered from the four search engines that did not appear as search results themselves. This combination of structured and unstructured search strategies was chosen to enable us to account for a more comprehensive set of initiatives that was better representative of the current state of Canada's system of AI governance. A full list of the methods through which each of the sources included in our study were located (search engine, content referencing, or external knowledge) can be found in Appendix 3A.

Phase 2: Relevance screening. The sources collected from the four search engines along with the supplementary sources were read and screened with reference to three relevance criteria. These criteria were used to determine whether or not the source described a unique AI governance initiative of relevance to the study:

Criterion 1: The source was published in 2017 or later (the year that Canada's national AI strategy was launched).

Criterion 2: The source acknowledges that AI has or will have social, economic, political, and/or ecological impacts.

Criterion 3: The source represents an initiative in that it enacts or describes an actionable intervention in those impacts.

Relevant sources were included in the next phase of the review, while non-relevant sources were excluded. A full inventory of all of the relevant initiatives we analyzed during our study can be found in Appendix 3A. As shown in Figure 3.1, the search process produced a total of 699 sources, of which 84 sources (53 federal and 31 provincial) were identified as relevant to our study.

Phase 3: Content Analysis. Our content analysis began in January 2022 and concluded at the end of January 2023. Web content associated with the 84 relevant sources was iteratively analyzed in order to develop codes, themes, and categories pertaining to the types of governance interventions and governance areas covered by each initiative (see Figure 3.2). Codes, themes, and categories emerged inductively from analyzing 33 extended-length observational notes made during analysis of common activities and concerns across multiple initiatives (referred to in Figure 3.2 as "cross-initiative observations"). We then identified additional secondary sources that were coherent with and relevant to the findings of our content analysis, and could be referenced in reporting on the findings of our content analysis (Sections 3.4 and 3.5) to enhance their robustness. The content analysis was conducted in accordance with the "conventional content analysis" methods of inductive coding and iterative category development described by Hsieh and Shannon (2005) and the six trustworthiness criteria described by Nowell et al. (2017):

(1) *Credibility* was ensured through prolonged and persistent engagement with the content over the course of one year, triangulation of codes and categories through regular discussions between the three researchers, and ongoing peer debriefing to provide external checks on the validity of preliminary findings as they emerged. Observing similarities between our categories and categories developed through an earlier comparative case study of Canadian and other national AI strategies (Kung, Boskovic, & Stix, 2020) also helped us ensure the credibility of our analysis.

(2) *Transferability* to other research contexts was ensured by providing thick descriptions of the social, political, economic, and ecological contexts surrounding the AI governance

activities we observed (transferability is discussed in more detail in Sections 3.6 and 3.7).

(3) *Dependability* was ensured by keeping "logical, traceable, and clearly documented" records of our data collection and analysis activities throughout our research process, as well as by providing detailed reporting on our methodological decisions in this section.

(4) *Confirmability* was ensured by keeping records of and justifying our theoretical and methodological decisions throughout our reporting on the study.

(5) *Auditability* was ensured by documenting our research process, the data we collected, the codes, categories, and themes we developed, and our methodological decisions in collecting and analyzing the data. Our reporting on our data collection and analysis processes throughout this section, Appendix 3A, and Appendix 3B is intended to ensure that another researcher with "the same data, perspective, and situation could arrive at the same or comparable, but not contradictory, conclusions."

(6) *Reflexivity* was ensured early in the study by keeping a "reflexive journal" of extended-length cross-initiative observations and personal reflections, as well as later on in the study by organizing regular researcher meetings and a shared codebook in which we could collaboratively discuss our observations and establish shared interpretations of their significance for our study.

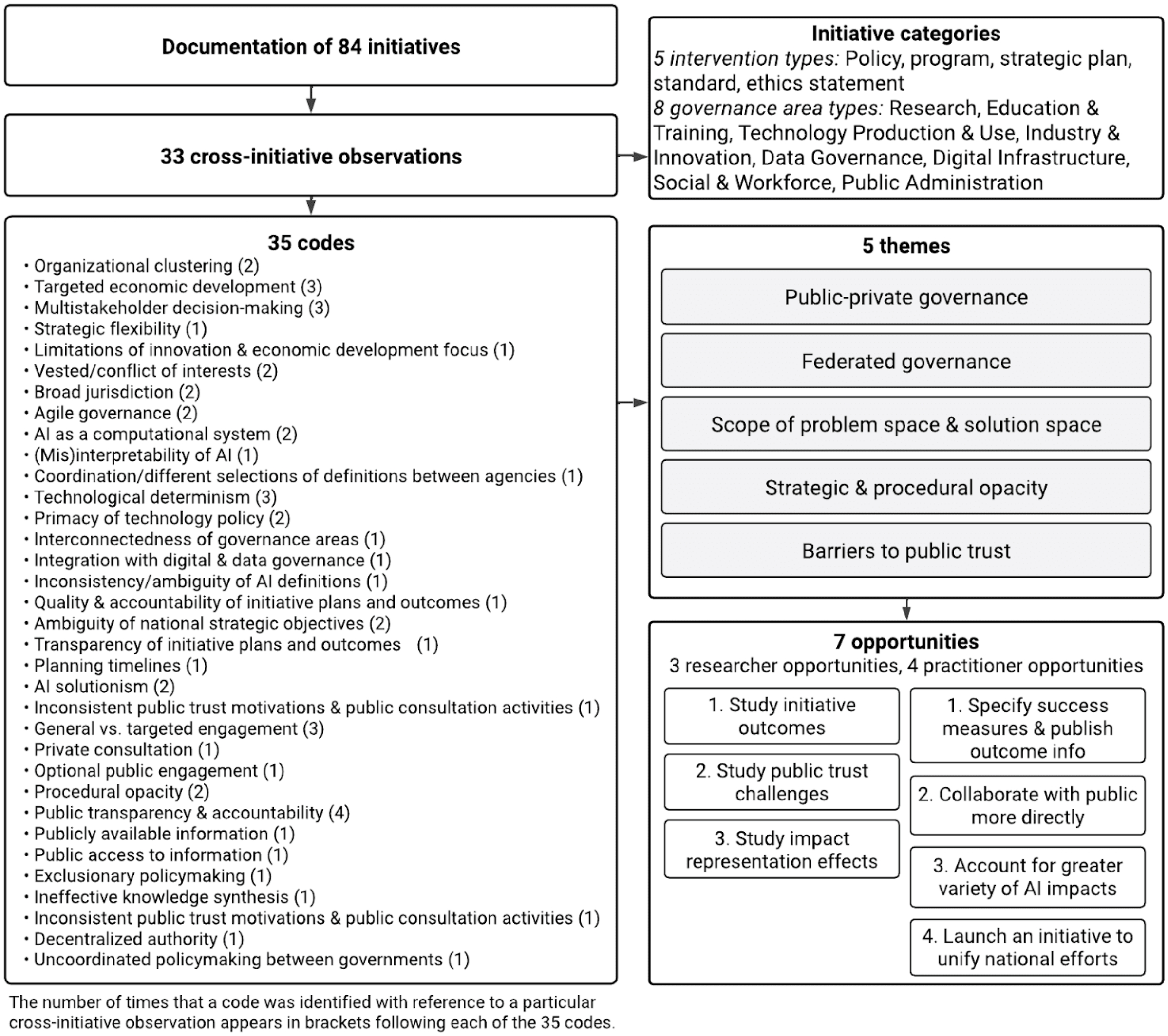

Figure 3.2: Diagram created by the authors illustrating the structure of the content analysis process through which categories, codes, themes, and opportunities for improvement were generated.

3.4. Findings

3.4.1. Overview of Findings

Across the 84 initiatives included in our study, we identified and defined five intervention types and eight governance areas, as shown Tables 3.1 and 3.2. A full inventory of the names, departments, governments, years of origin, intervention types, governance areas, discovery methods, and webpages associated with all 84 initiatives can be found in Appendix 3A. A full breakdown of the data we created pertaining to the level of government, intervention type, governance area, and year of origin associated with each initiative can be found in Appendix 3B. Charts depicting the total number of initiatives by year of origin, intervention type, and governance area can be found in Figures 3.4 and 3.5. For added clarity regarding the organizations we discuss and the initiatives they are responsible for, a diagram depicting a set of key federal entities, relations, and acronyms can be found in Figure 3.3.

Table 3.1: Intervention types that were identified and defined during content analysis.

(1) **Policy**: A "set of statements of principles, values and intent that outlines expectations and provides a basis for consistent decision-making and resource allocation in respect to a specific issue" (Government of Canada, 2021). Includes many types of instruments such as legislation, regulations, and directives, as well as the activities and recommendations involved in developing policy instruments.
(2) **Program**: A collection of projects that are managed together with "shared strategic intent" (Saunders, Mann, & Smith, 2008, p. 1096).
(3) **Strategic Plan**: A detailed action plan intended to "address the key strategic objectives" of an organization (Saunders, Mann, & Smith, 2008, p. 1106).
(4) **Standard**: An in-progress or completed document that is "established by a consensus of subject matter experts and approved by a recognized body that provides guidance on the design, use or performance of materials, products, processes, services, systems or persons." (ISO, n.d.)
(5) **Ethics Statement**: Statements of principles, declarations, or guidelines that "prescribe ethical direction to AIS [AI system] developers, AIS users, policymakers, and other stakeholders who seek to maximize the potential benefits and minimize the potential harms of AIS operations" (Attard-Frost, De los Ríos, & Walters, 2022, p. 1).

Table 3.2: Governance areas that were identified and defined during content analysis.

(1) **AI Research**: In response to perceived AI impacts, the initiative describes or enacts an intervention in scientific, social scientific and/or humanities research on AI.
(2) **AI Education & Training**: In response to perceived AI impacts, the initiative describes or enacts an intervention in education or training on AI design, development, and/or use.

(3) **Technology Production & Use**: In response to perceived AI impacts, the initiative describes or enacts an intervention in the design, development, and/or use requirements of AI technologies.
(4) **Industry & Innovation**: In response to perceived AI impacts, the initiative describes or enacts an intervention in economic activities related to AI innovation or AI-enabled industrial/sectoral development.
(5) **Data Governance**: In response to perceived AI impacts, the initiative describes or enacts an intervention in the collection, use, privacy, security, management, or stewardship of data.
(6) **Digital Infrastructure**: In response to perceived AI impacts, the initiative describes or enacts an intervention in the development or operation of digital hardware or infrastructure, such as computing environments, data centres, CPUs, or GPUs.
(7) **Social & Workforce**: In response to perceived AI impacts, the initiative describes or enacts an intervention in social or workforce development services.
(8) **Public Administration**: In response to perceived AI impacts, the initiative describes or enacts an intervention in the design, development, and/or use of AI applications in public administration contexts.

To maintain consistency with a theoretical understanding of “initiatives” as organizational activities, the definitions we assigned to intervention types were based on definitions provided by researchers and institutions that describe the intervention in terms of the organizational activities that it involves. The definitions we assigned to governance areas were based on the types of AI impacts acknowledged within the initiative, as well as based on the context of the initiative’s proposed or enacted intervention. Each initiative represented only a single type of intervention, but many initiatives were relevant to more than one governance area. Intervention types are therefore mutually exclusive while counts of governance areas are not mutually exclusive.

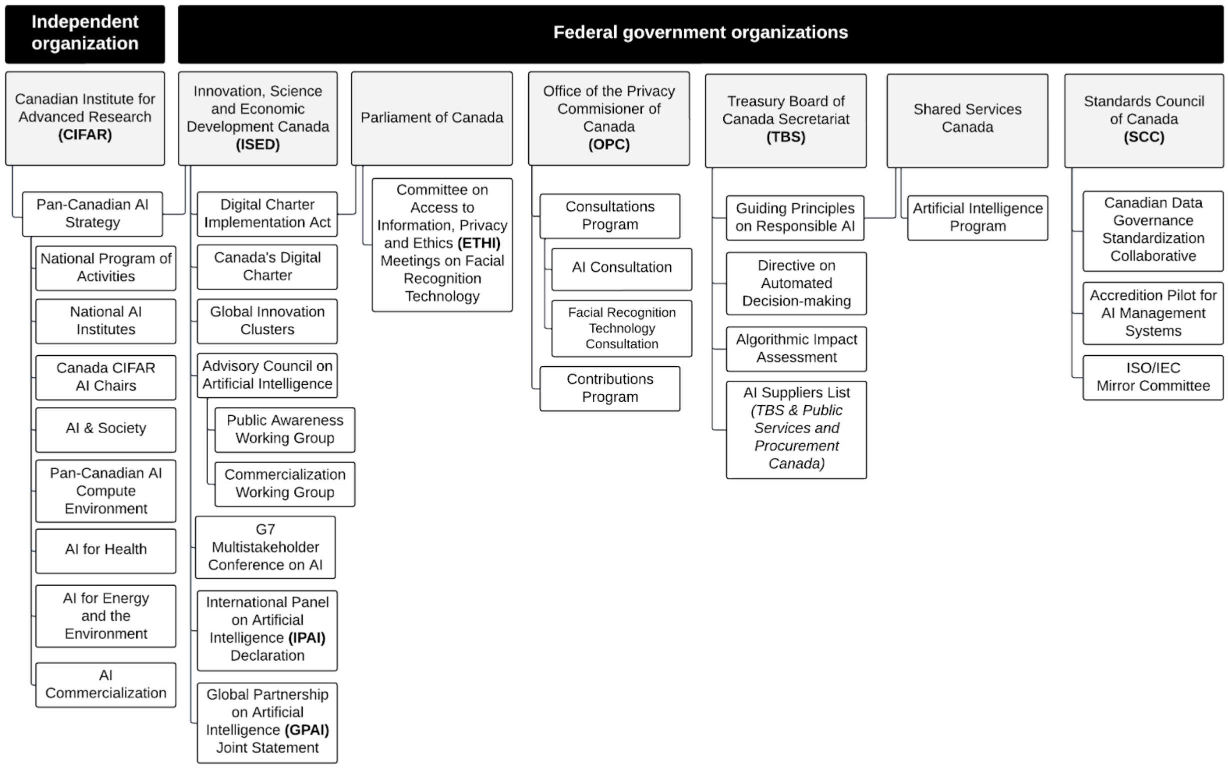

Figure 3.3: Diagram of the structure of a subset of federal organizations, initiatives, and lines of responsibility that are of particular significance to our findings and opportunities.

3.4.2. Intervention Types

As illustrated in Figure 3.4, of the 84 AI governance initiatives included in our study, 36 are *programs*, 20 are *policies*, 18 are *strategic plans*, six are *ethics statements*, and four are *standards*. Since 2018, programs have consistently been the most favored type of AI governance intervention at both the federal level and the provincial level, with 25 federal and 11 provincial programs launched to date. In 2019, policy interventions started to become more commonplace at both the federal and the provincial level, with 13 federal policy initiatives and seven provincial policy initiatives launched to date. Strategic plans have been consistently launched across the six years represented in our study, though unlike the other intervention types included in our study, there have been more strategic plans launched at the provincial level than at the federal level (10 provincial strategic plans vs. eight federal strategic plans). Ethics statements and standards are comparatively more scarce: one ethics statement initiative was launched each year between 2017 and 2020, two were launched in 2021, and no new ethics statement initiatives were launched in 2022; after one standards initiative was launched in 2017, no new standards initiatives were launched from 2018-2020, two new standards initiatives were launched in 2021, and one new standards initiative was launched in 2022.

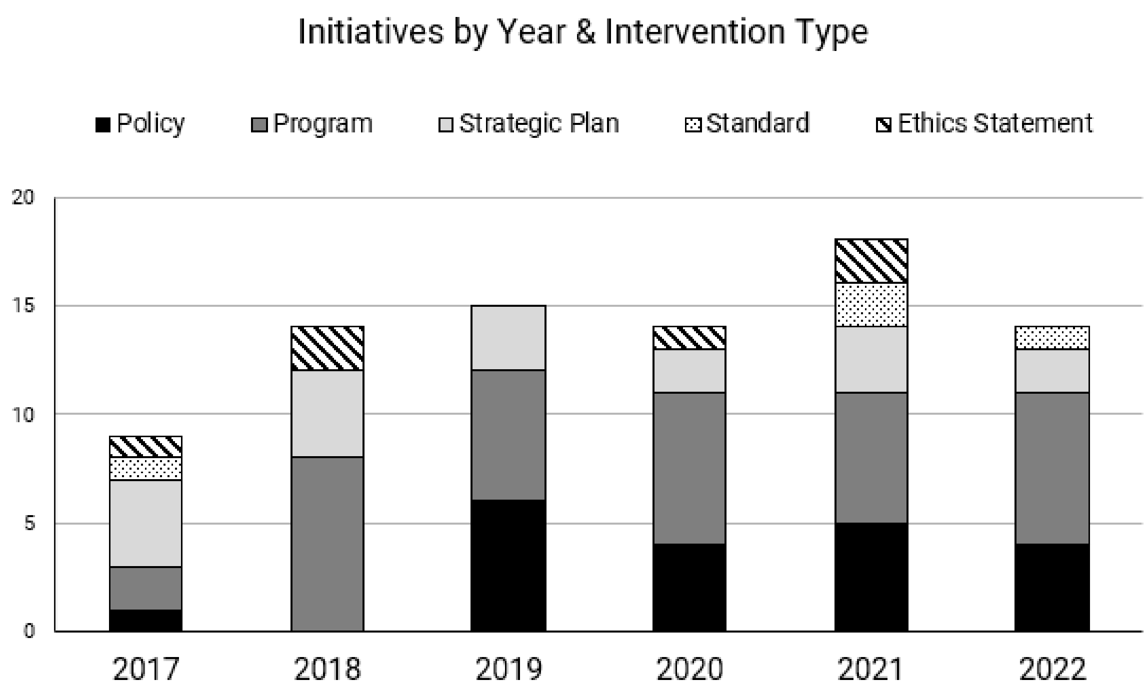

Figure 3.4: Bar chart depicting distribution of intervention types and years of origin across the 84 initiatives included in our study.

3.4.3. Governance Areas

In addition to identifying trends and gaps in intervention types, we are also able to identify several trends and gaps in governance areas. As illustrated in Figure 3.5, of the 84 AI governance initiatives included in our study, 50 seek to intervene in impacts related to *industry and innovation*, 40 in impacts related to *technology production and use*, 28 in impacts related to

AI research, 27 in impacts related to *public administration*, 20 in impacts related to *data governance*, 19 in impacts related to *social and workforce* issues, 10 in impacts related to *AI education and training*, and six in impacts related to *digital infrastructure*. Note that the governance areas we identify are not mutually exclusive–most initiatives seek to intervene in multiple governance areas (see Appendix 3A). With the exception of 2022, industry and innovation has been represented in more initiatives than any other governance area across every year included in our study. Technology production and use as well as AI research have also seen significant representation in most of the initiatives included in our study, with a substantial decline in AI research initiatives in 2021 and a substantial increase in technology production and use initiatives in 2022. The renewal of AI research initiatives in 2022 is attributed to the launch of the Canadian Institute for Advanced Research (CIFAR)'s second phase of the Pan-Canadian AI Strategy, as well as the rollout of multiple new policy and strategy initiatives with a significant research component in Québec and Alberta. There has been relatively consistent year-after-year representation of social, workforce, and public administration issues across all of the initiatives included in the study, sporadic representation of AI education, training, and data governance issues across the years, and comparatively little representation of digital infrastructure issues.

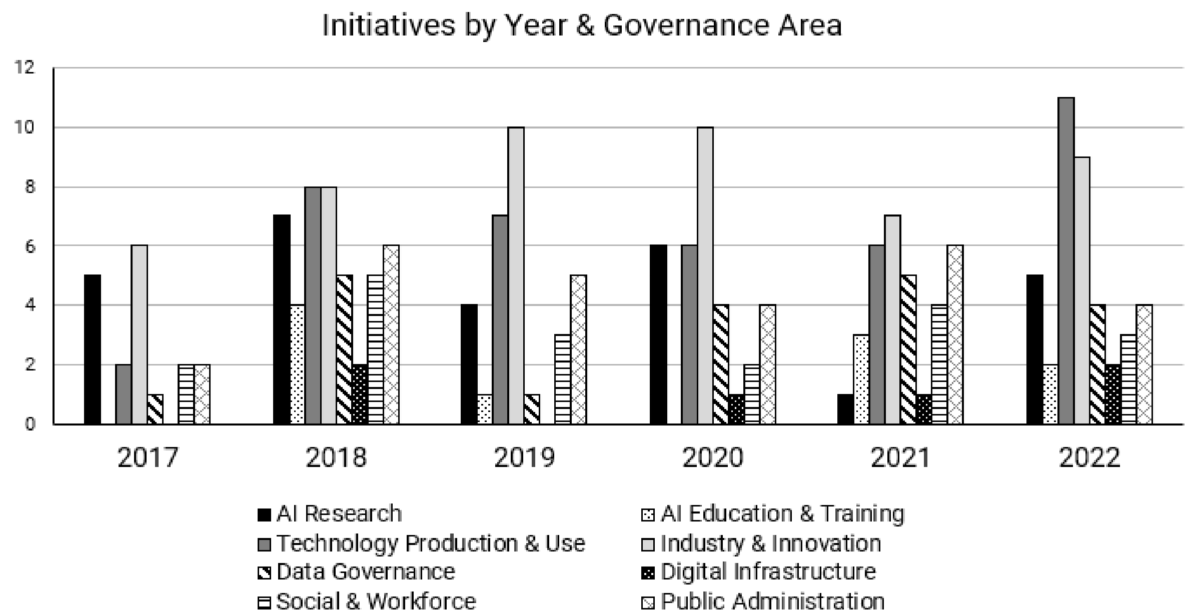

Figure 3.5: Bar chart depicting distribution of governance areas and years of origin across the 84 initiatives included in our study.

3.4.4. Program Initiatives

At present, Canadian AI governance primarily operates through interventions made by government-funded programs. A particularly notable portfolio of programs are those associated with the Pan-Canadian AI Strategy (referred to as initiative #F01 in Appendix 3A). The Pan-Canadian AI Strategy is administered by CIFAR–an independent, non-profit research organization–under appointment from Innovation, Science and Economic Development Canada (ISED), a federal department responsible for a broad portfolio of industry and innovation issues. The first phase of the Pan-Canadian AI Strategy was launched in 2017, which included 4 programs of relevance to our study: the National AI Institutes (initiative #F02 in Appendix 3A), the Canada CIFAR AI Chairs (#F10), the National Program of Activities (#F08), and AI & Society (#F09). The National AI Institutes are a program through which three centers for AI research and innovation have been established: the Vector Institute in Toronto, Ontario; Mila in Montréal, Québec; and Amii in Edmonton, Alberta. The Canada CIFAR AI Chairs program aims to recruit and retain the world's top AI researchers, while the National Program of Activities aims to support training and collaboration on AI development across the country. Through the AI & Society program, CIFAR has led a variety of activities with the goal of enabling responsible AI development, including a series of cross-country policy workshops, the design of AI governance solutions for low-middle income countries, as well as multiple reports on the global landscape of AI governance and responsible AI practices. In 2022, CIFAR launched the second phase of the Pan-Canadian AI Strategy, which in addition to renewing their phase one initiatives, introduced four new programs of relevance to our study: the Pan-Canadian AI Compute Environment (#F48), intended to support the computing infrastructure needs of Canada's AI researchers and developers; AI for Health (#F49), a strategic priority and program intended to support AI-based healthcare innovations; AI for Energy and the Environment (#F50), a strategic priority and program intended to support AI-based energy and environmental innovations; AI Commercialization (#F51), a strategic priority and program intended to support Canadian companies in bringing AI technologies to market.

In addition to CIFAR, National Research Council Canada funds a variety of AI research programs and projects geared toward technical innovation and industrial applications of AI (initiatives #F05, #F23, #F25, #F26, #F34, and #F44 in Appendix 3A). Other notable programs at the federal level include those associated with broader efforts to support AI-based innovation and economic development. For example, ISED's Global Innovation Clusters program (#F07) was launched in 2018 to support the industrial uptake of various AI technologies in the development of five regional economic clusters across Canada, while the International Innovation Program administered by the Canadian Trade Commissioner Service (#F42) has funded a variety of international partnership and technology co-development projects, including

AI development projects. To ensure the personal information of Canadians is protected amidst AI development and innovation, the Office of the Privacy Commissioner of Canada (a nonpartisan agency tasked with researching, auditing, and promoting public awareness of privacy-related issues in Canada's federal government) has also launched consultations on AI and funding for AI-related data governance and privacy research through its pre-existing Contributions Program (#F11) and Consultations Program (#F12). Some federal programs have a more inward-facing focus, such as Shared Services Canada's 2022 Artificial Intelligence Program (#F53) to guide federal departments in "attaining their AI goals through the application of best practices, engagement, capabilities, and technologies," or the Canada School of Public Service's 2022 "Artificial Intelligence is Here" educational series for federal public servants (#F41).

At the provincial level, AI governance programs also strongly focus on issues of economic development and workforce development. In Ontario, the Digital Export Market Development Initiative launched in 2020 (initiative #ON03 in Appendix 3A) supports Ontario's businesses in exporting their digital products and services–including many AI products and services–to foreign markets, while sector-specific and application-specific programs such as the 2019 Smart Grid Fund (#ON10) and 2019 Automated Vehicle Pilot Program (#ON11) support the innovation and adoption of specific AI applications within Ontario. In Alberta, the Major Innovation Fund launched in 2018 (#AB01) supports AI development in areas such as autonomous systems, open data technologies, and smart agriculture, while the 2021 Micro-credential Pilot Program (#AB03) establishes partnerships with industry and educational institutions to support worker training and credentialing in a variety of industries and sectors, including a micro-credential specifically in "machine learning and artificial intelligence." In

Québec, development of the AI workforce is being addressed by a pilot program launched in 2021 (#QC03) that provides permanent immigration to foreign workers in Québec's artificial intelligence and other information technology sectors, while the 2022 Intelligence numérique en éducation program (#QC10) supports Québec's public school system in integrating AI applications and other digital technologies into its administrative and educational activities.

3.4.5. Policy Initiatives

Several policy initiatives also play a vital role in attempting to control the design, development, and use of AI systems in Canada. At the federal level, Bill C-11 (also known as the Digital Charter Implementation Act) was tabled in the Parliament of Canada in 2020 with the goal of legislating new and revised laws concerning consumer data privacy. In 2022, the Digital Charter Implementation Act was re-tabled as Bill C-27 (initiative #F47 in Appendix 3A), this time containing an additional act within it–the "Artificial Intelligence and Data Act"–intended to set rules on the operation of AI systems by individuals and commercial entities. Notably, neither version of the Digital Charter Implementation Act applies to public sector entities. The operation of AI systems in the federal public sector is instead covered primarily by the Directive on Automated Decision-making (#F19), the Algorithmic Impact Assessment Tool (#F20), and the List of Interested Artificial Intelligence Suppliers (#F24), a suite of policy instruments launched by the Treasury Board of Canada Secretariat in 2019 (with the Suppliers List being launched in collaboration with Public Services and Procurement Canada) that are collectively intended to mitigate the risks of automated decision-making in many of Canada's federal institutions.

In addition to their work on drafting the Digital Charter Implementation Act, ISED also led the 2019 launch of Canada's multisectoral Advisory Council on Artificial Intelligence (#F21), intended to provide strategic and policy guidance on AI, and in 2020, a Public

Awareness Working Group (#F38) was added to the Advisory Council with the intent of fostering greater public awareness of and trust in AI technologies. Facial recognition technology (FRT) also represents an area of particular concern to Canadian policymakers: in 2022, Parliament's Standing Committee on Access to Information, Privacy and Ethics held two meetings on the use and impact of FRT (#F52), while the Office of the Privacy Commissioner of Canada (OPC) published a report on FRT guidance for police agencies based on public consultations that were conducted throughout 2021 (#F12). In addition to their public consultation on FRT, the OPC has also conducted a public consultation on AI technologies more generally in 2020 (#F12), the results of which were intended to guide the development of the Digital Charter Implementation Act.

Fewer policy initiatives exist at the provincial level, and those that do exist tend to focus on the use of AI technologies in regional economic development or public administration activities. For example, the Government of Ontario's 2021 Digital and Data Directive (#ON08) aims to ensure "the delivery of high-quality digital services and access to public government data" through the governance of public sector digital assets such as algorithms, datasets, and computational models. In Québec, a committee of cybersecurity and digital experts was formed in 2022 (#QC11) to advise the Minister of Cybersecurity and Digital on policies and strategies related to digital technologies–including AI technologies–with a particular interest in applying those technologies to public administration activities. Meanwhile, initiatives such as "Intellectual Property in Ontario's Innovation Ecosystem" (#ON04) and "Future of Work in Ontario" (#ON09) are intended to recommend policies and best practices to provincial agencies on industry, innovation, and workforce issues, including some concerns related to the potential impacts of AI on Ontario's economy and workers. A notable exception to the economic development and public administration focus of provincial policies is Québec's Law 25, a legislative act first introduced in 2020 under the title "Bill 64" (#QC12) with the goal of amending Québec's data privacy laws. With regard to potential AI impacts, Law 25 has a similar function to the data privacy protections included in the federal Digital Charter Implementation Act, calling for automated decision systems operating in Québec to comply with specific data collection, use, and governance requirements. However, Law 25 notably differs from its federal counterpart in including public sector entities within its scope (Daly & Orct, 2022) and in providing greater protections for privacy rights and human rights outside of commercial contexts (Scassa, 2022a).

3.4.6. Strategic Plan Initiatives

The most significant strategic plan at the federal level is the Pan-Canadian Artificial Intelligence Strategy (#F01), which outlines several high-level national strategic objectives that are being pursued through the various policies and programs of ISED and their partners. ISED is also responsible for Canada's Digital Charter (#F22), a 2019 strategic plan intended to use the findings of a series of national consultations on AI and other digital technologies to implement a series of programs, policies, and other strategic plans, including the Digital Charter Implementation Act (#F47). Another particularly notable set of strategic plans are the international planning initiatives undertaken by ISED leading up to the formation of the Global Partnership on Artificial Intelligence: the 2018 Canada-France Statement on Artificial Intelligence (#F15), the G7 2018 Multistakeholder Conference on Artificial Intelligence (#F13), the 2019 Declaration of the International Panel on Artificial Intelligence (#F27), and the 2020 Joint Statement from the founding members of the Global Partnership on Artificial Intelligence (#F33). Other strategic plans are more narrowly focused on the role of AI technologies in accomplishing the strategic objectives of particular federal agencies, such as the Canada Revenue Agency's 2018 to 2025 corporate business plans (#F18) or the Public Service Commission of Canada's 2018 Integrated Intelligence Evaluation Report (#F17).

More strategic plans have been developed at the provincial level, where strategic planning initiatives are often used to envision the future role of AI technologies in a particular economic sector and to organize provincial resources around the pursuit of AI-enabled economic development within that sector. For example, Ontario's 2019 "Driving Prosperity" initiative (#ON13), Québec's 2017-2022 and 2022-2027 research and innovation strategies (#QC04 and #QC08), as well as Alberta's 2017 Research and Innovation Framework (#AB04) and 2022 Technology and Innovation Strategy (#AB05) all acknowledge the potential economic benefits of supporting the innovation and uptake of AI technologies within particular strategic industries and sectors. In addition to strategic plans related to industry and innovation, Ontario's 2021 "Digital and Data Strategy" (#ON06) and 2017 "Putting Justice Within Reach" (#ON05) also outline several specific strategic objectives for implementing AI and other digital technologies within Ontario's public sector and justice system.

3.4.7. Ethics Statement Initiatives

A notable ethics statement at the federal level is the interdepartmental guiding principles on "Responsible Use of Artificial Intelligence" published jointly by the Treasury Board of Canada Secretariat and Shared Services Canada (initiative #F40 in Appendix 3A). These five guiding principles–(1) understand and measure the impact of AI, (2) be transparent, (3) provide meaningful explanation, (4) be open, and (5) provide sufficient training–are intended to be observed by federal institutions in their production and use of AI technologies. Other notable ethics statements at the federal level include the 2020 Freedom Online Coalition Joint Statement on Artificial Intelligence and Human Rights (#F32) and the 2018 Charlevoix Common Vision for the Future of Artificial Intelligence (#F14). Both of these ethics statements are outputs of international initiatives to establish guidelines on and commitments to ethical practices of AI development and use.

At the provincial level, Ontario's AI Guidance (#ON1) builds upon the findings of a public consultation process that began in 2021 to provide detailed guidelines for the transparent and ethical use of AI technologies within the Ontario public service. Capturing a broader scope of ethical considerations, the 2018 Montréal Declaration (#QC06) arose from a partnership between the Government of Québec, the Government of Canada, as well as several industry, academic, and civil society stakeholders to provide a set of guidelines for the ethical development and use of AI within society more generally.

3.4.8. Standards Initiatives

Although all three standards initiatives at the federal level have yet to produce finalized standards, these initiatives are all multistakeholder efforts led by the Standards Council of Canada (SCC), an organization tasked with promoting voluntary standardization activities. All three initiatives aim to establish shared technical and operational frameworks for the design, development, and use of AI technologies in Canada. The Canadian Data Governance Standardization Collaborative launched by SCC in 2019 (#F43) brings together experts and working groups from government, industry, academia, Indigenous organizations, and civil society with the goal of co-creating standards for the governance of data collection, organization, access, sharing, retention, analytics, and commercialization. More recently, an accreditation pilot for AI management systems launched by the SCC in 2022 (#F46) aims to assess the effectiveness of a selection of AI certification schemes and conformity assessment standards in preparation for developing a larger AI standardization and accreditation program in the future. In addition, SCC continues to organize and recruit experts for a Canadian mirror committee to the international technical committee that has been developing the ISO/IEC JTC 1/SC 42 standards on artificial intelligence since 2017 (#F06). At the provincial level, Ontario's 2021 Digital Service Standard (#ON07) specifies thirteen standardized principles and best practices for service design that provincial agencies are instructed to follow in order to meet the commitments set out by the province's 2021 Digital and Data Directive (#ON08).

3.5. Opportunities for Strengthening AI Governance in Canada

3.5.1. Opportunities for Researchers

Opportunity 1: Researchers should study initiative outcomes. The initiatives included in our review set ambitious objectives and promise beneficial outcomes to the Canadian public. Unfortunately, in reviewing the public-facing documentation of these initiatives that has been made available, the extent to which those objectives are being fulfilled in the actual outcomes of the initiatives is often unclear. This gap in clarity is especially evident in programs, policies, and strategic plans that have long-term objectives and complex outcomes that impact multiple governance areas. While some initiatives such as ISED's Global Innovation Clusters (#F07), Advisory Council on Artificial Intelligence (#F21), and the Pan-Canadian AI Strategy (#F01) regularly report on initiative outcomes relative to performance targets and strategic goals (Accenture & CIFAR, 2020; Innovation, Science and Economic Development Canada, 2022, 2024), other initiatives such as Ontario's Digital and Data Strategy (#ON01) and Alberta's Technology and Innovation Strategy (#AB02) set highly specific long-term objectives and performance measures for AI-related activities, but do not specify any performance targets or provide clear and regular reporting on the extent to which they have succeeded in achieving those objectives. Some initiatives with a significant public administration focus–such as Shared Services Canada's government-wide Artificial Intelligence Program (#F53), or the AI adoption goals described in the Canada Revenue Agency's corporate business plans (#F18)–do not publish any information about specific success measures, performance targets, or initiative outcomes. There are also some initiatives with robust reporting requirements that are only being partially fulfilled. In the case of TBS's Directive on Automated Decision-making and Algorithmic Impact Assessment initiatives (#F19 and #F20), impact assessments for 10 automated decision systems across multiple federal departments were published in the Government of Canada (2024) Open Government Portal at the time of study's conclusion at the end of January 2023, but there has been no public disclosure of any outcomes pertaining to the Directive's requirements for regular system monitoring, regular data validation, regular risk assessment, or regular reporting on the "effectiveness and efficiency" of systems in meeting program objectives. Challenges related to outcomes disclosure have recently been recognized by TBS in a periodic review of the Directive (2022), as well as by multiple scholars who were invited by TBS to comment on the Directive as part of that review process (Attard-Frost, 2022; Brandusescu & Sieber, 2022; Scassa, 2022b). In the case of other policy initiatives such as the federal Digital Charter Implementation Act (#F47) and parliamentary committee meetings on the use of facial recognition technology in policing contexts (#F52), many policy recommendations and criticisms have been raised by the Office of the Privacy Commissioner of Canada (2021a, 2021b, 2020) as well as by observers in academia and civil society (Ifill, 2022; Scassa, 2022a, 2022c, 2022d, 2022e; Witzel, 2022). However, in the case of the Digital Charter Implementation Act, the process through which those recommendations were considered and integrated into policy development is unclear, and as a result, many Members of Parliament and civil society observers voiced concern that the Act does not reflect best practices or public preferences for mitigating harmful AI impacts (Centre for Digital Rights, 2022; Scassa, 2022c, 2022f; Tessono et al., 2022; Williams, 2022; Villemure, 2022).

Further research is required to determine the extent to which Canada's AI governance initiatives have been successful in meeting their stated objectives and to determine the causes and effects of their successes and/or failures. Recognizing that the currently published information on initiative outcomes is significantly limited (a topic discussed at greater length in the following opportunity), potential methods of conducting further research on initiative outcomes could include: key informant interviews with initiative leaders and stakeholders in relevant organizations; submitting freedom of information requests to relevant organizations and analyzing records obtained through the freedom of information process; organizing town hall meetings, workshops, or other participatory research activities to share information on initiative outcomes, identify quality gaps, and discuss potential solutions.

Opportunity 2: Researchers should study challenges to public trust. Many of the initiatives included in our review acknowledge that the Canadian public's trust in AI systems and governance initiatives might be lessened by a variety of gaps in accountability and transparency to the public (e.g., the Directive on Automated Decision-making, as explained in a 2022 internal review by Bitar, Deshaies, and Hall; Canada's Digital Charter [initiative #F22 in Appendix 3A]; the Charlevoix Common Vision for the Future of Artificial Intelligence [#F14]), by gaps in public participation in the design and implementation of AI systems and governance initiatives (e.g., Innovation, Science and Economic Development Canada, 2023; the Advisory Council on Artificial Intelligence [#F21]; Bitar, Deshaies, & Hall, 2022), and by gaps in public awareness of AI development, use, and governance activities (e.g., Innovation, Science and Economic

Development Canada, 2023; 2020-2021 the Advisory Council on Artificial Intelligence [#F21]; Ontario's AI Guidance [#ON01]; the policy labs and 2020 facilitation guide produced as part of CIFAR's AI & Society initiative [#F09]). Those concerns echo similar concerns voiced by observers from academia, industry, and civil society, such as: inadequate public consultation and oversight in the development and proposed application of the Digital Charter Implementation Act (Centre for Digital Rights, 2022; Tessono et al., 2022); insufficient public access to information pertaining to the application and outcomes of the Directive on Automated Decision-making (Attard-Frost, 2022; Brandusescu & Sieber, 2022; Scassa, 2022b); public mistrust of AI systems due to the perceived opacity of their operations, their ability to cause harmful impacts, as well as a lack of clarity and confidence in the accountability mechanisms that might be able to mitigate those harms (Ipsos, 2022; Deloitte, 2019). Taking a broader view of the political economy of Canadian AI governance, Lepage-Richer and McKelvey (2022) reveal challenges to public accountability, transparency, and awareness of AI activities in their characterization of Canadian AI policy as centered around a "discretized, result-driven vision of government, where managerial and political control could be exercised by being outsourced to complex techno-corporate systems" (p. 9). Similarly, empirical studies of Canada's AI innovation ecosystem have characterized it as fraught with complex public-private partnerships that pose significant risks to public accountability and transparency: Brandusescu (2021) has found that in Canada, "public investments in AI technologies primarily benefit the private sector," and that "concentrations of power provide advantages to a handful of entities" (p. 7), while Frost (2020) has found that Canada pursues a "middling strategy of building national consensus on AI governance through knowledge brokerage, co-innovation, and value co-creation" (p. 14) across the public and private sectors.

Further research is required to more precisely identify the public accountability, transparency, and awareness challenges that exist across more AI governance initiatives, as well as to determine potential solutions to those challenges. Overcoming challenges to public trust will be made especially difficult by the federal government's more general administrative incapacity to provide adequate public access to government information and data, a weakness which has long been noted by many observers (e.g., Cardoso, 2022; Clarke, 2019; Information Commissioner of Canada, 2002, 2022; Roberts, 2006; Sieber & Johnson, 2015). However, by establishing a more comprehensive evidence base of specific challenges to public trust associated with specific AI governance initiatives, researchers can enable the policymakers and public servants responsible for those initiatives to identify and address those challenges through evidence-based interventions. We discuss this opportunity for practical intervention at greater length in Section 3.5.2.

Opportunity 3: Researchers should study the effects of impact representation on governance outcomes. Among the 84 initiatives included in our study, federal and provincial governments have clearly prioritized interventions in industry and innovation impacts (50 initiatives) and technology production and use impacts (40 initiatives) over interventions in social and workforce impacts (19 initiatives), education and training impacts (10 initiatives), and digital infrastructure impacts (6 initiatives). Although there are some initiatives intended to address these underrepresented types of AI impact at a national scale–for example, CIFAR's AI & Society initiative (#F09), education and training activities, and Pan-Canadian AI Compute Environment (#F48)–the significant disparity we find in impact representation reflects an industry-first, technology-push approach to AI governance, characterized by Djeffal, Siewert, and Wurster (2022) as a "stimulation approach." This approach of attempting to fulfill societal needs as a by-product of fulfilling industrial needs has also been observed in studies that find Canadian AI governance strongly focuses on intervening in job creation, innovation, and economic development (Babashahiashtiani, 2021; Frost, 2020; Lepage-Richer & McKelvey, 2022), as well as on empowering the private sector through strategic public investments (Brandusescu, 2021). Strategic plans such as the Charlevoix Common Vision for the Future of Artificial Intelligence (#F14) and the G7 Multistakeholder Conference on Artificial Intelligence (#F13) indicate that the AI governance approach of Canada (along with many other nations) rests on long-contested assumptions about the effectiveness of governments at producing broad-based societal benefits through targeted interventions in industries and markets (for discussions of these assumptions, see Harvey, 2005; Scott, 1998). These initiatives assume that by investing public funds to increase industry's ability to competitively produce, use, and innovate new AI technologies, the resulting economic benefits to industry will eventually produce broader societal benefits through a series of top-down market logics. For example, the summary report of the G7 2018 Multistakeholder Conference asserts that if governments can establish "competitive and sustainable industries, institutions, and businesses," then AI will "positively impact societies by contributing to value-added outcomes" (p. 2). In Canada's stimulation approach to AI governance, societal benefit is typically assumed to be a secondary epiphenomenon of economic benefits accrued through technological innovation.

There is no empirical evidence to support such assumptions about the socio-economic impacts of AI systems. In fact, many studies of the political economy of AI indicate that without broad-based and cross-cutting interventions in industries, technologies, societies, workforces, and digital infrastructures, a stimulation approach to AI governance might instead cause negative societal impacts. Without adequate counterbalances, expanding industry's capacity to develop and use AI systems may compound existing concentrations of capital, technology, data, and other resources in a small handful of dominant industry actors (described by Ahmed, Wahed, & Thompson, 2023; Dyer-Witheford, Kjøsen, & Steinhoff, 2019; Lehdonvirta, 2022; Whittaker, 2021), widen existing social, political, and economic inequities (described by Benanav, 2020; Crawford, 2021; Miceli & Posada, 2022; Pasquale, 2020), and create governance gaps or power imbalances that result in the civil society and the public sector becoming overly dependent on AI governance capabilities provided by the private sector (described by Auld et al., 2022; Mazzucato et al., 2022). More empirical research is required to determine if Canada's public investments in AI are actually translating into broader societal benefits, and if not, what negative effects an industry-first approach to AI governance might have on underrepresented governance areas such as society and the workforce, education and training, and digital infrastructure.

3.5.2. Opportunities for Practitioners

Opportunity 1: Policymakers and public servants should specify success measures for initiatives and routinely publish information on initiative outcomes. As discussed in the previous section, researchers and the general public often encounter a lack of clarity in seeking information about the goals and outcomes of Canadian AI governance initiatives. To provide greater clarity, policymakers and public servants can specify precise strategic goals, success measures, and performance targets for their initiatives, as well as routinely publish information on their degree of success in achieving their performance targets and strategic goals. Pursuing this opportunity could entail: following the example of the Ontario Digital and Data Strategy (#ON06), specifying precise strategic goals and success measures across multiple time horizons; following the examples of the Global Innovation Clusters (#F07) and Pan-Canadian AI Strategy (#F01), publishing regular reports on initiative outcomes relative to specific performance targets and goals on websites, information portals, registries, or databases; following the example of the Directive on Automated Decision-making (#F19), building quality assurance and reporting requirements into internal policy instruments, and ensuring those requirements are upheld in practice. In the case of initiatives that purport to be aligned with “ethical” or “responsible” AI principles (such as the principles outlined in the federal government’s guiding principles on responsible AI [#F40], or in Ontario’s AI Guidance [#ON01]), it is especially important to measure and report on the extent to which the initiative’s actual outcomes were aligned with those principles, as well as the interventions that might be required to better align future AI practices with those principles. In addition to potentially benefiting the performance of initiatives, improvements to the clarity, transparency, and rigor of strategic planning and reporting activities could also help practitioners secure public trust in their initiatives, an opportunity we discuss in greater detail below.

Opportunity 2: Policymakers and public servants should collaborate more directly with the public on designing and implementing initiatives. By providing the public with greater access to information about AI governance initiatives and their outcomes, policymakers and public servants can reduce a significant barrier to public trust. However, beyond issues of public awareness and access to information, insufficient public participation in the design and implementation of initiatives presents another opportunity for practitioners to strengthen Canadian AI governance. Recent recommendations made by the AI Public Awareness Working Group of Innovation, Science and Economic Development Canada (2023) emphasize this opportunity: their report suggests that the Government of Canada should make greater efforts not only in “launching sustained and government-led public information campaigns”, but also in "engaging people in Canada in meaningful public dialogues" and "promoting the value and necessity of engagement with citizens" (p. 12).

The initiatives included in our study involve varying degrees of public participation: while most initiatives show no evidence of any public participation in their design or implementation, some initiatives engage with the general public and publish detailed information about the outcomes of engagement (e.g., the Office of the Privacy Commissioner of Canada's consultations on AI and facial recognition technologies [#F12]; Ontario's AI Guidance [#ON01]), while other initiatives publish detailed information about the outcomes of their public engagement processes, but only engage with a relatively small, targeted group of stakeholders (e.g., the Directive on Automated Decision-making and Algorithmic Impact Assessment [#F19 and #F20]; CIFAR's AI & Society workshops [#F09]; Canada's Digital Charter [#F22]). However, across all of these initiatives, public participation is limited to informing decision-makers rather than directly collaborating with decision-makers on initiative design and implementation. This can present barriers to public accountability, transparency, and trust in situations where there appears to be a significant discrepancy between the information provided by the public and the initiative's eventual design and implementation. For example, despite having been informed by the Digital Charter consultations as well as by consultations conducted by the Office of the Privacy Commissioner of Canada, many observers have commented that the Digital Charter Implementation Act [#F47] does not adequately reflect the findings of public consultation processes (Centre for Digital Rights, 2022; Office of the Privacy Commissioner of Canada, 2021b; Scassa, 2022e; Tessono et al., 2022).

There is a risk that public engagement processes that do not faithfully represent the public's contributions in their outcomes may be viewed as tokenizing, performative, or selflegitimizing, and ultimately jeopardize public trust rather than secure it (Adu-Daako & Sieber, 2022; Arnstein, 1969; Cornwall, 2008). To mitigate that risk, policymakers and public servants can apply best practices and principles for public participation in AI design and governance as described by researchers (e.g., Birhane et al., 2022; Krafft et al., 2021; Sieber & Brandusescu, 2021) as well as by advisory and advocacy organizations (e.g., Law Commission of Ontario, 2021; RSA, 2018) to more directly, continuously, and transparently collaborate with the Canadian public on long-term practices of initiative design and implementation. In a comparative study of 16 national AI strategies, Wilson (2022) finds "little evidence that public engagement values and mechanisms are salient in the consolidation of AI governance regimes" (p. 8). By taking this opportunity to strengthen public participation in AI governance, Canada could fill this international gap and position itself as the global leader in participatory AI governance.

Opportunity 3: Policymakers and public servants should account for a greater variety of AI impacts when designing and implementing initiatives. Canada has yet another opportunity to show global leadership by adopting a more expansive view of AI impacts across its diverse range of AI governance initiatives. Many researchers have noted that because "artificial intelligence" is such an ambiguous concept, different interpretations of what types of phenomena constitute "AI" and what the impacts of AI therefore involve will cascade into different practices of AI governance (Büthe et al., 2022; Nitzberg & Zysman, 2022; Crawford, 2021; Mishra, Clark, & Perrault, 2020). The ambiguity of "AI" and its wide range of potential impacts poses a complex and multicontextual challenge for AI governance practices, particularly in designing policies, standards, and ethics statements that establish a concrete definition of "AI" as their epistemic grounding and build their principles, provisions, and requirements upon that grounding. For example, the 2022 Digital Charter Implementation Act defines an "artificial intelligence system" as a "technological system that, autonomously or partly autonomously, processes data related to human activities through the use of a genetic algorithm, a neural network, machine learning or another technique in order to generate content or make decisions, recommendations or predictions" (Parliament of Canada, 2022). This definition–with its scope narrowed to technological phenomena and informational inputs and outputs–is bounded off from many potential application contexts in which the system and its components may be developed or used (Centre for Digital Rights, 2022), from much of the human knowledge and labor required to develop and use the system (as discussed by Miceli & Posada, 2022; Miceli, Posada, & Yang, 2022; Sambasivan et al., 2021), as well as from the potentially intensive energy, water, and other material resources consumed by the system and the subsequent environmental impacts caused by the system (as discussed by Bender et al., 2021; OECD, 2022c; Ligozat et al., 2022; GPAI, 2021; Patterson et al., 2021; Strubell, Ganesh, McCallum, 2019).

Policymakers and public servants therefore have an opportunity to expand their practices of governing the production and use of AI technologies to identify, analyze, and intervene in more potential impacts associated with the development and use of the technologies. One approach to this expansion could involve integrating or cross-referencing standards that are intended to govern data or digital services (e.g., the Canadian Data Governance Standardization Roadmap [#F43], Ontario's Digital Service Standard [#ON07]) with other initiatives that are intended to govern the production and use of AI technologies (e.g., the Standards Council of Canada's accreditation pilot for AI management systems [#F46], Ontario's AI Guidance [#ON01) to create more comprehensive frameworks for AI standardization that combine data standards, algorithm design and use standards, along with other digital technology and service standards. Cross-referencing between different types of interventions (e.g., referencing the requirements of an AI standard within a policy instrument such as the Digital Charter Implementation Act) could also help to create more expansive and adaptable governance frameworks as AI technologies continue to evolve. To account for and intervene in a broader range of social, material, and environmental impacts, practitioners could apply principles, impact assessment frameworks, and standards for fair digital labor practices (e.g., Cole et al., 2022; Fairwork, 2023; GPAI, 2022; Institute for the Future of Work, 2023), sustainable AI (e.g., Gupta, 2021; Wilson & van der Velden, 2022; Wu et al., 2022), and ethical sourcing in other industries (e.g., Fair Trade Federation, 2023; SQF Institute, 2017) to their design and implementation of future ethics, standards, and policy initiatives. In addition, with support from CIFAR's AI & Society initiative, an extensive set of Indigenous perspectives, principles, and practices for governing many different types of AI systems and potential AI impacts has been produced (Lewis et al., 2020). Policymakers and public servants should now collaborate with Indigenous peoples to design and implement AI governance initiatives that apply those and other Indigenous perspectives, principles, and practices (e.g., Abdilla et al., 2021; Irwin & White, 2019; Lewis et al., 2018; Munn, 2023; Ricaurte, 2022).

Opportunity 4: Policymakers and public servants should launch a new initiative to cultivate a more unified national approach to AI governance. The centralization of responsibility for large-scale national initiatives for governing AI such as the Pan-Canadian AI Strategy (#F01), Global Innovation Clusters (#F07), Canada's Digital Charter (#F22), and the Digital Charter Implementation Act (#F47) into a single federal department–Innovation, Science, and Economic Development Canada (ISED)–seems intended to support a nationally unified approach to AI governance. However, aside from the challenges to public accountability, transparency, and participation we previously described regarding ISED-led initiatives, our findings also indicate that over-reliance on ISED for administering multiple AI governance initiatives of great national significance may limit Canada's ability to effectively intervene in a wide range of AI impacts. In describing the AI governance initiatives administered by ISED, the Office of the Prime Minister's 2021 mandate letter to the minister responsible for ISED centers issues of innovation, technological and economic development, and consumer protection. Consequently, the need for initiatives that mitigate potential AI-based impacts to workers, the environment, human rights, vulnerable groups, and public administration (such as the public consultations and parliamentary meetings on the use of facial recognition in policing, as well as internal policy instruments such as the DADM and AIA tool) are not discussed in any detail in ISED's mandate letter, largely placing these types of AI impacts outside the scope of ISED's mandated AI governance activities. This has contributed to a highly fragmentary and loosely coordinated approach to AI governance in Canada that stands in contrast to regions such as the European Union, where initiatives such as the European AI Strategy (2018) and AI Act (2021) account for a wide range of potential AI applications and impacts, and have been developed through extensive intergovernmental collaboration. The decentralized structure of Canadian AI governance is largely a reflection of Canada's more general federalist model of governance, which in pursuing national unity, is flexible and adaptable so as to accommodate the interests of and gain legitimacy from Canada's many geographic regions and identity groups, but also prone to intergovernmental conflict and ineffective policy collaboration and coordination (Bakvis & Skogstad, 2020; Cameron & Simeon, 2002; Skogstad & Wilder, 2020). Researchers have observed similar dynamics of legitimacy-seeking, intergovernmental conflict, and ineffective collaboration in Canadian AI governance practices (Frost, 2020; Lepage-Richer & McKelvey, 2022), indicating an opportunity for policymakers and public servants to cultivate a more unified national approach to AI governance by designing initiatives that are intended to integrate resources and perspectives from a more diverse range of stakeholders, reduce barriers to collaboration and coordination, resolve conflicts, and build greater legitimacy.

To pursue this opportunity, Canadian policymakers and public servants can look to other national and supranational contexts for examples of AI governance initiatives that are currently being used to facilitate resource integration, collaboration, conflict resolution, and legitimacy-building amongst large, federated groups of public sector, private sector, and civil society stakeholders, such as the United States National AI Initiative (2023) and the European AI Alliance (2023). Similarly, implementing the recent recommendation made by the AI Public Awareness Working Group of Innovation, Science and Economic Development Canada (2023) to create a "national AI Community of Practice" would support researchers, practitioners, and the general public in coordinating their governance efforts and collaborating on initiatives.

3.6. Limitations

Although our findings and proposed opportunities represent a timely and broad assessment of key patterns and gaps across Canada's system of AI governance, our study was limited by three categories of epistemological and methodological constraints:

1) *Fluidity of field:* Canada's system of AI governance is rapidly and continuously evolving, with departments and governments taking increasingly greater interest in AI governance, new initiatives frequently being launched, as well as existing initiatives frequently being changed. Though we endeavored to study a large selection of initiatives that are broadly representative of the state of Canadian AI governance during the period of our content analysis (January 2022 to January 2023), our review and analysis are not intended to be exhaustive and

may have minor gaps in timeliness and completeness, particularly in the many emerging AI governance initiatives in government with a public administration focus.

2) *Choice of sources*: Our primary sources are limited to web content and information that has been publicly disclosed by the organizations responsible for the initiatives included in our study. Due to the public access barriers we described in Section 3.5, our reliance on publicly available information limited our ability to analyze the design features, implementation processes, and outcomes of many initiatives. Additionally, although we reviewed and analyzed primary sources pertaining to all of the 84 initiatives included in our study, our reporting on findings and opportunities focused on a relatively smaller subset of initiatives that–due to their design features, implementation processes, and strategic goals–are most representative of particular patterns and gaps we noted during our analysis.

3) *Organizational scope:* Our study is limited to initiatives launched or contracted by public sector organizations within the federal government and three provincial governments. Provincial and territorial governments with relatively nascent systems of AI governance (e.g., British Columbia) or without any significant AI governance initiatives were excluded from our study. Additionally, although many municipal and private sector AI governance initiatives have been launched across Canada (e.g., Toronto Police Services Board, 2022; Digital Governance Standards Institute, 2021; the municipal AI lab of the City of London, Ontario, as described by Towards Data Science, 2021), we excluded these initiatives from our study. These exclusion decisions enabled us to focus our review and analysis on larger regions in which more AI governance initiatives existed and were further along in their implementation.

3.7. Conclusions

Our objectives in this study were to review a diversity of Canadian AI governance initiatives and synthesize our findings into a set of opportunities for strengthening Canada's system of AI governance. With those objectives fulfilled, we are able to identify the contributions of our study and the future directions for AI governance research and practice revealed by our study.

This study makes an empirical contribution to the research literature on Canadian AI governance. To date, this body of literature has typically focused on the design and implementation of a small handful of flagship federal policies and programs (in particular, the Digital Charter Implementation Act, Directive on Automated Decision-making, and Pan-Canadian AI Strategy), rather than a more comprehensive synthesis of many types of interventions in a diversity of AI impacts across multiple governments and departments.

In addition, this study makes a methodological contribution to AI governance research more generally: a macro-scale review and synthesis of this kind has not been conducted to study the AI governance system of any nation. Although our research design relies on a significant degree of researcher judgment and pre-existing knowledge of the Canadian context, many components of our research process and methodology could be adapted to similar studies of AI governance in other contexts.

Finally, this study also makes a practical contribution to AI governance researchers and practitioners in Canada and elsewhere who may benefit from enacting our proposed opportunities for improving Canadian AI governance. Although the opportunities we suggest are centered around the Canadian context, if similar challenges related to initiative outcomes, public trust and participation, impact representation, and national unification are encountered in other

AI governance contexts, it is possible that some of the opportunities for improvement that we identify in Canada may also be applicable to those contexts.

Chapter 4

AI Governance Systems in Canada

About this article: This article was co-authored with Kelly Lyons beginning in January 2022 and was originally titled *AI Governance Systems: A Multi-scale Analysis Framework, Empirical Findings, and Future Directions*. The original article was published in *AI and Ethics* in September 2024 and can be found at: https://link.springer.com/article/10.1007/s43681-024-00569-5

Abstract

The governance of artificial intelligence (AI) has recently emerged as a field of research and practice, but the structural and functional components of AI governance (AIG) systems are not well understood by researchers and practitioners. To address that gap, we apply service system analysis methods and thematic analysis methods to develop a novel framework for conceptualizing and analyzing AI governance systems across multiple scales of activity, including international, national, subnational, sectoral, and organizational systems of governance. We apply our analysis framework to an empirical study of Canada's national AIG system. Drawing upon qualitative data collected from 20 leaders of Canadian AIG initiatives and subject matter experts, we identify and discuss the actors, impacts, resources, networks, activities, logics, norms, and rules involved in structuring and operating a national AIG system, using Canada as a case study. Based on our findings, we propose three directions for future research: (1) conduct additional analysis of the 610 topics in our dataset, (2) further investigate institutional and ecosystem-level structures and dynamics in Canada's national AIG system, (3) apply our framework, data, and findings to study AIG systems in other contexts. We also outline four strategic objectives for strengthening Canada's AIG system: (1) implement new collaboration and coordination mechanisms, (2) create guidance for designing and implementing participatory AIG initiatives, (3) expand access to key resources needed for effective AIG practices, (4) advance diversity, equity, and inclusion in AIG activities.

Keywords: Artificial intelligence, governance, policy, strategy, expert interviews

4.1. Introduction

AI governance is often characterized as a practice intended to maximize the beneficial impacts and minimize the harmful impacts of artificial intelligence (AI) systems (Dwivedi et al., 2021; Nitzberg & Zysman, 2022; OECD, 2022; Taeihagh, 2021). As a consequence of what Suchman (2023) calls the “strategic vagueness” of the term “AI,” there is no universally agreed upon definition of *AI system*, nor is there universal agreement on the types of actors, activities, benefits, or harms that constitute practices of *AI governance*. According to Suchman, “AI” is interpreted differently–and with different strategic goals–depending on the social, political, economic, and organizational contexts in which it is perceived. Despite the strategic vagueness of “AI,” practices of AI governance in any context can be understood as constituting an *AI governance system* (Mäntymäki et al., 2022). As a system, AI governance (AIG) is an interdependent set of components that are (1) situated in a context, (2) structured by the perceptions of “AI” that exist within that context, and (3) intended to maximize benefits and minimize harms that actors within that context perceive as being caused by “AI.” AIG systems operate within and across many different contexts and scales of activity, including systems of international AIG (Schmitt, 2022; Tallberg et al., 2023; Veale, Matus, & Gorwa, 2023), national AIG (Attard-Frost, Brandusescu, & Lyons, 2024; Liebig et al., 2023; Radu, 2021; Wilson, 2022), municipal AIG (Kinder et al., 2023; Wan & Sieber, 2023), sectoral AIG (Peretz-Andersson et al., 2024; Stahl, 2022; Zuiderwijk, Chen, & Salem, 2021), and organizational AIG (Cihon, Schuett, & Baum, 2021; Mäntymäki et al., 2022; Maragno et al., 2023).

Although AIG is a fast-growing field of research and practice, AIG is not well understood or studied as a *systemic* phenomenon composed of several interdependent structural and functional components. To correct this knowledge gap, Birkstedt et al. (2023) suggest broadening the focus of AIG research “from organizations to networks and ecosystems” (p. 157), noting that “in future research the entire ecology or network of AIG stakeholders should be studied to understand the roles and functions of stakeholder groups” (p. 156). Mishra, Clarke, and Perrault (2020) note another gap in knowledge of AIGS, describing the resource inputs and outputs of AI systems as “difficult to measure as they are often service-based and their intangibility remains a measurement challenge” (p. 2). Birkstedt et al. and Mishra, Clarke, and Perrault find that the most significant knowledge gaps in AIGS research relate to specific system components: uncertainties regarding the networks, ecosystems, stakeholder roles, functions, services, intangible resources, and measurements involved in AIGS. These gaps indicate a need for more integrative theories, analysis methods, and empirical studies of the components of AIG systems across multiple contexts and scales of activity.

In this paper, we address those gaps by integrating a novel set of theories and research methods into an analysis framework, and by applying that framework to an empirical study of a national AIGS. Our study is motivated by three research questions:

RQ1 – Ontological Question: What components exist within an AIGS?

RQ2 – Analytical Question: How can those components be empirically studied within an AIGS?

RQ3 – Ecological Question: How can those components be empirically studied in a macro-scale AIG system that contains a population of AIG systems within it, such as a national AIGS?

To investigate those three research questions, we conducted a case study of Canada’s national AIGS. As the first country to launch a national AI strategy in 2017 (Radu, 2021), Canada is especially notable for the complexity and maturity of its AIG initiatives relative to the AIG initiatives of many other nations (Attard-Frost, Brandusescu, & Lyons, 2024). The relative complexity and maturity of Canada’s AIGS makes Canada an ideal case for a study of AIG practices across multiple contexts and scales of activity. In the following section, we describe the theoretical background for our study of Canada's AIGS. In section 3, we contextualize our study by summarizing the significant federal AIG initiatives and organizations in Canada. In section 4, we describe the research design and methodology of our study. In section 5, we present findings that emerged from interviews with 20 government leaders and subject matter experts on Canadian AIG. A full dataset containing 610 topics aggregated from across the interviews can be found in Appendix 4A. In section 6, we acknowledge the limitations of our study. In section 7, we discuss the implications of our study for future AIG research and practice. We conclude in section 8 with a brief reflection on the results and contributions of our study.

4.2. Theoretical Background

4.2.1. Artificial Intelligence

Although some researchers, practitioners, and policymakers attempt to essentialize the precise phenomena that are constitutive of AI systems–such as machine learning and machine cognition, technological embeddedness, and autonomy from human actors (Dwivedi et al., 2021; Madan & Ashok, 2023; Nitzberg & Zysman, 2022; OECD, 2024)–others argue that foundational concepts of "intelligence" and "AI" are ontologically ambiguous, contextually contingent, and developed through shared perceptions and social practices (Ashok et al., 2022; Attard-Frost, 2023; Blili-Hamelin, Hancox-Li, & Smart, 2023; Suchman, 2023). The vast scale of AI systems further compounds those ambiguities, contingencies, and social complexities. The actors, activities, benefits, harms, and other phenomena involved in AI systems and their governance are distributed across many scales of human activity (Attard-Frost, Brandusescu, & Lyons, 2024; Crawford, 2021; Crawford & Joler, 2018; Ricaurte, 2022), ranging from micro-scalar phenomena (e.g., localized individuals, organizations, and communities) to macro-scalar phenomena (e.g., nations, global value chains, and ecosystems). Studying AI systems and AIG practices across contexts and scales of activity therefore requires highly flexible theories and methods of systems analysis that are adaptable to ontological ambiguities, contextual contingencies, and stakeholder perceptions across multiple scales of human activity.

4.2.2. AI Governance Systems at Micro and Macro Scales

Mäntymäki et al. (2022) argue that an organization's AIG practices can be understood as constituent components of an *AIG system*:

> AI governance is a system of rules, practices, processes, and technological tools that are employed to ensure an organization's use of AI technologies aligns with the organization's strategies, objectives, and values; fulfills legal requirements; and meets principles of ethical AI followed by the organization . . . AI governance is a system whose constituent elements should be interlinked to form a functional entity (p. 604).

In this systemic perspective, an AIGS is composed of an interdependent set of structural and functional components, all of which are situated within an organizational context and influenced by values and normative perceptions of AI within that context. Other researchers have also observed organizational AIG components through a systemic perspective, such as AIG resources, networks, and ecosystems (Birkstedt et al., 2023; Cihon, Schuett, & Baum, 2021; Mishra, Clarke, & Perrault, 2020)

In addition to the scale of individual organizations, AIG components (e.g., resources, networks, and ecosystems) have been empirically studied within larger scale systems, such as the AIG systems of economic sectors, geographic regions, nations, and international groups (Liebig et al., 2022; Peretz-Andersson et al., 2024; Veale, Matus, & Gorwa, 2023; Wilson, 2022). These larger scale AIG systems contain within them populations of smaller scale AIG systems, such as the AIG systems of individual organizations within a sector, or the AIG systems of subnational regions within a nation. This multi-scalar conceptualization reflects a long-standing perspective on scale in the organizational research literature, in which the *micro-organizational scale* refers to activity within an individual organization, while the *macro-organizational scale* refers to the activity of a population of individual organizations that interact with one another (Parsons, 1956; McKelvey & Aldrich, 1983; Vibert, 2004; Kuhn, 2012). As systems that span micro- and macro-organizational scales, rigorous analysis of AIG systems and their components requires analysis methods that can account for macro-organizational contexts while also accounting for the micro-organizational contexts contained within the macro scale. In their study of interactions between national and subnational AIG initiatives in Germany, Liebig et al. (2022) demonstrate that a multi-level perspective is needed to analyze macro-scale systems that contain multiple levels of smaller scale systems within them, such as Germany's national AIG system that contains within it the AIG systems of the nation's subnational regions, economic sectors, and individual organizations.

4.2.3. Service Systems

To conduct our study of Canada's national AIGS, we applied an analysis framework that was developed by Frost, Cheng, and Lyons (2019) for the purpose of analyzing *service systems* across multiple levels of organization. Service systems are interdependent sets of structural and functional components through which service-based organizations integrate resources, co-create value, and engage in governance. The framework was developed through two systematic reviews of the research literature on service systems components and analysis methods (Frost & Lyons, 2017; Lyons & Tracy, 2013), making the framework uniquely capable of accounting for both a breadth and depth of service system components and component types. The framework contains five high-level system structures, seven types of service system components, and 33 sub-types of system components. The framework also provides a questionnaire and a 13-step analysis sequence for applying the framework to empirical studies of service-based organizations.

The service systems analysis framework was originally developed for context-independent analysis of micro-scale organizational contexts (e.g., activity within companies, government departments, research institutes) rather than macro-scale organizational contexts that contain a population of organizations within them (e.g., governments, sectors, nations). We therefore adapted the framework to the context-specific structures and gaps in our particular research context: the macro-scale organization of Canada's national AIGS. The adapted framework contains 12 analytical dimensions and 12 corresponding interview questions that we expected would enable context-sensitive interview data collection and analysis of Canada's national AIGS. The process of our framework design, data collection, and analysis is described in more detail in section 4.

4.3. Research Context: AI Governance in Canada

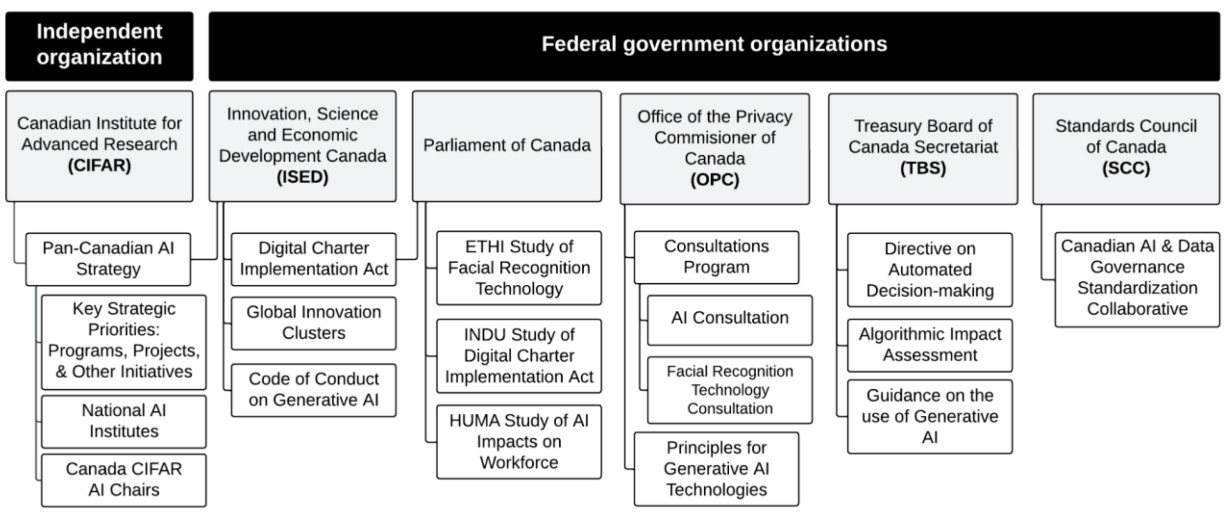

Figure 4.1: Key Canadian federal organizations, AIG initiatives, and lines of responsibility as of May 2024.

AIG is practiced across many scales in Canada, but the AIG initiatives created by Canada's federal government are especially influential in directing Canada's national system of AIG (Attard-Frost, Brandusescu, & Lyons, 2024; Brandusescu, 2021; Frost, 2020; Lepage-

Richer & McKelvey, 2022). Figure 4.1 depicts Canadian federal organizations that have a significant role in Canada's AIGS along with the AIG initiatives they are responsible for. In this section, we briefly describe these organizations and their AIG initiatives to provide contextual background for the research design, findings, and implications we present in subsequent sections.

Innovation, Science and Economic Development Canada (ISED): ISED is a federal department responsible for a sizable portfolio of industry and innovation issues. ISED has a broad mandate that includes two key features: advancing the growth of the Canadian economy, and building a fair marketplace for Canadian industry (ISED, 2024). This dual mandate has resulted in ISED occupying a highly central role in Canada's national AIGS, with ISED leading many AIG initiatives that are intended to both promote the growth of the Canadian AI industry and to create a regulated marketplace for the Canadian AI industry. ISED has undertaken many AIG initiatives in support of this dual mandate, but the two most notable are the Pan-Canadian AI Strategy (PCAIS) launched in 2017, and the Artificial Intelligence and Data Act (AIDA) tabled in Parliament in 2022 as one of three parts of a larger bill entitled the Digital Charter Implementation Act). Reflecting ISED's dual mandate, the PCAIS is intended to advance Canadian AI research, AI innovation, and the commercialization and adoption of AI applications across Canadian industry (ISED, 2022), while the AIDA is intended to establish a cross-sectoral framework for regulating commercial activities involving AI systems throughout Canada (Parliament of Canada, 2024a). In 2023, ISED created another new policy instrument: a Voluntary Code of Conduct intended to support organizations across Canada in responsibly developing and managing generative AI systems (ISED, 2023a). ISED is also responsible for the Global Innovation Clusters, an initiative launched in 2018 to support innovation in five regional economic clusters across Canada, including innovations in AI technologies (ISED, 2023b).

Canadian Institute for Advanced Research (CIFAR): CIFAR is an independent, nonprofit research organization. Under appointment from ISED, CIFAR is responsible for administering the PCAIS along with the many programs, projects, and other initiatives that fall within the scope of the PCAIS's strategic priorities (CIFAR, 2024). The PCAIS also includes three centers of research and innovation known as the National AI Institutes: the Vector Institute in Toronto, Mila in Montréal, and Amii in Edmonton, Alberta. Each of the three national AI institutes hosts leading AI researchers who are funded and supported by CIFAR's Canada CIFAR AI Chairs program, in addition to managing a range of other programs for AI research and development, training and education, and business services.

Parliament of Canada: Three standing committees of the Parliament of Canada have undertaken studies of AI systems and their societal impacts. The Standing Committee on Access to Information, Privacy and Ethics (ETHI) completed a study of the use and impact of facial recognition technologies in October 2022 (Parliament of Canada, 2022). The Standing Committee on Industry and Technology (INDU) launched a study of the AIDA and the larger Digital Charter Implementation Act it is part of in September 2023 (Parliament of Canada, 2024b). The Standing Committee on Human Resources, Skills and Social Development and the Status of Persons with Disabilities (HUMA) launched a study of the potential impacts of AI systems on the Canadian labour force in October 2023 (Parliament of Canada, 2023).

Office of the Privacy Commissioner of Canada (OPC): OPC is an independent federal agency responsible for overseeing the enforcement of privacy laws and regulations in Canada's public sector and private sector, as well as for conducting research and promoting public awareness of privacy-related issues. Through their various programs and projects, OPC has funded research on the privacy impacts of AI and launched public consultations on AI (Office of the Privacy Commissioner of Canada, 2018, 2022). In December 2023, OPC published guidance on principles and practices for protecting privacy in the development, deployment, and operation of generative AI systems (Office of the Privacy Commissioner of Canada, 2023).

Treasury Board of Canada Secretariat (TBS): TBS is an agency responsible for providing the federal public service with guidance on issues of public finance, management, and regulation. TBS has created a portfolio of policy instruments that are intended to mitigate the risks of automated decision-making and generative AI applications across many federal institutions, including the Directive on Automated Decision-making and Algorithmic Impact Assessment Tool put into force in 2019 and updated in 2023 (Government of Canada, 2023a, 2023b), and the voluntary Guide on the use of Generative AI published in 2023 and updated in 2024 (Government of Canada, 2024).

Standards Council of Canada (SCC): SCC is a state-owned enterprise tasked with promoting the development and implementation of voluntary standards across Canada. In July 2023, SCC announced that it would merge many of its existing AI and data governance initiatives into a new initiative entitled the AI and Data Governance Standardization Collaborative. The Collaborative is intended to build awareness of AI standardization activities, harmonize domestic and international AI standards, and develop AI standards that support Canadian AI adoption, innovation, and Indigenous leadership (Standards Council of Canada, 2023).

4.4. Research Design & Methodology

We conducted our study in five phases: (1) framework design, (2) data collection, (3) primary data analysis, (4) quality assurance, and (5) secondary data analysis. The activities involved in each phase of our study are described in detail in the following sub-sections and illustrated in Figure 4.2.

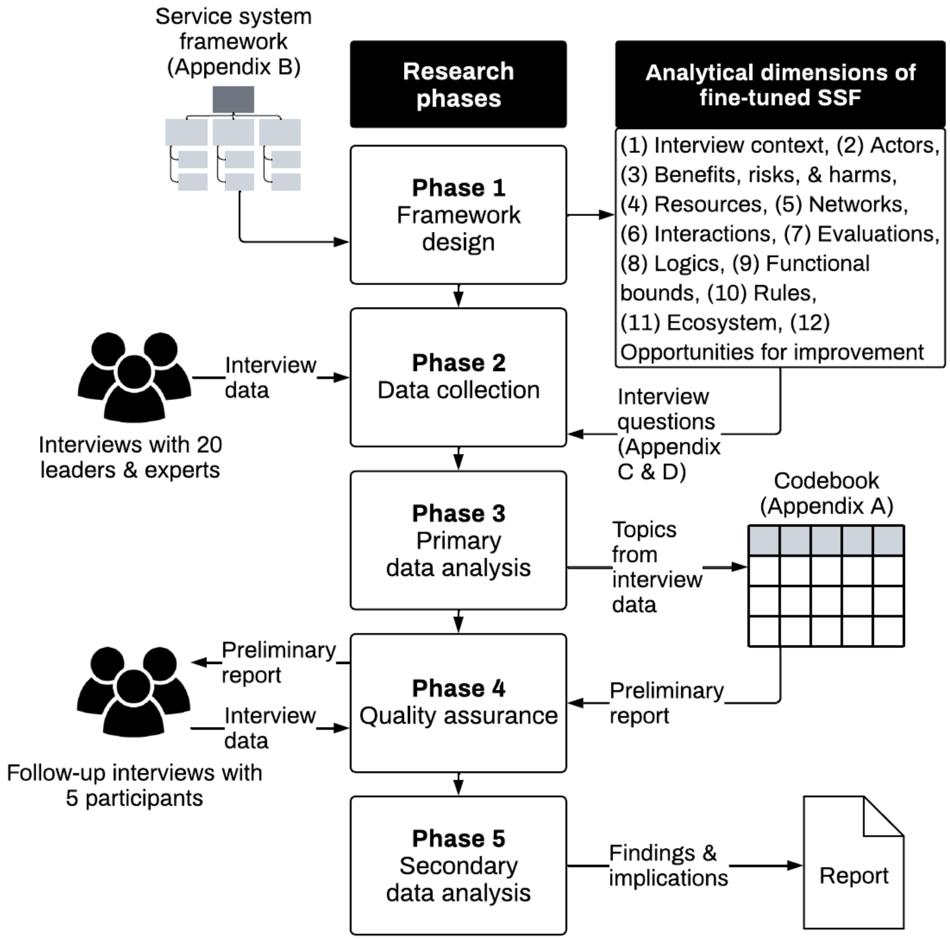

Figure 4.2: Diagram of the research phases and activities involved in our study.

Phase 1: Framework Design

We structured our study's data collection and analysis activities in relation to the 12 analytical dimensions of the service systems framework that were described in section 2: (1) interview context, (2) actors, (3) benefits, risks, & harms, (4) resources, (5) networks, (6) interactions, (7) evaluations, (8) logics, (9) functional bounds, (10) rules, (11) ecosystem, (12) opportunities for improvement. Diagrams of our framework and a process model of how the AIG system structures and activities contained in our framework relate to one another can be found in Appendix 4B. Our interview questionnaires based on the framework can be found in Appendices 4C and 4D.

Phase 2: Data Collection

From February 2023 to July 2023, we conducted 20 interviews with government leaders and subject matter experts from across Canada. Participants were only included in the study if their leadership role met our inclusion criteria of *active leadership*, *initiative impact*, and *knowledge breadth & depth*, or if their subject matter expertise met our inclusion criteria of *knowledge breadth & depth* and *public voice* (see Appendix 4E for more details about these inclusion criteria). Of the 20 participants, nine were leaders of federal, provincial, or municipal AI governance initiatives, seven were subject matter experts employed within the private sector, and four were subject matter experts employed within the academic or civil sectors. During each of the 20 interviews, we asked participants 12 questions that were intended to elicit their perceptions of the most significant phenomena involved in Canada's AIGS. Each of the 12 questions corresponds to one of the 12 analytical dimensions of our framework. For example, participants were asked questions such as: "What AI governance initiatives in Canada are you aware of?", "Which organizations, groups, industries, communities, or other types of social actors are involved in or affected by those initiatives?", and "What resources do those actors require in order to be involved in those initiatives?" During interviews, we collected qualitative data in the form of textual notes and audio recordings. Data was anonymized during collection and throughout the remainder of the study to encourage participants to share their perceptions more candidly. The full questionnaires that were used to guide each interview can be found in Appendix 4C.

Phase 3: Primary Data Analysis

From July 2023 to October 2023, we conducted our first analysis of the textual notes and audio recordings that were collected over the course of the 20 interviews. During this primary analysis phase, we applied "conventional content analysis" methods of inductive coding and iterative category development, as described by Hsieh and Shannon (2005). These methods enabled us to code key topics from across the interview data, develop categories and typologies, and identify quantitative trends and gaps in the data. The topics, categories, typologies, and quantitative data that emerged from our analysis were recorded in a codebook (see Appendix 4A for the final dataset we produced from our codebook). We also applied the six trustworthiness criteria for thematic analysis described by Nowell et al. (2017) to ensure that our analysis process and our documentation of the process (provided in this section and in our appendices) were credible, transferable, dependable, confirmable, auditable, and reflexive.

Phase 4: Quality Assurance

From October 2023 to December 2023, we conducted a second round of interviews with a subset of five of the original 20 participants. We again collected qualitative data during interviews in the form of textual notes and audio recordings. We selected participants for inclusion in this subset based on two criteria: (1) their leadership roles or subject matter expertise required them to have knowledge of the topics, categories, trends, and gaps that most frequently emerged during the primary data analysis phase, and (2) their leadership roles or subject matter expertise required them to have knowledge of many public sector, private sector, and civil society perspectives on AI governance. The purpose of these follow-up interviews was to determine if those topics, categories, trends, and gaps constituted an accurate and reasonably complete representation of Canada's AIGS. Prior to each follow-up interview, we distributed to the participant a short preliminary report on findings that emerged from primary data analysis. During interviews, we posed three questions to participants that were intended to elicit their perceptions on subjects of interest to them in the report, as well as their assessment of the report's accuracy and completeness (see Appendix 4D). Though all of the participants shared observations about the topics of greatest interest to them in the report and the implications of our report for their own work, none of the five participants expressed concerns about the accuracy, completeness, or overall quality of the preliminary report.

Phase 5: Secondary Data Analysis

From November 2023 to December 2023, we conducted a second phase of content analysis. This secondary data analysis was conducted on the initial topics, categories, and quantitative data recorded in our codebook in addition to the new interview data that was collected during the quality assurance phase. By conducting combined analysis of aggregate data from the preliminary interviews along with data from the follow-up interviews, we identified implications of our study for researchers, practitioners, and policymakers. The findings and implications of our study are detailed in the following sections.

4.5. Findings

Table 4.1: Summary of key findings. Findings include breakdowns of component types and topics most frequently identified across the 20 interviews.

Section	**Analytical Dimension**	**Summary of Findings**
4.5.1	Work Contexts	*Key topics & interview counts:* Regulatory development (11), organizational AI governance & enablement (8), governance gaps & needs assessments (4), data governance mechanisms (4), regulatory compliance (4), national AI governance frameworks (4)
4.5.2	Actors	*120 unique actors*: 44 organizational actors, 23 socio-technical actors, 53 actor instances
4.5.3	Benefits, Risks, & Harms	*Key topics & interview counts:* Public mistrust (7), epistemic risks of misinterpreting “AI” (7), risk mitigation (6), improved efficiency (6), improved service quality (6), harms to worker wellbeing & workplace quality (5), unfair decision outcomes (5), financial & economic gains (5)
4.5.4	Resources	*142 unique resources*: 75 knowledge & cognitive resources, 26 policy & legal resources, 18 financial resources, 17 data & computational resources, 6 cultural resources

4.5.5	Networks	*40 unique networks*: 18 resource integration networks, 22 governance networks
4.5.6	Interactions	N/A (See explanation in Section 4.5.6)
4.5.7	Evaluations	*Key topics & interview counts:* Benefits to public realized & risks mitigated (6), AI systems are aligned with organizational values (6), reduction of stakeholder frictions (5), international harmonization on AI regulation & standards (4), trust of public & other stakeholders secured (4), shared meanings established (4)
4.5.8	Logics	*Key topics & interview counts:* Maximize benefits & mitigate risks of AI development & use (8), maximize profit & shareholder value from AI adoption (6), achieve balance between public & private interests in AI outcomes (6), align Canadian AI governance with international partners (6), align AI systems with organizational values (5), facilitate cross-organizational & cross-sectoral collaboration (4), strengthen public/consumer trust in AI applications (4), ensure AI systems are compliant with relevant laws & standards (4), ensure benefits realized & harms prevented for vulnerable groups (4).
4.5.9	Functional Bounds	*Key topics & interview counts:* Knowledge & expertise resource limitations within organizations (10), gaps in stakeholder literacies & awareness (8), knowledge & expertise resource limitations within Canada's AI ecosystem (6), gaps in knowledge of stakeholder needs & requirements (6), financial limitations (5)
4.5.10	Rules	*Key topics & interview counts:* AI-related laws & regulations (9), limitations on scope & jurisdiction of AI policies (7), impact of organizational norms & culture on AI policies (7), requirements set by voluntary standards & guidance (4), privatization of audit, compliance, & regulatory services (4), norms of & imbalances in political and economic power (4)
4.5.11	Ecosystem	*Key topics & interview counts:* International alignment on AI governance frameworks (13), feedback loops & learning across AI governance initiatives (11), AI Brussels effect (7), transferable & open-source models for organizational AI governance (6), Canadian access to international AI markets (5), clustering & scope of co-governance activities (5)

4.5.12	Opportunities for Improvement	*Key topics & interview counts:* Stronger shared understandings of AI systems & best practices for AI governance (8), creating more opportunities for public participation in AI governance (7), cultivating a more strongly integrated ecosystem (6), cultivating more diversity in AI governance activities (6), implementing stronger participatory design & governance practices (6)

4.5.1. Work Contexts

Across the 20 interviews, participants perceived *regulatory development* (11 interviews), *organizational AI governance & enablement* (8 interviews), *standards development* (6 interviews), *governance gaps & needs assessments* (4 interviews), *data governance mechanisms* (4 interviews), *regulatory compliance* (4 interviews), and *national AI governance frameworks* (4 interviews) as the topics most relevant to their own work contexts. For example, activities related to the development and implementation of the proposed Artificial Intelligence and Data Act, the development of context-specific AI policies and governance practices within organizations and sectors, and the development of national and international industry standards for AI by the International Standards Organization (ISO) and Standards Council of Canada (SCC) stood out as topics of particular interest across many of these interviews.

Many of the topics that participants described as most relevant to their own work contexts covered more specialized issues or interests, and were therefore perceived by the participants less frequently: 22 of 34 topics were described by only one or two participants. A complete breakdown of all 34 topics that one or more participants perceived as being most relevant to their own work contexts can be found in Appendix 4A. Topics that were perceived by participants less frequently are not necessarily less important to Canada's AIGS–participants may perceive similar issues through different epistemic frames, or may perceive issues that other participants do not perceive due to individual differences in their awareness, attention, work interests, and areas of specialization. Lower frequency topics related to participant work contexts cover a wide range of issues, such as the supply chains and ecosystems needed to develop AI systems, AI workforces, and AIG initiatives; different types of governance initiatives, approaches, and mechanisms; and different types of ethical and practical considerations that motivate the participants in their work.

4.5.2. Actors

Across the 20 interviews, participants perceived a total of 120 unique actors as being involved in or impacted by Canada's AIGS. The actors were described at varying degrees of scale, including entire nations and sectors down to specific government departments or specific companies that were perceived by interviewees as having a significant role in Canada's AIGS. We categorized the actors perceived by the participants within a typology that contains two actor types, 67 actor sub-types, and 53 instances of specific actors (see Figure 4.3). What we call *organizational actors* are a type of actor that participants described with reference to the organizational structures or functions of the actors, or with reference to a group of organizations with similar business offerings and similar governance needs (e.g., industries or sectors). What we call *socio-technical actors* are a type of actor that were described by participants with direct reference to the actor's role in the creation and/or operation of AI systems (e.g., actors involved in or impacted by the design, development, deployment, use, or management of an AI system). These categories are not mutually exclusive–specific instances of actors can be categorized under multiple types or sub-types depending on the organizational, socio-technical, and/or sectoral activities they are involved in or impacted by.

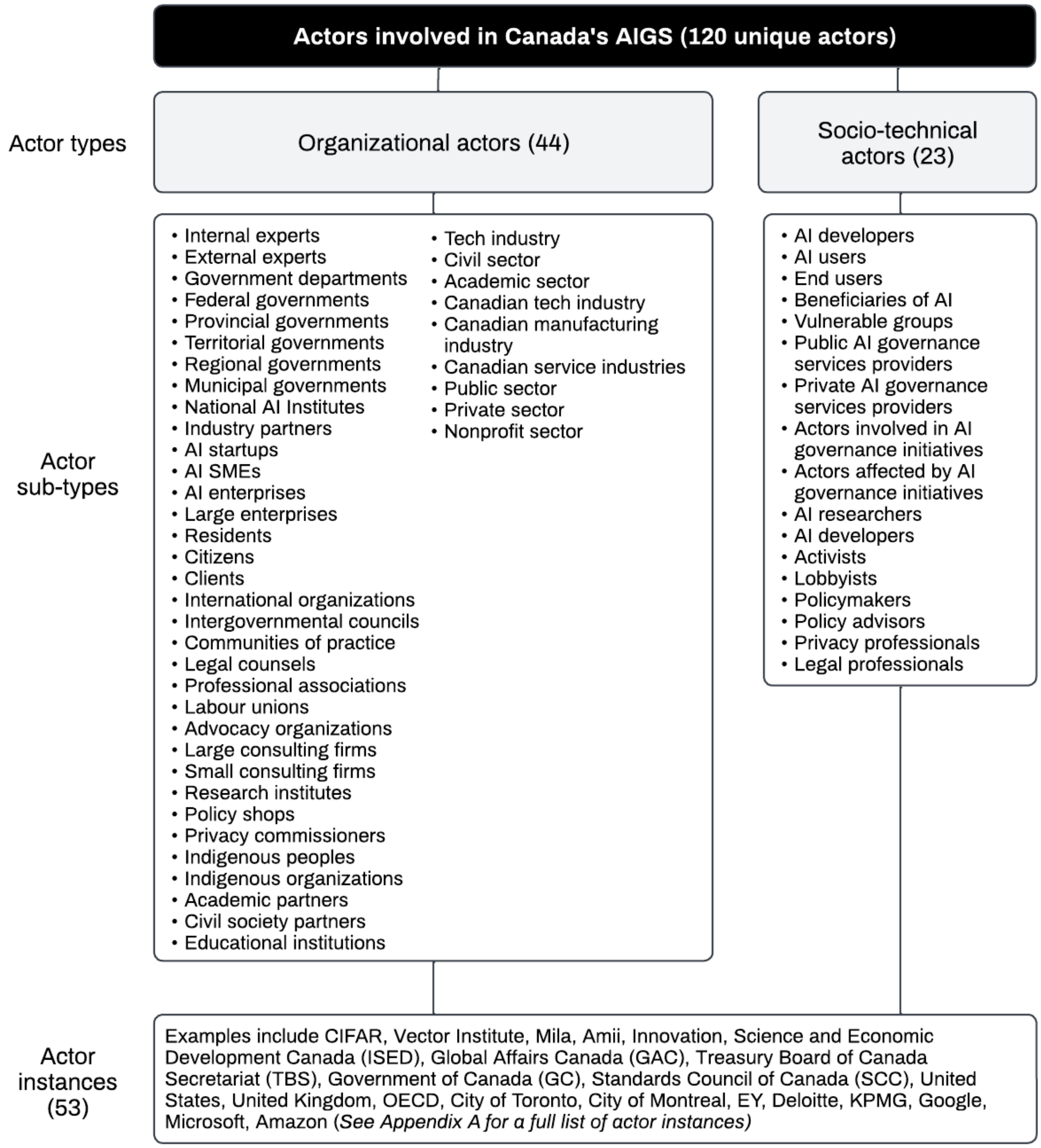

Figure 4.3: Diagram depicting a typology of actor types, sub-types, and instances that were perceived by participants.

4.5.3. Benefits, Risks, & Harms

Participants perceived many potential benefits, risks, and harms of AI systems and/or AI governance activities within Canada's AIGS. The most frequently perceived potential benefits were *risk mitigation* (6 interviews), *improved efficiency* (6 interviews), *improved service quality*

(6 interviews), and *financial & economic gains* (5 interviews). The most frequently perceived potential risks or harms were *public mistrust* (7 interviews), *epistemic risks of misinterpreting "AI"* (6 interviews), *harms to worker wellbeing & workplace quality* (5 interviews), *unfair decision outcomes* (5 interviews), *tensions between responsible AI & profitable AI* (4 interviews), *rights-based harms* (4 interviews), and *reinforcement of structural injustices* (4 interviews).

Many benefits, risks, and harms were perceived by participants less frequently. Lower frequency benefits, risks, and harms cover a wide range of issues. Many of the less frequently perceived topics cover the potential benefits of responsibly governed AI to various aspects of Canadian leadership and competitiveness in AI, AI innovation and technological development, financial and economic gains, service quality and cost savings in specific types of organizations, as well as potential benefits of effective AIG practices to scalability, robustness, accountability, transparency, and trust in AI systems. Other lower frequency topics cover potential risks and harms caused by AI systems, including risks and harms to marginalized groups, workers, businesses, and the environment. Some participants also indicated that ineffective AIG practices may pose complex systemic risks to political institutions and institutional capacities, regulatory systems, supply chains, legacy digital infrastructures, personal privacy and security, social and economic equity, and economic equilibrium. A complete breakdown of all 72 benefits, risks, and harms that were perceived by the participants can be found in Appendix 4A.

4.5.4. Resources

Across the 20 interviews, participants perceived a total of 142 unique resources as being involved in Canada's AIGS. Of those 142 resources, we categorized a significant majority as *knowledge & cognitive resources* (75 unique resources). This category includes resources such as knowledge of AI development and use, knowledge of AI ethics, technical expertise, business expertise, legal expertise, implementation expertise, working groups, advisory and decision-making bodies, AI literacies, and AI education and training for marginalized groups.

In addition to knowledge & cognitive resources, we categorized many resources as *policy & legal resources* (26 unique resources). This category includes resources for rule-giving such as laws, regulations, directives, guidance documents, technical and industry standards, auditing and assessment frameworks, compliance management programs, and human rights. This category also includes resources for rule-making such as legal teams, advocacy groups, policy models, and best practices. Another category we identified is *financial resources* (18 unique resources), which includes resources such as budgets, public and private funding and investment, commercialized AI models, business models, capital, and funds for marginalized groups, civil society, and private sector organizations to participate in Canada's AIGS. *Data & computational resources* (17 unique resources) includes a variety of technical resources required for the functioning of responsibly governed AI systems, such as productionized machine learning models, cleaned and labeled data, high quality and domain-specific training and testing data, reliable data pipelines and technical services supply chains, along with reliable computing resources, hardware, and digital infrastructures. *Cultural resources* (6 unique resources) are intangible resources that enable capacity-building and legitimacy-building within organizational cultures, such as cultural awareness of AI ethics and governance issues, approvals and buy-in for creating and implementing AIG initiatives, and positive public perceptions of AI systems and their governance. A complete breakdown of all 142 resources that were perceived by the participants can be found in Appendix 4A.

4.5.5. Networks

Across the 20 interviews, participants perceived a total of 40 unique networks as being involved in Canada's AIGS. Of those 40 unique networks, we categorized 18 as types of *resource integration networks* through which multiple actors attempt to exchange resources to achieve their individual goals. We categorized the remaining 22 unique networks as types of *governance networks* through which multiple actors attempt to coordinate their actions, align their goals, and collectively act upon their shared values. Resource integration networks cluster around particular types of resources (e.g., knowledge sharing networks among departments, sectors, governments, or nations; networks for exchanging expertise, practical guidance, or for providing AI education and training; networks and platforms for sharing code, data, or social communications). Governance networks cluster around the actors and resources involved in maintaining particular organizational structures and/or in enabling particular organizational functions (e.g., the network of actors and resources involved in maintaining the Pan-Canadian AI Strategy and enabling organizations through it; the networks involved in developing, legislating, reviewing, and assessing compliance with the Artificial Intelligence and Data Act; the networks involved in formalizing a set of mechanisms and practices for governing AI systems within an organization). A complete breakdown of all 40 networks that were perceived by the participants can be found in Appendix 4A.

4.5.6. Interactions

During primary data analysis, we excluded interview data pertaining to interactions between specific actors from further study to preserve participant anonymity. In many cases, the context-specificity of the actor-actor interactions described by participants (e.g., interactions between highly specific organizations even within aggregated data) presented a risk of making it possible to identify individual participants from the aggregated data. We were able to infer findings pertaining to interactions from the actors, networks, activities, and ecosystem-level phenomena described by the participants. Findings pertaining to interactions are reported on in more detail in the following sections.

4.5.7. Evaluations

The actors in Canada's AIGS apply many criteria to evaluate the success of AIG practices. Across the 20 interviews, the most frequently perceived evaluation criteria for successful AIG practices were *benefits to public realized & risks mitigated* (6 interviews), *AI systems are aligned with organizational values* (6 interviews), *reduction of stakeholder frictions* (5 interviews), *international harmonization on AI regulations & standards* (4 interviews), *trust of public & other stakeholders secured* (4 interviews), and *shared meanings established* for foundational, scope-setting terms such as "AI systems" and "AI governance" (4 interviews).

Many evaluation criteria were perceived by participants less frequently. Lower frequency evaluation criteria cover a wide range of qualities that participants perceived as being characteristic of successful AIG practices, including strategic and operational effectiveness, efficiency, competency, interoperability, scalability, transferability, competitiveness, profitability, protectiveness, compliance, accountability, transparency, and legitimacy. These evaluation criteria were often described by participants in relation to specific organizational or institutional outcomes, such as high applicability and transferability of policy instruments across contexts, balancing of generalizability and context-sensitivity in governance frameworks, and enablement of effective knowledge-sharing, organizational management, and standardization practices. A complete breakdown of all 48 evaluation criteria that were perceived by the participants can be found in Appendix 4A.

4.5.8. Logics

The actors in Canada's AIGS are motivated by a wide range of intentions and desires. Across the 20 interviews, the most commonly perceived logics for practicing AIG were to *maximize benefits & mitigate risks of AI development & use* (8 interviews), *maximize profit & shareholder value from AI adoption* (6 interviews), *achieve balance between public & private* *interests in AI outcomes* (6 interviews), *align Canadian AI governance with international partners* (6 interviews), *align AI systems with organizational values* (5 interviews), *facilitate cross-organizational & cross-sectoral collaboration* (4 interviews), *strengthen public/consumer trust in AI applications* (4 interviews), *ensure AI systems are compliant with relevant laws & standards* (4 interviews), and *ensure benefits realized & harms prevented for vulnerable groups* (4 interviews).

When participants were prompted to describe logics for practicing AIG that they perceived within Canada's AIGS, many participants described similar issues to the issues that they previously covered in describing evaluation criteria for successful AIG practices. However, in describing logics for practicing AIG, participants often framed their perceptions differently than they had framed their perceptions of evaluation criteria. Participants often framed logics for practicing AIG in relation to perceived social, organizational, political, economic, material, or psychological needs of one or more actors. Evaluation criteria, however, were often framed by participants as qualitative measures or performance indicators that could be used to assess if those needs had been met. The difference between these two framings is subtle but important: participant perceptions of evaluation criteria relate most directly to the quality and measurability of AIG activities, whereas perceptions of logics relate more directly to the rationales and aspirational goals that motivate those AIG activities. We represented this framing difference between logics and evaluation criteria in our coding of the topics. For example, the evaluation topic *AI systems are aligned with organizational values* (ID#6.2 in Appendix 4A) seems nearly identical to the logics topic *align AI systems with organizational values* (ID#7.5). However, the former topic implies that some actors perceive value alignment as a measurable quality of organizational AIG activities; the latter topic implies that some actors perceive the alignment of

AI systems with organizational values as a rational and desirable goal in carrying out their organizational AIG activities.

The logics motivating the various actors of Canada's AIGS are grounded in tacit assumptions about the rationality and desirability of AI adoption and AIG activities. Implicit to many of these logics is an assumption that AI adoption is intrinsically desirable, and that extrinsic AIG practices can enable rational actors to responsibly fulfill their AI adoption desires through a utilitarian risk-benefit calculus. The logics motivating Canada's AIGS cover many rationales and desired goals. Participants indicated that many actors desire AIG practices that strike reasonable balances between perceived dichotomies such as benefits/risks, public/private interests, and innovation/regulation. Participants also perceived desires to align AI systems with human values across many scales of activity, including alignment with stakeholder values, shareholder values, organizational values, national values, and international markets. Securing the trust of consumers, vulnerable groups, markets, and the Canadian public in AI systems and their governance is another significant desire motivating Canada's AIGS, as well as strengthening the capacities, capabilities, and competitiveness of Canadian institutions and businesses. The frequency with which topics related to financial and economic desires were raised by participants–most strikingly, *maximize profit & shareholder value from AI adoption* (ID#7.4, 6 interviews)–indicates that Canada's AIGS is motivated by strong economic logics. It is unclear from our interviews how such powerful incentive structures might be reconciled with conflicting logics perceived by participants–such as trust-building, harm prevention, accountability, and justice–that seek to disrupt undesirable political and economic power structures. A complete breakdown of all 39 logics that were perceived by the participants can be found in Appendix 4A.

4.5.9. Functional Bounds

The actors in Canada's AIGS face many cognitive, informational, and other functional limitations in pursuing their AI governance goals. Across the 20 interviews, the most frequently perceived functional boundaries on AIG practices were *knowledge/expertise resource limitations within organizations* (10 interviews), *gaps in stakeholder literacies & awareness* (8 interviews), *knowledge/expertise resource limitations within Canada's AI ecosystem* (6 interviews), *gaps in knowledge of stakeholder needs & requirements* (6 interviews), *financial limitations within organizations* (5 interviews), *administrative incapacities within organizations* (4 interviews), and *interpretive barriers & ontological uncertainties* in making shared meanings and collectively acting upon foundational terms such as "AI systems" and "AI governance" (4 interviews).

The functional bounds described by participants reflect constraints on the availability and use of many of the knowledge resources, cognitive resources, and cultural resources that we discussed in section 4.5.4. The frequency with which participants perceived various limitations in the availability of AIG knowledge and expertise–both within specific organizations, and across the broader ecosystem–indicates that Canada is facing significant constraints in its AIG knowledge supply chains. AIG knowledge shortages were perceived by participants as limiting the development and quality of Canada's AIGS at both micro- and macro-organizational scales. Awareness of and literacy with a variety of AI-related topics is also an important cognitive resource, and here too, participants frequently perceived gaps in the awareness and literacies of many different actors as constraining the development and quality of Canada's AIGS. Further cognitive constraints are imposed by ambiguous ontologies of "AI" and epistemic disconnects between policymakers and the contexts of AI development and use they desire to govern. Participants frequently noted that "AI" is a highly ambiguous object of governance. This ontological ambiguity cascades into epistemic and practical limitations on interventions in "AI systems," as policymakers and other AIG practitioners must balance generalized cross-contextual awareness and specialized context-sensitive awareness in scoping the "AI systems" and impacts they aim to intervene in. Participants also perceived limitations on financial resources as constraining the AIG activities of many under-resourced actors, particularly small businesses, small government policy teams, civil society organizations, and advocacy groups.

Data resource constraints and computational resource constraints were not frequently perceived as functional limitations on AIG activities. Only two functional bounds directly relate to data resources or computational resources: *technical debts* (ID#8.11) and *cultural limitations of data sourcing & documentation practices* (#8.15). Rather than being purely technical constraints, both of these topics strongly relate to constraints in organizational culture. Technical debts are constraints imposed within organizational cultures that promote non-maintenance of legacy technical infrastructures that support new AI applications through API interfaces. Data sourcing and documentation are constraints imposed within cultures that promote responsible data governance and compliance with data protection regulations.

Many participants perceived that capacity gaps can emerge from multiple interlocking resource constraints. The capacity gaps perceived by participants include incapacities of organizational actors, government institutions, civil society, and the broader ecosystem to effectively develop, use, manage, make sense of, and share knowledge about AI systems, impacts, policy instruments, and other governance mechanisms. A complete breakdown of all 25 functional bounds that were perceived by the participants can be found in Appendix A.

4.5.10. Rules

In addition to functional limitations, the actors in Canada's AIGS also face many normative and rule-based limitations in pursuing their AI governance goals. Across the 20 interviews, the most frequently identified normative and rule-based limitations on AI governance practices were *requirements set by AI-related laws & regulations* (9 interviews), *limitations on scope & jurisdiction of AI policies* (7 interviews), *impact of organizational norms & culture on AI policies* (7 interviews), *requirements set by voluntary standards & guidance* (4 interviews), *privatization of audit, compliance, & regulatory services* (4 interviews), and *norms of & imbalances in political and economic power* (4 interviews), *avoidance of conflicts of interest in policy co-design* (3 interviews), and an *emerging norm of accountability & enforceability gaps* (3 interviews).

While participants typically perceived functional bounds as emerging from a combination of knowledge, cognitive, cultural, and financial resource constraints, participants typically perceived rules as emerging from a combination of constraints imposed by policy resources, legal resources, and cultural resources. Participants perceived a spectrum of AI policy instruments with varying tradeoffs between legal force and administrative agility, ranging from legislation and regulation (high force/low agility) to voluntary guidance documents (low force/high agility). Participants also perceived that legal, regulatory, and policy frameworks for AI may impose differing limitations across actors depending upon the scope of AI development and use activities covered by the framework and the jurisdiction through which the framework is to be enforced.

The normative limitations perceived by participants are often more complex and restrictive than limitations based in policy or law. Participants frequently perceived that political, economic, and organizational norms impose tacit rules and structural patterns upon AIG activities and outcomes. Various norms of political and economic power–and particularly, the consolidation of political and economic power into a small set of government institutions and industry actors–were frequently perceived as imposing structural limitations and power imbalances upon the activities of Canada’s AIGS. Norms of power were perceived as influencing the public accountability processes of government and industry actors, the development and enforcement of legal and regulatory frameworks for AI, the ability of government institutions to avoid conflicts of mandate in co-designing AI policies with one another, and the ability of government to avoid conflicts of interest in co-designing AI policies with industry actors (who often possess greater knowledge of and expertise with AI than government actors). Participants also perceived that incentive structures are placing limitations on regulatory development and enforcement: many government institutions face limitations in financial resources and administrative capacity, so they are incentivized to outsource regulatory activities such as AI system audits, compliance assessment and certification, and professionalization of AI auditors and other AI policy experts to private sector actors with greater capacities; meanwhile, the private sector is incentivized to support the government in privatizing those administrative capacities, as privatization will enable businesses to create lucrative new markets for AI auditing, compliance management, and other professional services. Participants connected these incentive structures to long-standing political and cultural norms of outsourcing in Canada’s public sector, and to an emerging norm of privatized AI regulation. Some participants advocated for this emerging norm, noting potentials for regulatory efficiencies and cost reductions; others expressed concern that excessive regulatory privatization may ultimately contribute to the further erosion of administrative and regulatory capacities in Canada’s public sector.

Participants also perceived organizational norms such as industry self-regulation of AI, agile organizational cultures, and culturally encoded narratives of AI futures and technological futurism as tacit rules that limit the structural possibilities of Canada’s AIGS. A complete breakdown of all 21 rules that were perceived by the participants can be found in Appendix 4A.

4.5.11. Ecosystem

Canada's national AIGS exists within a broader ecosystem of Canadian and international governance systems that impact Canada's AIGS. Across the 20 interviews, the most frequently perceived ecosystem-level topics were *international alignment on AI governance frameworks* (13 interviews), *feedback loops & learning across AI governance initiatives* (11 interviews), the *AI Brussels effect* through which European Union AI policy is able to strongly influence Canada's development of AI policy (7 interviews), *transferable & open-source models for organizational AI governance* (6 interviews), *Canadian access to international AI markets* (5 interviews), and *inter-departmental clustering & scope of co-governance activities* (5 interviews).

When prompted to describe ecosystem-level phenomena either within or beyond Canada's national context, participants described four main types of external AIG systems that Canada's national AIG system interacts with: international AIG systems, subnational AIG systems, sectoral AIG systems, and organizational AIG systems. Participants perceived many international AIG activities of significance to Canada's national AIGS, including international alignment of regulatory models and harmonization of market frameworks with key trading partners (e.g., the European Union, the United States, and other G7 nations), as well as knowledge sharing, policy transfer, and development of technical standards and best practices through international organizations and communities of practice (e.g., the OECD, UNESCO, NATO, and ISO). In addition, some participants observed that geopolitical tensions are contributing to volatility and uncertainty in Canada's broader strategic environment, and to systemic risks of misperceiving AI opportunities and threats, intensifying AI arms race dynamics, and proliferating military applications and malicious uses of AI.

At the subnational scale, participants perceived many collaborative AIG activities occurring across clusters of federal, provincial, territorial, and municipal government institutions. Though Canada's national AIGS is largely directed by federal strategy and policy initiatives, many AIG initiatives also exist at the provincial and municipal levels of government. For example, Ontario's *Trustworthy AI Framework* (2023) provides guidance for AI development and use across the province's public sector, and a version of the federal *Directive on Automated Decision-making* has been adapted for use in the municipal government of the City of London, Ontario (as described in Towards Data Science, 2021). Participants frequently described intergovernmental knowledge sharing, formal collaboration, and informal collaboration on AIG initiatives as significant components of Canada's national AIGS. In addition to government institutions, participants also perceived many collaborative AIG activities occurring within and across industries, sectors, labour unions, professional associations, communities of practice, and civil society. Though they exist within the subnational scale, sectoral and organizational AIG systems have wide-ranging impacts on both the Canadian private sector and public sector. Sectoral and organizational AIG activities include the development, implementation, and integration of context-specific AI strategies, standards, and policies between and within organizations. Participants also noted the importance of amending and coordinating existing sectoral and organizational governance frameworks (e.g., for data governance, privacy protection, compliance management, and impact assessment) with new AIG frameworks so as to ensure continuity and coherence between the frameworks. Figure 4.4 depicts an aggregated view of the key ecosystem-level entities and relations that were perceived by participants. A complete breakdown of all 25 ecosystem-level phenomena that were perceived by participants can be found in Appendix 4A.

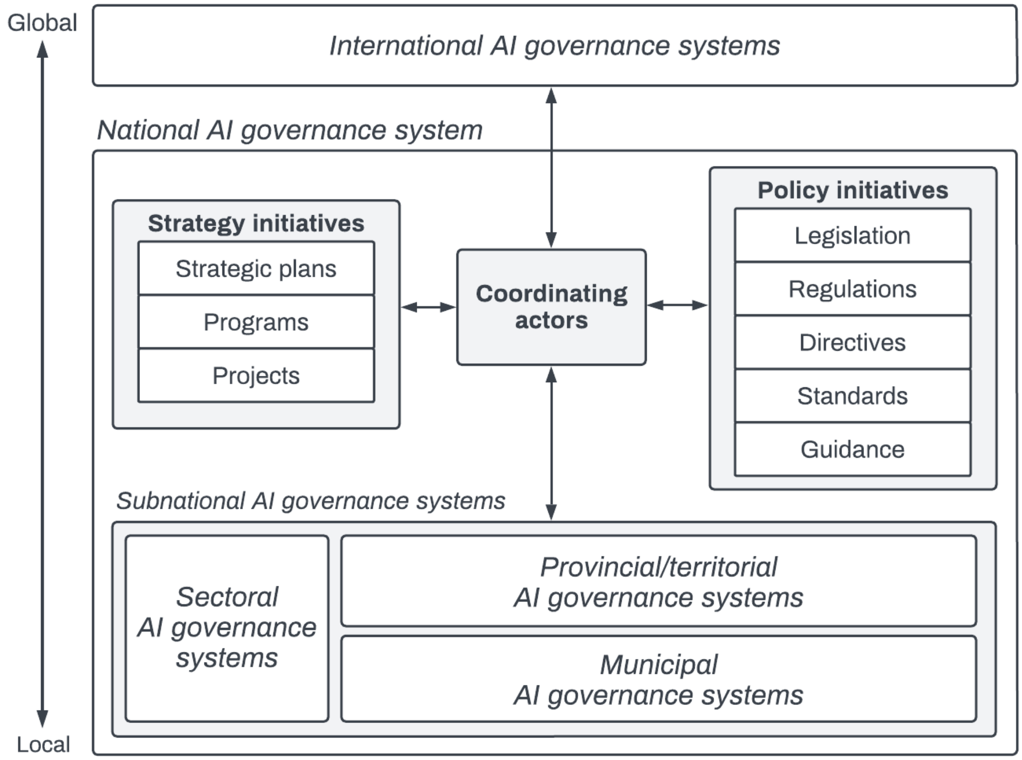

Figure 4.4: Diagram of key ecosystem-level entities and relations that participants perceived as significant to Canada's national AIGS. Structures of strategy initiatives, policy initiatives, and coordinating actors also exist within international and subnational AIG systems but are omitted from this diagram for the purpose of visual simplicity.

4.5.12. Opportunities for Improvement

Across the 20 interviews, participants perceived many opportunities for improving the overall state of Canada's national AIGS. The most frequently described opportunities for improvement were *stronger shared understandings of AI systems & best practices for AI governance* (8 interviews), creating *more opportunities for public participation in AI governance* (7 interviews), cultivating a *more strongly integrated ecosystem across departments, governments, & sectors* (6 interviews), cultivating *more diversity in AI governance activities* (6 interviews), and implementing *stronger participatory design & participatory governance practices* (6 interviews).

The opportunities for improvement described by participants can be characterized as four types of high-level gaps in Canada's national AIGS: (1) gaps in institutional arrangements, (2) gaps in resource availability, (3) gaps in actor alignment, and (4) gaps in diversity, equity, and inclusion. Participants frequently described structural and functional gaps in institutional arrangements as opportunities for improvement. These institutional gaps include: loose coordination on AIG initiatives across government departments, levels of government, and sectors; planning, implementation, experimentation, and scope challenges in strategic initiatives and policy initiatives; gaps in the public sector's administrative, regulatory, and training capacities; and systemic challenges with agility, adaptability, resilience, scalability, public accountability, and public trust across Canada's AIGS. Gaps in the availability of many types of resources were also frequently described as opportunities for improvement. Notable resource gaps include limited knowledge and awareness of AI impacts, of stakeholder needs, and of best practices for AI governance throughout Canada's AIGS, particularly in public sector organizations and organizations with less mature AI governance cultures. Participants also perceived gaps in the supply of AI policy expertise, AI commercialization and business scale-up expertise, and other supports for AI policy, commercialization, and scale-up. Another opportunity for improvement noted by participants was to reduce financial barriers and other resource constraints for small businesses, civil society organizations, and marginalized groups wishing to participate in policy co-design and standardization activities.

In addition to institutional and resource gaps, participants frequently perceived a need to create new governance mechanisms to better address challenges of actor coordination, collaboration, and alignment across multiple scales of AIG activity. These include challenges of intra-organizational alignment (alignment of values and goals of individual workers, teams, and departments within an organizational AIGS), inter-organizational alignment (alignment of values and goals of multiple organizations in an AI value chain), national alignment (alignment of the national AIGS with values and goals of individuals and organizations across a nation), and international alignment (alignment of multiple national AIG systems in a transnational AI value chain).

Finally, participants frequently described gaps in diversity, equity, and inclusion as opportunities for improvement. Gaps in racial, gender, disciplinary, or sectoral diversity of the individuals and organizational actors represented in AIG initiatives were frequently described. Participants also described shortages of collaboration opportunities and resources for publics, civil society organizations, marginalized groups, and Indigenous peoples wishing to participate in government-led AIG initiatives. Some participants perceived inequitable AI impacts and AIG outcomes as potential consequences of governance practices that do not take a diverse and broadly inclusive approach to public participation. A complete breakdown of all 44 opportunities for improvement that were perceived by the participants can be found in Appendix 4A.

4.6. Limitations

Our data and findings represent the most significant components of Canada's national AIGS as perceived by our participants during the data collection phase of our study (February 2023 – July 2023). AIG is a rapidly evolving field of research and practice, and AIG initiatives are continuously emerging and developing throughout Canada and around the world. Our study is not intended to provide an exhaustive or immutable account of Canada's AIGS–our data and findings provide a time-sensitive, accurate, and reasonably complete representation of Canada's AIGS, as confirmed by participants during the quality assurance phase of our study (October 2023 – December 2023). Our data and findings may contain minor gaps in timeliness and completeness due to the complexity and fluidity of the field, and new minor gaps will likely continue to emerge as Canada's AIGS continues to evolve.

Our discussion of findings is intended to provide an integrative analysis of many topical trends and topical gaps that we observed across each analytical dimension of our dataset. Our discussion of findings is not intended to provide a comprehensive account of all of the findings that could possibly be inferred from topics, topical patterns, and topical gaps in our dataset. As discussed in section 4.4.1, the topics in our dataset are aggregated from across 20 interviews that are subject to the perceptual limitations of the participants, such as individual differences in attention, knowledge, and interest. Additionally, our analysis is subject to our own perceptual constraints on attention, knowledge, and interest as researchers. Our data and our discussion of findings are therefore also subject to the perceptual limitations of our participants and ourselves. We mitigated those perceptual limitations to the greatest extent possible by including participants from multiple sectoral and disciplinary backgrounds, by conducting our 20 initial interviews over the span of six months to minimize the potential influence of external events on group-level attentional biases, and by conducting follow-up interviews several months later to ensure our data and findings remained accurate.

The size of our sample imposes minor limitations on our data and our findings. We determined that our data collection process had reached saturation after 20 interviews, as clear distributions of topics perceived more frequently versus topics perceived less frequently had emerged across each of the analytical dimensions. We do not expect that these distributions or the relative weighting of the higher-frequency topics would substantially change if further interviews are conducted with additional participants. Though additional lower-frequency topics perceived by one or few participants could emerge through further interviews, these additional topics would not have a significant impact on the overall quality or accuracy of our findings.

Despite these limitations, we confirmed through our quality assurance interviews that there are no significant gaps in the accuracy or completeness of our study.

4.7. Implications for Future Research, Practice, & Policy

4.7.1. Implications for Research

Our findings reveal an extensive agenda for future research on AIG in the Canadian context, AIG at national scales, and AIG practices more generally. We divide this agenda into three future research directions: (1) conduct additional analysis of topics in our dataset, (2) further investigate institutional and ecosystem phenomena, (3) apply our framework, data, and findings to other AIG research contexts.

Research Direction 1: Conduct additional analysis of topics in our dataset. Our study contributes a rich set of data and empirically grounded findings to future studies of AIG. The dataset we created over the course of this study (Appendix 4A) provides researchers with a total of 610 topics that were aggregated from across our 20 interviews. Collectively, these topics represent the most significant components of Canada's national AIGS as perceived by the government leaders and subject matter experts who participated in our study. Our discussion of findings in the previous section was structured according to the 12 analytical dimensions of our service system framework, and we used content analysis methods to identify topics, categories, trends, and gaps in the data within each of the 12 analytical dimensions. Theories other than our theoretical framework of service system analysis could be used to structure the data into different analytical dimensions and to contextualize and interpret the data differently, thereby generating additional categories, trends, gaps, and findings. Additional findings could also be generated by conducting thematic analysis or entity extraction of multiple analytical dimensions at once, or of the entire dataset without separating the data into discrete analytical dimensions. Alternative analysis methods could also be applied to the dataset, such as sentiment analysis (Mäntylä, Graziotin, & Kuutila, 2018), topic-based frame analysis (Ylä-Anttila, Eranti, & Kukkonen, 2022), and topic-based discourse analysis (Brinkmann, 2019) to generate more detailed findings on the positive/negative sentiments, epistemic and political frames, and communicative patterns that are influencing Canada's national AIGS. Future studies could also analyze topical and thematic gaps in the dataset, focusing analysis and reporting on topics that were not perceived by participants or were perceived by participants less frequently, despite being well-evidenced in other data sources and research literature.

Research Direction 2: Further investigate institutional & ecosystem phenomena. Our most complex findings relate to *institutional arrangements* (the logics, functional bounds, and rules structuring Canada's national AIGS) and *ecosystem-level relationships* between Canada's national AIGS and other AIG systems. Many researchers have analyzed the institutional arrangements shaping the development of the Directive on Automated Decision-making (Karanicolas, 2019; McKelvey & MacDonald, 2019; Scassa, 2021) and the Artificial Intelligence and Data Act (Scassa, 2023; Tessono et al., 2022). Aloisi and De Stefano (2023) compare Canada's legal framework for AI to US and EU frameworks, analyzing institutional and cross-jurisdictional effects such as the AI Brussels effect, legal transfer, and regulatory arbitrage. Faveri and Auld (2023) analyze the effects that Canada's regulatory framework for AI may have on the structuring and functioning of regulatory institutions and audit markets. Brandusescu (2021) analyzes the tensions between public and private interests and the economic power structures that influence many Canadian AIG activities. Lepage-Richer and McKelvey (2022) analyze how political and cultural resources are mobilized and constrained as part of AI adoption and AIG initiatives within Canada's federal government. Frost (2020) analyzes the institutional arrangements through which Canada's national AI innovation system functions in service of Canada's broader economic and geostrategic objectives.

Though these studies shed greater light on many of our own findings regarding the political and economic logics, resource constraints, power structures, and international relations that influence Canada's national AIGS, many of our other findings regarding institutional and ecosystem phenomena have not been well-studied in the research literature. To address this gap in the literature, more detailed investigation is needed of many of the issues we discuss in sections 4.5.8-4.5.11. The issues requiring further study include conflicting logics and value misalignments in Canada's AIG activities; effects of organizational cultures and power imbalances on AI policy development processes and Canada's AIG outcomes; constraints in Canada's supply chains of AIG knowledge, expertise, and contextual awareness; incentive structures and potential effects of regulatory privatization in Canada's AIGS; barriers to civil society participation in Canadian AIG activities; and barriers to effective collaboration, coordination, and resource integration with international and subnational AIG systems.

Research Direction 3: Apply our framework, data, & findings to other AIG research contexts. Our study contributes not only to the advancement of research on Canadian AIG, but also to the advancement of AIG research more generally. Our framework and methodology are transferable to studies of other national AIG systems, and can be further fine-tuned as new findings emerge across research contexts. Though our data and findings are grounded in the Canadian context, they may also be applicable to other AIG contexts in which leaders and experts have similar work interests to the work interests described by our participants in section 4.5.1, such as regulatory development and compliance, organizational AI governance and enablement, governance gaps and needs assessments, data governance mechanisms, and national AIG frameworks. However, additional interviews should also be conducted with leaders and experts in other AIG contexts to identify their own work interests and to determine the extent to which our data and findings are applicable to their work contexts. Comparative analysis of

Canada's national AIGS and other contexts could then also be conducted to further enrich the data and findings generated from across the studies.

4.7.2. Implications for Practice & Policy

Based on participant perceptions of opportunities for improving Canada's AIGS and the four high-level gaps discussed in section 4.5.12, we identify four interconnected strategic objectives that–if acted upon by policymakers, public servants, and other practitioners of AIG–would strengthen the overall state of Canada's national AIGS:

(1) Implement new collaboration and coordination mechanisms.

(2) Create guidance for designing & implementing participatory AIG initiatives.

(3) Expand access to key resources needed for effective AIG practices.

(4) Advance diversity, equity, and inclusion in AIG activities.

Although these four strategic objectives are ambitious in scope and scale, there are many specific interventions in policies, programs, and strategic initiatives that Canadian AIG leaders can begin enacting to fulfill the four objectives. Figure 4.5 illustrates these potential interventions in relation to our four strategic objectives, our four high-level gaps, and the 44 opportunities for improving Canada's AIGS that were perceived by participants. Some of these objectives and interventions will be more difficult or costly to enact than others. Any organization acting upon these objectives will need to prioritize some interventions over others based on their available resources and their organization's short, medium, and long-term strategic plans. An in-depth discussion of these strategic objectives and potential interventions can be found in Appendix 4F.

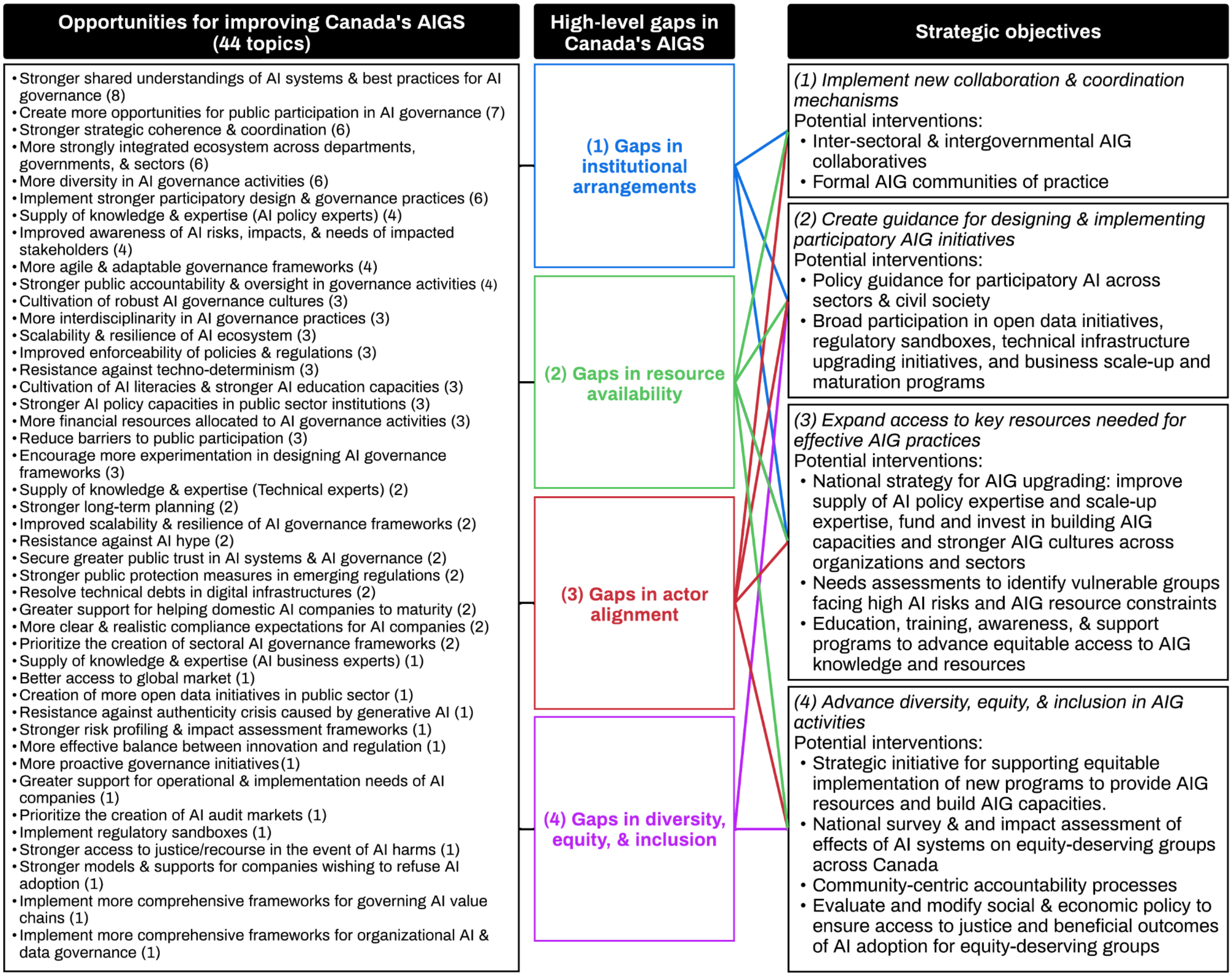

Figure 4.5: Relationships between participant perceptions of 44 opportunities for improving Canada's AIGS, four high-level gaps, four strategic objectives for addressing those gaps, and potential interventions for fulfilling each of the strategic objectives.

4.8. Conclusion

The findings and implications of our study address our three research questions (RQ1-RQ3 in section 4.1):

RQ1 – Ontological Question: What components exist within an AIGS?

RQ2 – Analytical Question: How can those components be empirically studied within an AIGS?

RQ3 – Ecological Question: How can those components be empirically studied in a macro-scale AIG system that contains a population of AIG systems within it, such as a national AIGS?

The governance of AI in Canada functions through a network of interdependent service systems that operate at international, national, and subnational scales (RQ1). These systems are structured by the perceptions, values, and activities of many actors within and outside of Canada, and can be empirically studied by applying service system analysis methods to analyze AIG activities within and across scales (RQ2 and RQ3).

Our study contributes new empirical data and findings to researchers, practitioners, and policymakers who wish to better understand or intervene in the systems through which a nation's governance of AI is structured and operated. Our study also contributes novel theoretical and methodological frameworks to researchers, practitioners, and policymakers who wish to transfer our theories and methods to their own studies of national AIG systems or AIG systems at other scales of activity. The theories and methods used in our study can be transferred to other studies of AIG systems by applying the concepts established in section 4.2 and Appendix 4B, as well as by applying the research activities described in section 4.4 and Appendices 4C-E to collect data from and analyze the service system components of different AIG research contexts.

As national pursuits of responsible AI governance continue, it will be vital for researchers, practitioners, and policymakers to remain mindful that their perceptions of "AI", "AI systems", "AI impacts", and "AI governance" are structured by the social, political, economic, and organizational contexts they are situated within. Many perceptions exist of AI systems, their impacts, and the types of activities that constitute their governance. To responsibly govern AI, individuals, organizations, and nations must expand the periphery of AI governance to account for a greater range of AI impacts, and to include a greater diversity of actors in AI governance activities. Only then can the many challenges of AIG be addressed in a way that will faithfully represent the perceptions and interests of everyone with a stake in AI.

Chapter 5

Conclusion:

A Future for Transfeminist AI Governance

5.1. Summary of Findings & Recommendations

This dissertation was motivated by an observation that AI governance is often ineffective at preventing harms caused by AI systems. Although it was clear that AI systems frequently cause harmful impacts to society and the environment, the reasons why AI governance practices are often ineffective at preventing those harms were relatively unclear. When read through a transfeminist lens—a lens centering challenges of power, inclusion, equity, community, and transformative change—the research findings presented throughout this dissertation shed greater light on the reasons why AI governance is so often ineffective at preventing harm. Table 5.1 outlines the three overarching objectives of this dissertation and the key findings that have emerged from research addressing each objective.

Table 5.1. Key findings from the research presented in chapters 2, 3, and 4 that emerged in response to the three overarching objectives of this dissertation.

Research Objective	Key Findings from Research
1. AI Impacts: Determine what types of benefits and harms AI systems are capable of causing, the actors responsible for those benefits and harms, the actors impacted by those benefits and harms, and the activities through which those benefits and harms are caused.	• AI systems are capable of causing benefits to efficiency, productivity, and quality of products and services. • AI systems are capable of causing harms to marginalized populations and regions, workers, political and economic freedoms, social trust, democratic institutions, public services, justice systems, and ecosystems. • Beneficial and harmful impacts of AI systems are caused by developers, users, service providers, data center operators, manufacturers, governments, communities, and workers. • AI value chains enable co-creation of value in AI systems through integration of software resources, hardware resources, knowledge resources, financial resources, and governance resources.
2. AI Governance Initiatives: Determine what types of AI governance initiatives have been created to intervene in those impacts, and the degree to which those initiatives are effective or	• Policy instruments, programs, strategic plans, standards, and ethics statements have been created to intervene in the impacts of AI systems. • In Canada, AI governance initiatives have often been effective at prioritizing interventions in Canadian industry and innovation, AI technology production and use, AI research, and public administration applications of AI.

ineffective at intervening in those impacts.	• In Canada, AI governance initiatives have often been ineffective at reporting on post-implementation outcomes, securing public trust in AI, intervening in a diversity of AI impacts, and cultivating strong coordination between governments, sectors, and civil society.
3. AI Governance Systems: Determine how those initiatives function as part of larger AI governance systems that exist across multiple contexts and levels of scale.	• AI governance initiatives function as activities within AI governance systems, which can be conceptualized as a type of service system. • The activities that occur within AI governance systems are co-created by actors, resources, networks, logics, bounds, and rules that interact across multiple contexts and levels of scale. • AI governance systems and the initiatives that are designed and implemented within them are structured by the values, perceptions, and capacities of many actors.

In Chapter 1 (*Introduction*), I set out to address three interrelated research objectives regarding the impacts and governance of AI systems. To fulfill those three research objectives, I developed a meta-theoretical framework grounded in the literature on service science, feminist science and technology studies, and queer and trans studies. By applying service realist and transfeminist theories of ontology, epistemology, and ethics, I conceptualized AI governance as an intermediary system that exists between AI systems and their socio-material contexts. I introduced AI governance systems as social systems consisting of dominant and marginalized actors, activities, resources, networks, logics, rules, and norms that enable intervention in beneficial and harmful impacts caused by AI systems across many contexts and scales of activity. I demonstrated that the application of transfeminist principles and practices to AI governance activities can support the protection and flourishing of marginalized communities harmed by AI systems.

With research objectives and a meta-theoretical framework established, we determined in Chapter 2 (*The Impacts of AI*) that the global scale and complex value chains of AI systems make it difficult to effectively intervene in the impacts that AI causes to society and the environment. These impacts include improvements to process efficiency and quality of products and services, reinforcement of systemic discrimination and injustice, erosion of social trust and democratic institutions, labor exploitation and labor displacement, losses of agency and freedoms, and intensifications of climate change and ecological degradation. We determined that these impacts are co-created by a variety of actors–such as data subjects and data providers, model and application developers, model and application users, data center operators, hardware manufacturers, governments, tech companies, investors, communities, and workers–who input resources into and receive resources from AI systems. These resources include many forms of software resources, hardware resources, human resources, financial resources, and governance resources that are integrated through AI systems over the course of their lifecycles. We identified ethical concerns and governance implications of AI systems that arise from these resourcing activities, and we explored opportunities for conducting further research into the impacts and ethics of AI value chains, for applying more theories and methods to study the value chains of specific AI systems, and for intervening in AI impacts by implementing ethical sourcing practices throughout the AI value chain.

In Chapter 3 (*AI Governance Initiatives in Canada*), we more closely analyzed how and why governments are launching AI governance initiatives to intervene in the impacts caused by AI systems. Looking at 84 federal and provincial AI governance initiatives launched in Canada from 2017-2022, we analyzed government programs, policy initiatives, strategic plans, standards, and ethics statements that have been created to intervene in a variety of impacts caused by AI systems. These include impacts to Canadian industry and innovation, technology production and use, public administration, education and training, social and workforce wellbeing, research creation, data governance, and digital infrastructure. We determined that Canada's AI governance initiatives face challenges related to co-governance and coordination between actors, narrowly scoped problem spaces and governance solutions, opacity of strategic goals and governance activities, and barriers to securing public trust in AI systems and their governance. In response, we recommended opportunities for researchers to strengthen Canada's AI governance initiatives by studying the post-implementation outcomes of initiatives more closely, by studying drivers of public trust and mistrust more closely, and by studying the effects of prioritizing intervention in some types of AI impacts over other types (e.g., prioritizing government intervention in AI impacts on industry over impacts on the workforce). We also recommended opportunities for public servants and policymakers to strengthen Canada's AI governance initiatives by specifying initiative success measures and post-implementation outcomes more clearly, by collaborating with civil society actors and the general public more closely, by intervening in a greater variety of AI impacts, and by launching a new initiative to better unify the many activities occurring across Canada's national AI governance system. Although our study is situated within the Canadian context and within a particular timeframe, our findings and their implications for practice and policy are transferable to other AI governance contexts facing similar challenges.

In Chapter 4 (*AI Governance Systems in Canada*), we determined that the AI governance initiatives and activities introduced in Chapter 3 exist within larger systems of governance that are composed of many structural and functional components and span multiple scales of organizational activity. We synthesized theoretical perspectives on AI, AI governance, and service systems to conceptualize AI governance systems as service systems: interdependent sets of actors, activities, resources, networks, logics, functional bounds, and rules that enable intervention in the beneficial and harmful impacts of AI systems within a particular context. We applied this framework to collect and analyze data from interviews with 20 Canadian AI governance leaders and subject matter experts. We found that Canada's national AI governance system is operated by networks of organizational and socio-technical actors who integrate five types of resources with one another (knowledge and cognitive resources, policy and legal resources, data and computational resources, financial resources, and cultural resources) and interact with external systems at other scales (international, subnational, sectoral, and organizational AI governance systems) in order to intervene in the impacts of AI systems. Crucially, we determined that these actors and their activities are structured by strategic, political, economic, and cultural logics for desiring or opposing AI adoption. Their capacities to act upon those logics are limited by functional, rule-based, and norm-based limitations. Based on these findings, we identified potential directions for future AI governance systems research, as well as potential interventions for addressing four high-level gaps in Canada's national AI governance system: gaps in institutional arrangements, gaps in resource availability, gaps in alignment between actors, and gaps in diversity, equity, and inclusion.

From the findings presented in Chapters 2, 3, and 4, we contributed a total of 17 recommendations for advancing future AI governance research and practice to researchers, practitioners, and policymakers. In summary, the recommendations are:

Recommendation 1: Researchers should conduct more empirical and action research into the specific ethical concerns, value chain actors, and resourcing activities.

Recommendation 2: Researchers should develop and apply theories and methods for systematically modeling AI value chains, analyzing a diverse range of ethical concerns in those value chains, and enacting interventions in those value chains.

Recommendation 3: Practitioners and policymakers should design and implement ethical sourcing practices across value chains that provide resource inputs to or receive resource outputs from AI systems.

Recommendation 4: Researchers should study the outcomes of Canada's AI governance initiatives.

Recommendation 5: Researchers should study challenges to public trust in Canada's AI governance initiatives.

Recommendation 6: Researchers should study the effects of AI impact representation on the outcomes of Canada's AI governance initiatives.

Recommendation 7: Policymakers and public servants should specify success measures for initiatives and routinely publish information on the outcomes of Canada's AI governance initiatives.

Recommendation 8: Policymakers and public servants should collaborate more directly with the public on designing and implementing Canada's AI governance initiatives.

Recommendation 9: Policymakers and public servants should account for a greater variety of AI impacts when designing and implementing Canada's AI governance initiatives.

Recommendation 10: Policymakers and public servants should launch a new initiative to cultivate a more unified national approach to AI governance in Canada.

Recommendation 11: Researchers should conduct additional analysis of topics in our dataset of perceptions of Canada's national AI governance system.

Recommendation 12: Researchers should further investigate institutional & ecosystem phenomena in Canada's national AI governance system.

Recommendation 13: Researchers should apply our AI governance systems analysis framework, data, and findings to other AI governance research contexts.

Recommendation 14: Practitioners and policymakers should implement new collaboration & coordination mechanisms in Canada's national AI governance system.

Recommendation 15: Practitioners and policymakers should create guidance for designing and implementing participatory AI governance initiatives in Canada.

Recommendation 16: Practitioners and policymakers should expand access to key resources needed for effective AI governance practices in Canada.

Recommendation 17: Practitioners and policymakers should advance diversity, equity, and inclusion in Canada's AI governance activities.

If effectively implemented, these 17 recommendations could support greater prevention of harms caused by AI systems. However, effective implementation of these 17 recommendations will face significant barriers in navigating the power relations inherent to AI systems and AI governance systems. In Chapters 2, 3, and 4, we determined that imbalances in political power, economic power, computational power, cultural power, and other forms of power between actors often present barriers to enacting change in AI systems and AI governance systems. Additional barriers to change are imposed by norms of exercising power to organize actors and to provide them with resources to do the work of governance.

Transformative change–change in which unjust norms and power structures are dismantled to prevent further harm to historically marginalized groups–is especially difficult to enact in AI systems and AI governance systems. These systems are built atop a stack of other unjust computational, technological, and social systems that long preceded the recent advent of data-intensive machine learning applications (Crawford, 2021; Crawford & Joler; 2023;

Crawford & Joler, 2019). How, then, can AI governance possibly be made more effective at preventing harms–and particularly, at preventing harms to marginalized communities–if studies, practices, and systems of AI governance are all so limited by power structures and by histories of domination and oppression?

There is no simple or clear answer to this question. However, as I've reflected on the findings and implications of my research, some more precise explanations for the ineffectiveness of AI governance have emerged, along with some necessary pathways forward for re-imagining what "AI governance" fundamentally is and what the work of AI governance ought to entail. The remainder of this concluding chapter details those explanations and future pathways through three thematically interlinked reflections on the research presented in this dissertation. In keeping with this dissertation's overarching adherence to a trans feminist standpoint epistemology, these explanations and future pathways are presented as a self-reflexive account that combines my lived experience of studying and practicing AI governance with empirical evidence presented in chapters 2, 3, and 4 and in additional research literature.

5.2. Reflections on Research: Further Discussion of Findings & Future Pathways

5.2.1. Reflection 1: AI governance is an exercise of power

As we have seen in the research presented across chapters 2, 3, and 4, AI governance is an exercise of power. AI governance frequently proves ineffective at preventing AI systems from harming people and the environment because studies, practices, and systems of AI governance are limited by various forms of power relations. These power relations include resource dependencies, rule-making and rule-taking activities, and norms of political and economic inclusion. Power relations are deeply embedded in the networks of actors and the logics, bounds, rules, and norms involved in their co-creation of AI governance activities. The effectiveness of an AI governance activity at preventing AI systems from causing harm is therefore derived from the perceptions and powers of the actors included in that activity: AI governance emerges from the underlying values and logics of those actors, the qualities those actors perceive as constituting “effective” vs. “ineffective” AI governance, and the resources and capacities those actors are able to exercise to enact their shared values and logics.

In researching and writing Chapter 2, I often felt overwhelmed by the immense challenges of context and scale implicated in the societal and environmental impacts of AI systems. The resources required to develop and operate AI systems are distributed across global value chains. This results in complex issues of jurisdiction and transnational impact that AI governance researchers, practitioners, and policymakers must contend with. Industry actors making voluntary commitments to more robust ethical sourcing standards at socially and environmentally impactful points throughout the AI value chain (as suggested in *Recommendation 3*) could help to address challenges of transnational context and global scale in principle. In practice, however, effective global implementation of ethical sourcing practices faces political and economic barriers arising from the power relations embedded in AI systems and their governance. Veale, Matus, and Gorwa (2023) highlight several tensions limiting the global implementation of ethical and regulatory standards for AI. These tensions include expertise consolidation and lobbying by dominant industry actors seeking to influence regulation, outsourcing of labor and development tasks to extraterritorial regulatory havens, and perverse incentives for obtaining foreign investment, resulting in race-to-the-bottom dynamics in the development of regulations and standards.

Given these tensions of political and economic power in AI governance, voluntary adoption of robust ethical standards for preventing harms to society and the environment across global AI value chains will have limited effectiveness. Although governments could impose legal and regulatory requirements for industry actors to comply with more robust ethical standards across value chains, these top-down pushes for comprehensive, compliance-based regulatory frameworks are also limited by power relations. In a forthcoming paper I co-authored with Helen A. Hayes entitled *Valuing Value Chains: The Governance of AI Value Chains in Canada* (Attard-Frost & Hayes, 2025), we compare Canada's approach to regulating the value chains of AI systems to the regulatory approaches of the European Union and the United States. Relations of political and economic power have limited the effectiveness of AI regulation across all three of these jurisdictions. We argue that although the European Union's *AI Act*–a general-purpose, cross-sectoral regulation for AI systems–was developed through extensive consultations with government, industry, and civil society stakeholders, the final text of the EU AI Act contains limited protections for workers and the environment. In addition to those limitations, some have criticized the EU AI Act for containing limited protections for human rights, and for scaling back the compliance requirements and protections set forth in earlier versions of the AI Act in response to pressure from industry stakeholders and lobbyists (Access Now, 2024; Amnesty International, 2023; Corporate Europe Observatory, 2023).

Similar power dynamics can be observed in the United States, where there is no general-purpose regulatory initiative comparable to the EU's AI Act, but rather, a patchwork of executive, administrative, and legislative initiatives at multiple levels of government and with varying degrees of legal force. Some voluntary instruments such as the White House's *Blueprint for an AI Bill of Rights* (The White House, 2022) and the National Institute of Standards and Technology's *AI Risk Management Framework* (NIST, 2023) were created in consultation with several public sector, private sector, and civil society organizations. However, the US AI industry exercises considerable political, economic, and cultural power across the nation's AI governance system. Companies that dominate the AI industry–such as Google, Amazon, Microsoft, and OpenAI–shape US AI governance to suit their strategic priorities and economic interests through regular engagement with government officials, extensive participation in the development of policy and regulation, and privileged access to multistakeholder meetings organized by Congressional and White House committees (Henshall, 2024; Hine, 2024; Johnson, 2023; Merica, 2024; The White House, 2023). Simultaneously, the US AI industry also works to normalize futurological narratives about AI's supposedly utopian promise and apocalyptic peril–and the ability of AI developers to navigate society away from peril and toward greater promise–as a cultural pretext for policymaking and for media coverage of AI governance activities (Altman, Brockman, & Sutskever, 2023; Barakat, 2024; Dandurand et al., 2023; Gebru & Torres, 2024; Herrman, 2024; OpenAI, 2023).

In researching and writing Chapter 3 of this dissertation, I became more aware of the extent to which industry power also dominates the AI governance priorities of Canada. This in itself is not a new finding: in her landmark empirical study of public and private investment activity in Canada's AI ecosystem, Ana Brandusescu (2021) demonstrates clearly that "concentrations of power provide advantages to a handful of entities with financial resources, data, and technologies across a few universities and affiliated research nonprofits, startups, and international (big) tech companies" (p. 7). The findings we present in Chapter 3 build upon and validate her earlier findings regarding power imbalances across Canada's AI governance initiatives.

Chapter 3 demonstrates that Canadian federal and provincial AI governance initiatives launched from 2017 to 2022 prioritize the development of Canadian industry, technological innovation, and economic growth over protections for vulnerable social groups, workers, and the environment. Even policy instruments created for the ostensible purpose of preventing harms caused by AI systems–such as the *Artificial Intelligence and Data Act* (AIDA) and the *Directive on Automated Decision-making*–have substantial gaps in the scope of harms they cover and in their enforceability (Attard-Frost, 2022; Attard-Frost, 2023a; Brandusescu & Sieber, 2022; Scassa, 2021; Scassa, 2023; Tessono et al., 2022). The AIDA in particular has been widely criticized for failing to meaningfully include civil society and marginalized communities in its policy development process, for not representing a wide range of collective harms and environmental harms in its scope, and for failing to guarantee adequate protections against harms to vulnerable groups caused by AI systems (The Dais & Centre for Media, Technology and Democracy, 2023).

Some more recent Canadian AI governance initiatives address a greater scope of societal and environmental impacts. For example, the *Voluntary Code of Conduct on the Responsible Development and Management of Advanced Generative AI Systems* was published by Innovation, Science and Economic Development Canada (ISED) in September 2023 (Innovation, Science and Economic Development Canada, 2023a). The Code of Conduct is intended to secure voluntary commitments from developers, deployers, and operators of generative AI systems to ensure safety, accountability, transparency, fairness and equity, human oversight, validity, and robustness in their systems. Also in September 2023, the Treasury Board of Canada Secretariat (TBS) published their *Guide on the use of Generative AI* (Government of Canada, 2024). The TBS Guide provides federal public servants with voluntary guidance for ensuring quality, protection of information, and human autonomy in their use of generative AI systems, as well as for distinguishing humans from machines and for mitigating bias, legal risks, and environmental impacts of generative AI. In addition to those voluntary instruments, amendments were proposed to the AIDA by the Minister of ISED in November 2023 to address gaps in the legislation (Minister of Innovation, Science and Industry, 2023). The amended version of the AIDA added specificity to the legislation's definition of "high-impact systems," added requirements for "general-purpose systems" (such as OpenAI's GPT systems), clarified stakeholder responsibilities and compliance requirements across the AI value chain, clarified the role of the proposed "AI and Data Commissioner" who would be tasked with enforcing the legislation, and moved the legislation's provisions and requirements into closer alignment with international regulatory frameworks such as the EU's AI Act.

As I reflected on these more recent initiatives and amendments to the AIDA, I noticed that although their scope represented a more clear and wide range of AI impacts than earlier initiatives, many underlying challenges of industry prioritization, public participation, and public accountability still remained. The ISED Code of Conduct was criticized for an extremely short consultation period of less than one month, for being developed out of invitation-only meetings with a limited group of civil society and industry stakeholders, and for limitations in its scope, accountability measures, and enforceability (Attard-Frost, 2023b; Innovation, Science and Economic Development Canada, 2023b; Karadeglija, 2023). The ISED Code of Conduct and the TBS Guide both continue to contribute to a lack of transparency and accountability in the outcomes of Canadian AI governance initiatives: although the TBS Guide has received some praise for its ambitious scope, and although 40 organizations have voluntarily signed onto the ISED Code of Conduct as of January 2025, it is not clear if these initiatives have had any significant effect on the operations of Canadian AI developers and users. There has been no formalized post-implementation monitoring, reporting, or evaluation of the effectiveness of these guidance documents at achieving their goals of risk mitigation and harm prevention.

Meanwhile, the Minister of ISED's proposed amendments to the AIDA were met with mixed reception: while some changes to the legislation were praised, there were also concerns that the amendments still did not sufficiently address criticisms of the legislation's gaps in scope, specificity, enforcement powers, and public consultation (Canadian Association of Professional Employees, 2024; Canadian Chamber of Commerce, 2024; OpenMedia et al., 2024). Some critics observed that the consultation activities that did occur greatly overrepresented industry actors and their interests (Castaldo, 2023; Clement, 2023). In their written brief to the parliamentary committee studying the legislation, the Assembly of First Nations (2023) announced that they may take legal action against the government for failing to uphold a constitutional responsibility to consult with Indigenous peoples during the drafting of the legislation.

As the parliamentary process of studying and amending the AIDA proceeded into 2024, the federal government's 2024 budget announcement further underscored the government's priorities: $2 billion was to be invested in expanding computing infrastructure for Canadian industry and AI researchers, $200 million was allocated to commercializing AI technologies and accelerating AI adoption across critical sectors, while only $50 million was allocated to "supporting workers who may be impacted by AI, such as creative industries" (Prime Minister of Canada, 2024). The proroguing of Parliament in January 2025 has cast some uncertainty upon the future of AI legislation and governance in Canada. However, these more recent developments in Canadian AI governance confirmed to me that our findings in Chapter 3 from reviewing 84 initiatives launched from 2017 to 2022 have remained valid: Canadian AI governance initiatives are still prioritizing the needs of industry over broader societal needs, are still failing to ensure meaningful participation from and accountability to the public and marginalized communities, and are still ineffective at preventing AI systems from causing harms to society and the environment. In Canada, support for and development of industry power is treated with greater importance than public empowerment.

5.2.2. Reflection 2: Top-down AI governance is limited by power imbalances

My experience witnessing repeated shortcomings of government accountability, transparency, and public participation across so many Canadian AI governance initiatives made me increasingly concerned that the Canadian nation-state may be too structurally and functionally limited to effectively prevent AI systems from harming society and the environment. My concern was compounded in researching and writing about Canada's national AI governance system for Chapter 4 of this dissertation. Conducting interviews with government leaders and subject matter experts throughout 2023 for Chapter 4 further clarified for me that Canada's industry-first approach to AI governance was in large part a product of the political, economic, and cultural power that industry exercises over government.

During these interviews, the logics, bounds, and rules that participants described to me made it strikingly clear that Canada's national AI governance system is shaped and constrained by imbalanced power relations. Several logics for engaging in AI governance activities that were described most frequently by participants pointed toward the ability of dominant industry actors to steer the direction of Canadian AI governance by exercising their overbearing political and economic power. *Maximize profit & shareholder value from AI adoption* (#7.4 in Appendix 4A, described in 6 interviews) indicates that AI governance is often practiced in Canada with the goal of securing financial benefits and greater economic power for private interests. *Achieve balance* *between public & private interests in AI outcomes* (#7.3 in Appendix 4A, described in 6 interviews) indicates that AI governance practices intended to secure more beneficial AI adoption outcomes for the public are somehow reconcilable with the goal of securing financial benefits and greater economic power for private interests. This balancing of public and private interests functions through a variety of other logics perceived by the participants, most notably, *strengthen public/consumer trust in AI applications* (#7.7 in Appendix 4A, described in 4 interviews) and *ensure AI development & use contributes to economic development* (#7.11, 3 interviews). The rationale implied by these logics is highly complex and speculative: by stimulating AI adoption and ensuring AI applications are trusted by Canadian consumers and the general public, the development and use of those trustworthy AI applications will contribute to Canada's economic development, which will in turn contribute to economic benefits being accrued by and fairly balanced between private stakeholders and the general public. In the process of analyzing these logics and writing Chapter 4, I reflected on a part of our analysis from Chapter 3:

> In Canada's stimulation approach to AI governance, societal benefit is typically assumed to be a secondary epiphenomenon of economic benefits accrued through technological innovation. There is no empirical evidence to support such assumptions about the socio-economic impacts of AI systems. In fact, many studies of the political economy of AI indicate that without broad-based and cross-cutting interventions in industries, technologies, societies, workforces, and digital infrastructures, a stimulation approach to AI governance might instead cause negative societal impacts. Without adequate counterbalances, expanding industry's capacity to develop and use AI systems may

compound existing concentrations of capital, technology, data, and other resources in a small handful of dominant industry actors (pp. 84-85 of this dissertation).

This empirically tenuous *rising tide lifts all boats* rationale is conditional on the success of a long series of complex, well-coordinated policy interventions intended to counterbalance the domineering power of industry. Despite the many speculative leaps this rationale requires, it is so deeply embedded into the logics of Canada's national AI governance system that it functions as a cultural narrative: it is an implicit pretext for engaging in AI governance activities that is observable across multiple studies, and it is made durable and enduring by strong support and repeated deployment of the narrative from private sector and public sector organizations. Other logics that were frequently perceived by participants in the research conducted for Chapter 4 also function in service to this cultural narrative, such as *align AI systems with organizational values* (#7.5, 5 interviews), *cultivate a trustworthy brand image* (#7.13, 3 interviews), *enable greater access to international markets* (#7.14, 3 interviews), and *create AI systems that are interoperable across markets & jurisdictions* (#7.15, 3 interviews). Values-based AI adoption, brand trustworthiness, interoperability, and market access do not primarily serve the public interest: these logics for engaging in AI governance primarily serve the AI adoption goals of Canadian industry and their private shareholders. These logics are premised on a speculative narrative that a balance of public and private benefits will somehow be achieved as a downstream effect of well-calibrated organizational values, trustworthy brands, and interoperable AI markets.

In conducting research interviews for Chapter 4, the bounds and rules described by participants further concerned me, as they signaled that Canada's national AI governance system has been structured by powerful norms that may already be too solidified to be susceptible to transformative change. Limitations of *AI governance knowledge and expertise* and *financial resources* within organizations are widespread (#8.1 and #8.5 in Appendix 4A, described in 10 and 5 interviews), making it difficult for smaller private sector and civil society organizations to effectively practice AI governance and meaningfully participate in AI governance initiatives. AI policy development processes are strongly influenced by the norms and cultures of organizations (#9.2, 7 interviews), meaning that if a small handful of powerful public sector organizations co-create AI policy with a small handful of powerful private sector organizations, a result will be that regular involvement of private interests in public policy processes will become a cultural norm. This emerging norm of private control over public AI policy is reflected in the domain of regulatory development, where the *privatization of audit, compliance, & regulatory services* (#9.5, 4 interviews) is a contentious issue. The normalization of dominant industry actors steering the politics and economics of Canada's national AI governance system also reinforces existing *norms of & imbalances in political & economic power* (#9.6, 4 interviews), contributing to an *emerging norm of accountability & enforceability gaps* (#9.8, 3 interviews) wherein the public sector becomes so dependent on specialized AI governance knowledge, expertise, and other resources provided by the private sector that it is normal for the public sector to design AI policy instruments with light-touch accountability requirements and weak enforcement capabilities. If Canada's national AI governance system so greatly prioritizes the regulatory needs and perspectives of dominant industry actors, then it will be difficult to effectively prevent AI systems from causing harm to marginalized communities and society more broadly.

Many of the government leaders and subject matter experts I interviewed were optimistic that power imbalances in Canada's national AI governance system can be counterbalanced through top-down government intervention. I am less optimistic. Though participants suggested opportunities for improvement such as *create more opportunities for public participation in AI governance* (#11.2, 7 interviews), *more diversity in AI governance activities* (#11.5, 6 interviews), *implement stronger participatory design & governance practices* (#11.6, 6 interviews), and *stronger public accountability & oversight in governance activities* (#11.10, 4 interviews), enacting these opportunities at a national scale will require a protracted and concerted effort from public sector, private sector, and civil society organizations. Working together closely, these organizations will need to build greater power and resources for marginalized groups, along with greater capacities for participatory governance, co-regulation by a diversity of actors, proactive accountability measures, and robust protections against harms caused by AI systems. Unfortunately, this kind of diverse, equity-seeking approach to collaboration seems highly unrealistic: in the case of powerful government institutions and industry actors that benefit from power imbalances in Canada’s national AI governance system, this would in many cases mean acting against their own established norms and their own perceived interests.

In addition to conflicting interests and political challenges, a national scale effort to diminish power imbalances in AI governance would encounter practical challenges navigating the gaps in coordination and collaboration discussed in Chapters 3 and 4. The resource integration networks and governance networks needed to support broadly inclusive and equitable collaboration across governments, sectors, and civil society are not well developed, making *Recommendations 8* and *17* difficult to implement effectively (collaborate more directly with the public on designing and implementing Canada’s AI governance initiative, and advance diversity, equity, and inclusion in Canada’s AI governance activities). In principle, implementation of *Recommendations 10* and *14* (implementing new collaboration and coordination mechanisms, and launching a new initiative to cultivate a more unified national approach to AI governance) would further develop these networks, creating the infrastructure needed to enable more participatory forms of governance, accountability, and regulatory protection. In practice, however, implementing these recommendations would be a resource-intensive endeavor, again requiring dominant actors to act against their own established norms and perceived interests by empowering actors at the margins of Canada's national AI governance system with greater agency to change the system's structures and operations. Without clear incentives for dominant actors to empower civil society and marginalized communities, other recommendations directed at public servants and policymakers are also likely to have limited effectiveness. The effective implementation of *Recommendations 15* and *16* (create guidance for participatory AI governance initiatives, and expand access to key resources needed for effective AI governance practices) are dependent on powerful actors with existing networks of knowledge, financial, computational, legal, and policy resources voluntarily redistributing their resources to relatively less powerful actors. The effective implementation of *Recommendation 7* (specify success measures for initiatives and routinely publish information on the outcomes of initiatives) is dependent on government institutions and individual public managers voluntarily subjecting their operations to greater public scrutiny, as well as on the existence of civil society organizations that have the resources and capacities needed to hold government institutions accountable when they are not transparent about success measures and outcomes. The effective implementation of *Recommendation 9* (account for a greater variety of AI impacts in governance initiatives) is dependent on government institutions going against established political, cultural, and epistemic norms of prioritizing the impacts of AI systems on industry and technological innovation in their AI governance activities over other societal and environmental impacts.

As I continued reflecting on the embedding of these power imbalances and resource dependencies in Canada's national AI governance system, I recognized that *top-down AI governance*–a system of governance in which marginalized communities are dependent upon resources and leadership provided to them by brokers of state power and industry power–is not sufficient for enacting a transfeminist approach to AI governance. To enact transfeminist principles of transformative change, anti-normativity, agency, fluidity, community, solidarity, and resistance in AI governance–and by extension, to more effectively prevent AI from causing harm to marginalized communities through the application of transfeminist AI ethics–systems of *bottom-up AI governance* will be necessary. While top-down AI governance is set in motion by alliances of state power and industry power, bottom-up AI governance is set in motion by the power of collective action at smaller scales of organizational activity, such as communities and workplaces.

AI governance activity already occurs regularly at smaller scales. A recent flashpoint in bottom-up AI governance can be seen in the resistance of creative communities and workers against harmful generative AI systems. Popular text-generating, image-generating, and audio-generating AI applications such as ChatGPT, DALL-E, Stable Diffusion, Midjourney, and MuseNet are trained on millions of copyrighted works that have been scraped from the web and used for model training without consent from or compensation to the creators of those works (Jiang et al., 2023). In response, online communities and social movements such as #CreateDontScrape (Create Don't Scrape, 2024) and labour unions such as the Writers Guild of America (WGA) and Screen Actors Guild - American Federation of Television and Radio Artists (SAG-AFTRA) have collectively organized against developers and operators of generative AI systems. The 2023 WGA strike resulted in the union establishing regulations for the training of generative AI models on union-protected materials and for workers to decide if and how generative AI should be used in their own work processes (Writers Guild of America, 2023). The 2023 SAG-AFTRA strike resulted in regulations being established for the development and use of synthetic digital replicas of performers (Patten, 2023). These community-led and worker-led initiatives offer a powerful lesson for the future of AI governance: if top-down AI governance systems are structurally limited by power imbalances, bottom-up AI governance can provide communities and workers a viable alternative to depending on resources and effective regulatory leadership from state and industry actors.

5.2.3. Reflection 3: AI governance must be transformed

AI governance research, practices, and systems must be radically transformed to more effectively prevent harm to marginalized communities. In her analysis of state-led and industry-led initiatives for governing generative AI systems, Inga Ulnicane (2024) coins the term "governance fix" to describe highly centralized, top-down initiatives that enact a "narrow and technocratic approach" to AI governance (p. 1). Instead of governance fixes that reinforce existing power imbalances, Ulnicane recommends enacting more pluralistic, polycentric, and participatory approaches to co-governance that seek to transformatively change structural imbalances of power. Mere governance fixes–new initiatives, policy amendments, and minor operational changes within the narrow perimeters of existing power structures–are insufficient without a deeper re-imagining and reconfiguration of the structures, norms, and possibilities of AI governance. By re-imagining AI governance through a transfeminist lens, we can begin to envision alternative possibilities for AI governance that are less harmful and more effective, with effectiveness understood in relation to transfeminist principles of transformative change, anti-normativity, agency, fluidity, community, solidarity, and resistance.

In light of the research presented in this dissertation, I outline here a preliminary agenda for a transfeminist approach to AI governance that seeks to transformatively change future AI governance research, practices, and systems. This agenda contains three high-level actions for researchers, practitioners, policymakers, community organizers, advocates, and other stakeholders to implement in parallel: (1) Expand the scope of AI and AI governance, (2) Oppose AI governance systems that exclude marginalized communities, (3) Co-create bottom-up AI governance systems to reduce dependency on industry and the state.

Action 1: Expand the scope of AI and AI governance. Chapters 2, 3, and 4 all show that "AI" is an ambiguous phenomenon. The meaning of AI–once a purely theoretical concern for an esoteric enclave of computer scientists, philosophers, and futurists–has in recent years become a matter of enormous practical and political consequence. In her 2021 book *Atlas of AI*, Kate Crawford observes that "each way of defining artificial intelligence is doing work, setting a frame for how it will be understood, measured, valued, and governed" (p. 7). The ways in which the term AI is interpreted has a cascading effect on how AI governance is studied, practiced, and systematized. The ontological boundaries of AI that are encoded into now-emerging AI laws, policy instruments, and standards will cascade into the system design requirements of AI developers and operators, as well as into business rationales for public and private sector AI investment, procurement, and management decisions. Theoretical differences in how to understand and interpret the meaning of AI will ultimately result in practical differences: AI systems will be developed, used, and governed differently based on what those systems are understood to consist of. If AI is understood as consisting purely of software resources–such as data, algorithms, machine learning models, and neural networks–then AI developers, users, and regulators will be bounded within a very different governance system than if AI were interpreted as also consisting of, for example, the human labor and knowledge, hardware components, digital infrastructure, energy, water, and minerals required to develop and operate AI systems.

Expanding the ontological perimeter of AI to encircle a broader range of phenomena supports more interpretive frames, discursive spaces, values and norms, and practical possibilities for AI governance. By allowing ambiguity and fluidity in how the term AI is perceived, interpreted, discussed, and acted upon, the innate *transness* of AI can function as a foundation for more inclusive forms of AI governance. In a 2023 essay, transfeminist scholars Mijke van der Drift and Nat Raha write that "radical transfeminism embraces trans as active and anti-normative, rather than defined as a stable form . . . Trans is thus a dynamic formation, which does not lay claim to simply *be*, but which functions by disrupting static categories of being" (pp. 13-16). At a time when so many new laws, policy initiatives, and technical standards are attempting to stabilize AI into a well-defined form (see for example the definitions of AI that form the basis of the EU's *AI Act*, Canada's *Artificial Intelligence and Data Act*, and the ISO/IEC 42001:2023 standard for AI management systems), it is imperative that a transfeminist approach to AI governance resists forced stabilization into narrow techno-legal definitions and forms of governance. A transfeminist approach to AI governance can instead embrace the transness of AI: the ambiguity, the fluidity, and the plurality of meaning embedded in the term AI all provide affordances for more diverse understandings of AI, epistemic communities, governance systems, and justice-seeking projects to form within the vast discursive nebula we refer to as "AI."

Like AI, the concept of "AI governance" also has some innate transness. AI governance is a dynamic and ambiguous phenomenon, simultaneously referring to many possible ways of researching, practicing, and systematizing interventions in "AI." In this dissertation, some disambiguation has been provided by framing AI governance as a system of practices intended to maximize benefits and minimize harms caused by AI systems. However, this framing is itself deliberately ambiguous, intended to transform AI governance from a domain dominated by legal, technical, business, and policy experts into a more open discursive space that can accommodate diverse perspectives, exploratory questions, and alternative answers. What types of benefits are included in AI governance, and why? What types of harms are included, and why? Who benefits and who is harmed, and through what types of practices do they intervene in AI systems to maximize benefits and minimize harms? Chapters 2, 3, and 4 all provide some potential answers to those questions, but the amorphous ontological perimeter of AI and the fluid contexts, scales, and scopes of AI governance systems leave space for many alternative answers. There is no definitive, objective answer to what AI governance is and what the work of AI governance ought to entail: in a service realist view, ontologies and ethics of AI governance are co-created by many networks of actors applying their many resources, logics, rules, and norms with the intent of planning and enacting governance activity. In Chapter 4, we saw that AI governance contains a plurality of values, perspectives, and practices. It could be said that writing and reading this dissertation are both practices of AI governance: we write and read the research presented here with the intent of cultivating greater knowledge and cognitive resources that can be applied to maximize benefits and minimize harms caused by AI systems.

AI governance is not the preserve of well-funded experts, academics, lawyers, and public intellectuals. A transfeminist approach to AI governance resists what van der Drift and Raha (2023) refer to as “neoliberal encapsulation:” the containment of transness to “the future that liberalism has wanted for us – a future based in limited forms of social inclusion and legal rights” (p. 13). A transfeminist approach to AI governance works against neoliberal encapsulation by expanding the ontological scope of AI governance to include a more comprehensive set of actors and activities than what is typically found within the narrow technocratic confines of state-led and industry-led AI governance initiatives. Communities and workers impacted by AI systems all possess their own domain expertise: they possess context-sensitive, direct knowledge of AI impacts and resource needs within their communities and workplaces. Collective organizing of communities and workers, community-led policy and strategy initiatives, direct action against harmful AI systems and governance systems, development of small-scale resource-sharing and capacity-building networks, discursive interventions, and public protest can all be directed toward maximizing benefits and preventing harms caused by AI systems. Therefore, these are all legitimate practices of AI governance. By expanding the perimeter of AI governance to encompass this larger scope of practical possibilities, AI governance is transformed into a more inclusive democratic project, one that escapes neoliberal encapsulation. Rather than consolidating more power, expertise, and agency into already powerful state and industry organizations, a transfeminist approach to AI governance must build greater agency and solidarity within marginalized and under-resourced communities that have so far had their needs largely ignored by state and industry actors.

Action 2: Oppose AI governance systems that exclude marginalized communities. As it has become more clear that AI systems frequently cause harm, a vibrant discussion has emerged regarding the accountability processes through which impacted actors may oppose AI developers and operators whose systems caused harm to them. Principles of public accountability, contestability, and resistance are of particular importance in these discussions, and conceptual frameworks such as “algorithmic resistance” (Bonini & Treré, 2024; DeVrio, Eslami, & Holstein, 2024; Ganesh & Moss, 2022; Velkova & Kaun, 2021), “resisting AI”

(McQuillan, 2022) and "contestable AI" (Alfrink et al., 2023; Alfrink et al., 2024; Balayn et al., 2024) have emerged as banners under which diverse practical possibilities for opposition against AI systems are studied. Although these frameworks share a common interest in studying methods through which the developers and operators of AI systems can be opposed by vulnerable and harmed stakeholders, contestability-based frameworks and resistance-based frameworks do have some crucial differences and limitations.

Alfrink et al. (2023) characterize *contestability in AI* as distinctively procedural and mechanistic. Contestability requires institutional mechanisms and affordances for open and iterative debate about AI design practices, responsiveness to complaints about an automated decision or other output of a particular AI system, and commitments to remediation and procedural justice throughout the lifecycle of the system. Although contestable AI provides a process-oriented approach for opposing context-specific harms caused by AI systems, the effectiveness of contestability mechanisms is largely dependent upon voluntary commitments to participatory design and procedural justice from AI developers, operators, and policymakers–many of whom would, again, be acting against their own perceived interests by empowering marginalized actors with robust mechanisms with which to contest their practices of AI development, management, and policymaking. Indeed, Alfrink et al. imply that contestable AI faces complex structural barriers to democratic control even in public sector contexts where it is more likely to be effective, noting the "duty of care government organizations have toward citizens; and the (*at least nominal*) democratic control of citizens over public organizations" (p. 632, emphasis added). Meanwhile, *resistance against AI* refers to a broader category of oppositional acts that emerge from structural conflict with power within an AI system, including (but not limited to) acts of contestation and complaint. In their review of the research literature on algorithmic resistance, Bonini and Treré (2024) observe that resistance against AI systems and other data-intensive algorithmic systems entails an *agonistic* relationship with power, a type of structural relation between dominant and marginalized actors that is simultaneously both frictive and generative of new possibilities for building greater agency amongst marginalized actors. Bonini and Treré are careful to diminish the revolutionary romanticism of resistance, noting that although resistance is "born of a response to existing systems of power, control, and domination," (pp.17-18), acts of resistance are not limited to historically marginalized actors: resistance is practiced by a range of actors seeking to cultivate greater agency that they can then exercise to enact a variety of benevolent and malevolent intentions.

The limitations of these frameworks led me to write an essay entitled *AI Countergovernance* intended for a wide audience of researchers, practitioners, public servants, community organizers, and activists seeking strategies and tools for opposing harmful AI systems (Attard-Frost, 2023c). In the essay, I apply transfeminist principles of fluidity, agency, community, resistance, and transformative change to outline an integrative framework of oppositional action that reconciles the mechanistic view of contestability-based frameworks with the structural view of resistance-based frameworks. I build upon Dean's (2018) analysis of contestability mechanisms and resistance practices involved in building counterpower against state actors whose governance systems failed to serve the needs of their publics (a process that Dean refers to as "counter-governance"), introducing *AI countergovernance* as a conceptual and practical framework for opposing AI governance systems that exclude the needs of marginalized communities. Synthesizing lessons learned from four real-world cases of collective organizing against state-led and industry-led AI governance systems (see Table 5.2), I argue that communities and workers can more effectively oppose harmful AI systems by organizing against their higher-level governance systems (e.g., by organizing against the full spectrum of organizational actors, networks, logics, and resources responsible for co-creating the harmful AI system), rather than by centering contestation and resistance around the more granular technological components of the AI systems and the socio-technical practices involved in designing, developing, deploying, and operating those components (e.g., the development and use of harmful datasets, algorithms, machine learning models, and computing infrastructures).

Table 5.2: Key lessons learned from the four cases of AI countergovernance discussed by Attard-Frost (2023c, 2024), highlighting comparative sources of power and counterpower, practices of countergovernance, and outcomes of each case.

Case	**Sources of power & counterpower**	**Practices of countergovernance**	**Outcomes**
Project Maven (Google military partnership with US Department of Defense)	• *Power*: Industry-state partnership • *Counterpower*: Workers	• Open letter to Google CEO • Worker resignations & walkouts • Media engagement	• Project Maven contract was not renewed in 2019 • Re-emergence & expansion of Google-DoD partnerships in subsequent years
Sidewalk Toronto ("Smart city" development project)	• *Power*: Industry-state partnership • *Counterpower*: Local community	• Public meetings • Alternative urban development & data infrastructure plans • Media engagement	• Sustained community backlash against project • Cancellation of project in May 2020 due to multiple factors (including resistance from local community)
Backlash against Canada's Artificial Intelligence & Data Act (AIDA)	• *Power:* State with significant industry influence • *Counterpower*: Public (Primarily policy experts & advocacy groups)	• Open letters to Members of Parliament • Public audits & assessment of the AIDA • Alternative policy proposals • Witness testimony & impact statements • Media engagement	• Sustained public backlash against AIDA • New consultations launched & amendments to AIDA that reflect some criticisms • Many criticisms currently remain unaddressed

Writers Guild of America Strike	• *Power*: Industry • *Counterpower*: Workers	• Collective bargaining • Policy & planning through workplace technology committee • Media engagement	• New collective agreement contains regulations for AI-generated writing, writer use of AI, and use of writers' material in model training

This countergovernance framework first expands the ontological scope of "AI" and "AI governance" to accommodate a more open horizon of practical possibilities, then moves the targets of opposition from the level of AI systems to the level of AI governance systems. In this framework, communities organize against ineffective AI governance systems by working transversally both inside and outside of established institutions, as well as by leveraging a fluid repertoire of governance practices and mechanisms to build agency and counterpower. Practices of AI countergovernance can include: awareness-building through media engagement, protests, petitions, and open letters addressed to sources of state and industry power, such as elected representatives and corporate executives; community-led audits, evaluations, and iterative contestation of harmful AI systems, policy instruments, and organizations; co-creation of shared knowledge resources and funds, guidelines and rules, and direct action campaigns to proactively prevent the development and operation of harmful AI systems within a community. Through its grounding in transfeminist principles and practices of governance, this framework enables actionable strategies of contestation, resistance, capacity-building, and justice-seeking within marginalized communities that are vulnerable to or have been harmed by AI systems.

Action 3: Co-create bottom-up AI governance systems to reduce dependency on industry and the state. Beyond merely opposing ineffective large-scale systems of top-down AI governance, a transfeminist approach to AI governance can also advance the bottom-up development of small-scale alternative systems in more localized contexts, such as communities and workplaces. If developed in accordance with transfeminist principles of transformative change, agency, community, and solidarity, a foundational goal of developing bottom-up AI governance systems should be to gradually reduce the dependency of vulnerable communities on AI governance resources provided by state and industry actors. Transfeminist AI governance must build within communities the resources, networks, and agency needed to exercise greater power over AI systems that pose a risk to the community. As governance practices, actors, networks, resources, logics, rules, and norms within the community gradually become more systematized, power can be more effectively built and exercised to prevent the development and operation of AI systems that may cause harm to the community. This will be a fluid and dynamic process, involving diverse and extensive co-creation activity between community organizers and researchers, practitioners, public servants, journalists, activists, and other sympathetic actors with the intent of supporting the community in building and exercising power.

Researchers can support community organizers with empirical research and action research intended to determine more clear, context-sensitive practices for effective community-led design and implementation of AI governance systems. **Practitioners** of technology design, management, and policy can support by training community organizers with skills and knowledge from their respective disciplines, and by assisting organizers in the design and implementation of projects and programs, policy instruments and guidance, community events and workshops, knowledge resources, technologies, and other governance tools required by the community. **Public servants** can support by directing a greater share of public funding and resources away from industry-centric AI governance initiatives and toward: knowledge-building initiatives and participatory policymaking initiatives in vulnerable communities; civic participation funds to enable under-resourced civil society organizations and communities to more actively participate in policy design and co-regulation initiatives; and, legal supports and policy workshops to build stronger regulatory infrastructure and capacities for resistance within communities. **Journalists and creatives** can support by investigating the impacts of AI on the community, giving voice to the experiences and stories of the community, and building public awareness of the realities, possibilities, and potential futures of community-led AI governance. This work is particularly crucial for helping to build the cultural power and counternarratives needed to oppose dominant futurological narratives of AI-fueled utopias, economic windfalls, and technocratic regulators staving off the existential risk of unaccountable "rogue AI" systems. Unaccountable AI systems already exist in the present moment; their developers and operators are already self-regulating rogues who pose an existential risk to marginalized communities. Counternarratives such as this are essential for subverting the discursive frames and regulatory pretexts of dominant actors.

At its core, a transfeminist approach to AI governance is rooted in the lived experiences of trans people, and the histories of domination, marginalization, and oppression that trans people–especially Black trans people and trans people of color–have been forced to endure. This dissertation has emphasized the application of transfeminist ethical principles to AI governance research and practice over more direct engagement with the lived experience of being impacted by an AI system as a trans person. However, a recent report by a group of researchers at the University of Virginia shows that AI and other data-intensive technologies are indeed causing trans communities around the world to experience numerous physical, psychological, social, political, and economic harms (Reia, Leach, & Li, 2025). Against a backdrop of rising anti-trans sentiment and anti-trans legislation around the world, their report underscores the urgent need to develop community-centric systems of AI, data, and digital governance that protect trans people and empower us to prosper in online and offline environments. As Susan Stryker (2017) shows in her pivotal work *Transgender History: The Roots of Today's Revolution*, the history of trans liberation movements provides ample evidence that trans communities cannot fully rely upon or entrust the governance systems of the state to protect us from harm. Not only are state-led systems of governance frequently developed without meaningful inclusion of trans and other marginalized communities, but in far too many instances, *state power has been actively exercised with the intent of causing harm to us*. It is only through arduous acts of community organizing, direct action, and bottom-up practices of governance involving fluid, decentralized networks of collaborators–both inside and outside the halls of institutional power–that some political victories have been won for some trans communities in some parts of the world.

Future AI governance research, practices, and systems have much to learn from trans ways of thinking, feeling, and doing governance. It is not sufficient for us to do nothing more than wait for alliances of state and industry power to reach accords on legal frameworks for AI, even if we are granted a seat at the table of their discussions. It is not sufficient for us to hope that those laws will be reliably, effectively, and justly enforced to protect trans lives and to ensure the flourishing of trans communities and other marginalized communities. A transfeminist future for AI governance begins within our communities and outside of the state. As Stryker puts it in concluding her account of the history of trans liberation movements (2017, p. 236): "We can do more than cross our fingers and hope for the best if we ourselves work together to bend our little corner of the universe."

References

References for Chapter 1 (*Introduction*)

A+ Alliance (2021). Feminist AI. https://feministai.pubpub.org/

Abercrombie, G., Benbouzid, D., Giudici, P., Golpayegani, D., Henandez, J., Noro, P., Pandit, H. . . . & Waltersdorfer, L. (2024). A collaborative, human-centred taxonomy of AI, algorithmic, and automation harms. *arXiv:2407.01294v1*. https://arxiv.org/pdf/2407.01294

Adarlo, S. (2023, September 26). Critics furious Microsoft is training AI by sucking up water during drought. *Futurism*. https://futurism.com/critics-microsoft-water-train-ai-drought

AIAAIC (2025). AI, algorithmic and automation incidents and issues. *AIAAIC*. https://www.aiaaic.org/aiaaic-repository/ai-algorithmic-and-automation-incidents

AI Incident Database (2025). CSETv0. https://incidentdatabase.ai/taxonomy/csetv0/

AI Incident Database (2025). Discover incidents. *AI Incident Database*. https://incidentdatabase.ai/apps/discover/?is_incident_report=true

Akaka, M. A., Koskela-Huotari, K., & Vargo, S. L. (2019). Further advancing service science with service-dominant logic: Service ecosystems, institutions, and their implications for innovation. In P. P. Maglio, C. A. Kieliszewski, J. C. Spohrer, K. Lyons, L. Patricio, & Y. Sawatani (Eds.), *Handbook of Service Science*, *Volume II* (641-659) Springer.

Akaka, M. A. & Parry, G. (2019). Value-in-context: An exploration of the context of value and the value of context. In P. P. Maglio, C. A. Kieliszewski, J. C. Spohrer, K. Lyons, L. Patricio, & Y. Sawatani (Eds.), *Handbook of Service Science*, *Volume II* (457-477). Springer.

Alter, S. (2008). Service system fundamentals: Work system, value chain, and life cycle. *IBM Systems Journal*, *47*(1), 71–85.

Barad, K. (2015). TransMaterialities: Trans*/matter/realities and queer political imaginings. *GLQ: A Journal of Lesbian and Gay Studies*, *21*(2-3), 387-422.

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? 🦜 *FAccT '21: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (610-623). https://dl.acm.org/doi/10.1145/3442188.3445922

Benjamin, R. (2019). *Race after technology: Abolitionist tools for the New Jim Code*. Polity Press.

Birhane, A. (2021). Algorithmic injustice: A relational ethics approach. *Patterns*, *2*(2), 100205.

Birkstedt, T., Minkkinen, M., Tandon, A., & Mäntymäki, M. (2023). AI governance: Themes, knowledge gaps and future agendas. *Internet Research*, *33*(7), 133-167.

Camilleri, M. A. (2023). Artificial intelligence governance: Ethical considerations and implications for social responsibility. *Expert Systems*, 2023;e13406.

Cihon, P., Schuett, J., & Baum, S. D. (2021). Corporate governance of artificial intelligence in the public interest. *Information*, *12*, Article 275.

Collins, P. H. (2000). *Black feminist thought: Knowledge, consciousness, and the politics of empowerment*. (2nd ed.). Routledge.

Common, M. & Stagl, S. (2012). *Ecological economics: An introduction*. Cambridge University Press.

Costanza-Chock, S. (2018). Design justice, A.I., and escape from the matrix of domination. *Journal of Design and Science*. https://jods.mitpress.mit.edu/pub/costanza-chock/release/4

Costanza-Chock, S. (2020). *Design justice: Community-led practices to build the worlds we need*. The MIT Press.

Crenshaw, K. (1991). Mapping the margins: Intersectionality, identity politics, and violence against women of color. *Stanford Law Review*, *43*(6), 1241-1299.

Davidson, M. (2007). Seeking refuge under the umbrella: Inclusion, exclusion, and organizing within the category *Transgender*. *Sexuality Research & Social Policy*, *4*(4), 60-80.

de la Bellacasa, M. P. (2017). *Matters of care: Speculative ethics in more than human worlds*. University of Minnesota Press.

DeVrio, A., Eslami, M., & Holstein, K. (2024). Building, shifting, & employing power: A taxonomy of responses from below to algorithmic harm. *FAccT '24: Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency* (1093-1106). https://dl.acm.org/doi/10.1145/3630106.3658958

Diberardino, N., Baleshta, C., & Stark, L. (2024). Algorithmic harms and algorithmic wrongs. *FAccT '24: Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency* (1725-1732). https://dl.acm.org/doi/10.1145/3630106.3659001

Djeffal, C., Siewert, M. B., & Wurster, S. (2022). Role of the state and responsibility in governing artificial intelligence: A comparative analysis of AI strategies. *Journal of European Public Policy*, *29*(11), 1799-1821.

Dotson, K. (2014). Conceptualizing epistemic oppression. *Social Epistemology*, *28*(2), 115-138.

Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C. Crick, T. . . . & Williams, M. D. (2021). Artificial intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management*, *57*, April 2021, 101994.

D'Ignazio, C. & Klein, L. F. (2020). *Data feminism*. The MIT Press.

Erdélyi, O. J. & Goldsmith, J. (2022). Regulating artificial intelligence: Proposal for a global solution. *Government Information Quarterly*, *39*(4), October 2022, 101748.

Frost, R. & Lyons, K. (2017). Service systems analysis methods and components: A systematic literature review. *Service Science*, *9*(3), 219-234.

Frost, R. B., Cheng, M., & Lyons, K. (2019). A multilayer framework for service system analysis. In P. P. Maglio, C. A. Kieliszewski, J. C. Spohrer, K. Lyons, L. Patricio, & Y. Sawatani (Eds.) *Handbook of Service Science, Volume II* (285-306). Springer.

Galpin, C., Gwenffrewi, G., & Stokoe, A. (2023). Transfeminist perspectives: Beyond cisnormative understandings of the digital public sphere. *European Journal of Women's Studies*, *30*(4), 502-515.

GPAI (2021). Climate change and AI: Recommendations for government action. https://gpai.ai/projects/responsible-ai/environment/climate-change-and-ai.pdf

Gray, J. E. & Witt, A. (2021). A feminist data ethics of care framework for machine learning: The what, why, who and how. *First Monday*, *26*(12), December 2021. https://firstmonday.org/ojs/index.php/fm/article/view/11833/10528

Hancox-Li, L. & Kumar, E. I. (2021). Epistemic values in feature importance methods: Lessons from feminist epistemology. *FAccT '21: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (817-826). https://dl.acm.org/doi/10.1145/3442188.3445943

Haraway, D. (1988). Situated knowledges: The science question in feminism and the privilege of partial perspective. *Feminist Studies*, *14*(3), 575-599.

Harding, S. (1992). Rethinking standpoint epistemology: What is "strong objectivity?" *The Centennial Review*, *36*(3), 437-470.

Harding, S. (1995). "Strong objectivity": A response to the new objectivity question. *Synthese*, *104*(3), 331-349.

Hernández, A. D., Krishna, S., Perini, A. M., Katell, M., Bennett, SJ, Borda, A., Hashem, Y. . . . & Leslie, D. (2024). Mapping the individual, social, and biospheric impacts of foundation models. *FAccT '24: Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency* (776-796). https://dl.acm.org/doi/10.1145/3630106.3658939

Holmes, A. G. D. (2020). Researcher positionality – A consideration of its influence and place in qualitative research – A new researcher guide. *Shanlax International Journal of Education*, *8*(4), 1-10.

Jaakkola, E., Aarikka-Stenroos, L., & Ritala, P. (2019). Institutionalization process of service innovation: Overcoming competing institutional logics in service ecosystems. In P. P. Maglio, C. A. Kieliszewski, J. C. Spohrer, K. Lyons, L. Patricio, & Y. Sawatani (Eds.), *Handbook of Service Science, Volume II* (497-516). Springer.

Katzan, H. (2009). Principles of service systems: An ontological approach. *Journal of Service Science*, *2*(2), 35-52.

Keller, J. & Kittay, E. F. (2017). Feminist ethics of care. In A. Garry, S. J. Khader, & A. Stone (Eds.) *The Routledge Companion to Feminist Philosophy* (540-555). Routledge.

Keyes, O. (2018). The misgendering machines: Trans/HCI implications of automatic gender recognition. *Proceedings of the ACM on Human-Computer Interaction*, *2*, Issue CSCW, Article 88. https://dl.acm.org/doi/10.1145/3274357

Kinder, T., Stenvall, J., Koskimies, E., Webb, H., & Janenova, S. (2023). Local public services and the ethical deployment of artificial intelligence. *Government Information Quarterly*, *40*(4), October 2023, 101865.

Kong, Y. (2022). Are "intersectionally fair" AI algorithms really fair to women of color? A philosophical analysis. *FAccT '22: Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency* (485-494). https://dl.acm.org/doi/10.1145/3531146.3533114

Kuziemski, M. & Misuraca, G. (2020). AI governance in the public sector: Three tales from the frontiers of automated decision-making in democratic settings. *Telecommunications Policy*, *44*(6), July 2020, 101976.

Latour, B. (2005). Reassembling the social: An introduction to actor-network-theory. *Oxford University Press*.

Lehuedé, S. (2024). An elemental ethics for artificial intelligence: Water as resistance within AI's value chain. *AI & Society*. https://link.springer.com/article/10.1007/s00146-024-01922-2

Li, P., Yang, J., Islam, M. A., & Ren, S. (2023). Making AI less "thirsty": Uncovering and addressing the secret water footprint of AI models. *arXiv:2304.03271v3*. https://arxiv.org/pdf/2304.03271

Luccioni, A. S., Jernite, Y., & Strubell, E. (2024). Power hungry processing: ⚡ Watts ⚡ driving the cost of AI deployment? *FAccT '24: Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency* (85-99). https://dl.acm.org/doi/10.1145/3630106.3658542

Lyons, K. & Tracy, S. (2013). Characterizing organizations as service systems. *Human Factors and Ergonomics in Manufacturing & Service Industries*, *23*(1), 19-27.

Madan, R. & Ashok, M. (2023). AI adoption and diffusion in public administration: A systematic literature review and future research agenda. *Government Information Quarterly*, *40*(1), January 2023, 101774.

Maglio, P. P., Vargo, S. L., Caswell, N., & Spohrer, J. (2009). The service system is the basic abstraction of service science. *Information Systems and e-Business Management*, *7*, 395-406.

Maglio, P. P. & Spohrer, J. (2007). Fundamentals of service science. *Journal of the Academy of Marketing Science*, *36*, 18-20

Malatino, H. (2020). *Trans care*. University of Minnesota Press.

Malatino, H. (2022). *Side affects: On being trans and feeling bad*. University of Minnesota Press.

Mäntymäki, M., Minkkinen, M., Birkstedt, T., & Viljanen, M. (2022). Defining organizational AI governance. *AI and Ethics*, *2*, 603-609.

Marčetić, H. & Nolin, J. (2022). Feminist data studies and the emergence of a new data feminist knowledge domain. *First Monday*, *27*(7), July 2022. https://firstmonday.org/ojs/index.php/fm/article/download/12295/10681

Marvin, A. (2019). Groundwork for transfeminist care ethics: Sara Ruddick, trans children, and solidarity in dependency. *Hypatia*, *34*(1), 101-120.

Mattern, S. (2018). Maintenance and care. *Places Journal*. November 2018. https://placesjournal.org/article/maintenance-and-care

McFadden, C., Crowley-Henry, M., Rumens, N., Rocco, T. S., & Collins, J. C. (2024). Doing transgender: Gender minorities in the organization. *Gender, Work & Organization*, *31*(5), 1754-1765.

McQuillan, D. (2022). *Resisting AI: An anti-fascist approach to artificial intelligence*. Bristol University Press.

Moriggi, A., Soini, K., Franklin, A., & Roep, D. (2020). A care-based approach to transformative change: Ethically-informed practices, relational response-ability & emotional awareness. *Ethics, Policy, & Environment*, *23*(3), 281-298.

Ng, I. C. L., Maglio, P., Spohrer, J., & Wakenshaw, S. (2018). The study of service: From systems to ecosystems to ecology. In S. Vargo & R. Lusch (Eds.), *The SAGE Handbook of Service Dominant Logic* (230-240). Sage.

Nitzberg, M. & Zysman, J. (2022). Algorithms, data, and platforms: The diverse challenges of governing AI. *Journal of European Public Policy*, *29*(11), 1753-1778.

Nownes, A. J. (2019). *Organizing for transgender rights: Collective action, group development, and the rise of a new social movement*. State University of New York Press.

OECD (2022). Measuring the environmental impacts of artificial intelligence compute and applications: The AI footprint. *OECD Digital Economy Papers*, No. 341. https://doi.org/10.1787/7babf571-en

OECD (2024). OECD AI Principles overview. *OECD.AI Policy Observatory*. https://oecd.ai/en/ai-principles

Radu, R. (2021). Steering the governance of artificial intelligence: National strategies in perspective. *Policy and Society*, *40*(2), 178-193.

Ren, S. & Wierman, A. (2024, July 15). The uneven distribution of AI's environmental impacts. *Harvard Business Review*. https://hbr.org/2024/07/the-uneven-distribution-of-ais-environmental-impacts

Ricaurte, P. (2022). Ethics for the majority world: AI and the question of violence at scale. *Media, Culture & Society*, *44*(4), 726-745.

Roberts, H., Hine, E., Taddeo, M., & Floridi, L. (2023). Global AI governance: Barriers and pathways forward. *SSRN*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4588040

Schelenz, L. (2022). Artificial intelligence between oppression and resistance: Black feminist perspectives on emerging technologies. In A. Hanemaayer (ed.), *Artificial Intelligence and Its Discontents* (225-249). Springer.

Scheuerman, M. K., Pape, M., & Hanna, A. (2021). Auto-essentialization: Gender in automated facial analysis as extended colonial project. *Big Data & Society*, *8*(2), July 2021.

Scheuerman, M. K., Paul, J. M., & Brubaker, J. R. (2019). How computers see gender: An evaluation of gender classification in commercial facial analysis and image labeling services. *Proceedings of the ACM on Human-Computer Interaction, 3,* Issue CSCW, Article 144. https://dl.acm.org/doi/10.1145/3359246

Shade, L. R. (2023). From media reform to data justice: Situating women's rights as human rights. In M. Gallagher & A. V. Montiel (Eds.), *The Handbook of Gender, Communication, and Women's Human Rights* (71-87). Wiley Blackwell.

Shelby, R., Rismani, S., Henne, K., Moon, A., Rostamzadeh, N., Nicholas, P., Yilla-Akbari, N. . . . & Virk, G. (2023). Sociotechnical harms of algorithmic systems: Scoping a taxonomy for harm reduction. *AIES '23: AAAI/ACM Conference on AI, Ethics, and Society* (723-741). https://dl.acm.org/doi/10.1145/3600211.3604673

Siltaloppi, J. & Wieland, H. (2018). Institutional change in service ecosystems. In S. Vargo & R. Lusch (Eds.), *The SAGE Handbook of Service Dominant Logic* (299-315). Sage.

Solaiman, I., Talat, Z., Agnew, W., Ahmad, L., Baker, D., Blodgett, S. L., Chen, C. . . . & Subramonian, A. (2024). Evaluating the social impact of generative AI systems in systems and society. Forthcoming in Hacker, Engel, Hammer, & Mittelstadt (Eds.), *Oxford Handbook on the Foundations and Regulation of Generative AI*. https://arxiv.org/pdf/2306.05949

Spade, D. (2015). *Normal life: Administrative violence, critical trans politics, & the limits of law*. (2nd ed.). Duke University Press.

Spash, C. L. (2012). New foundations for ecological economics. *Ecological Economics*, *77*, 36-47.

Spohrer, J. (2015). Thinking about value. [Slideshow presented at the Royal College of Art]. https://www.slideshare.net/spohrer/2030-inspire-students-to-build-it-better-20150113-v3

Spohrer, J. & Kwan, S. K. (2010). Service science, management, engineering, and design (SSMED): An emerging discipline – outline & references. In Information Resources Management Association (Ed.), *Electronic Services: Concepts, Methodologies, Tools and Applications* (1970-2002). IGI Global.

Stahl, B. C., Antoniou, J., Ryan, M., Macnish, K. & Jiya, T. (2022). Organisational responses to the ethical issues of artificial intelligence. *AI & Society*, *37*, 23-37.

Suchman, L. (2007). *Human-machine reconfigurations: Plans and situated actions*. (2nd ed.). Cambridge University Press.

Tallberg, J., Erman, E., Furendal, M., Geith, J., Klamberg, & Lundgren, M. (2023). The global governance of artificial intelligence: Next steps for empirical and normative research. *International Studies Review*, *25*(3), viad040.

Tessono, C. (2024). AI governance needs a climate change strategy. *Centre for International Governance Innovation*. https://www.cigionline.org/static/documents/DPH-paper-Tessono_p7QWmGF.pdf

Toupin, S. (2024). Shaping feminist artificial intelligence. *New Media & Society*, *26*(1), 580-595.

Turri, V. & Dzombak, R. (2023). Why we need to know more: Exploring the state of AI incident documentation practices. *AIES '23: Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society* (576-583). https://dl.acm.org/doi/10.1145/3600211.3604700

Ungless, E. L., Ross, B., & Lauscher, A. (2023). Stereotypes and smut: The (mis)representation of non-cisgender identities by text-to-image models. *Findings of the Association for Computational Linguistics: ACL 2023* (7919-7942). https://aclanthology.org/2023.findings-acl.502.pdf

van der Drift, M. & Raha, N. (2020). Radical transfeminism: Trans as anti-static ethics escaping neoliberal encapsulation. In J. Cooke (Ed.), *The New Feminist Literary Studies* (13-24). Cambridge University Press.

Vargo, S. L. & Lusch, R. F. (2004). Evolving to a new dominant logic for marketing. *Journal of Marketing*, *68*(1), 1-17.

Vargo, S. L. & Lusch, R. F. (2017). Service-dominant logic 2025. *IJRM: International Journal of Research in Marketing*, *34*, 46-67.

Vargo, S. L. & Lush, R. F. (2016). Institutions and axioms: An extension and update of service-dominant logic. *Journal of the Academy of Marketing Science*, *44*, 5-23.

Varon, J. & Peña, P. (2021). Building a feminist toolkit to question A.I. systems. *Notmy.ai*. https://notmy.ai/news/algorithmic-emancipation-building-a-feminist-toolkit-to-question-a-i-systems/

Verloo, M. & van der Vleuten, A. (2020). Trans* politics: Current challenges and contestations regarding bodies, recognition, and trans* organising. *Politics and Governance*, *8*(3), 223-230.

Wan, S. & Sieber, R. E. (2023). Artificial intelligence (AI) adoption in Canadian municipalities: In-house development versus outsourcing. *The 18th International Conference on Computational Urban Planning and Urban Management*. https://osf.io/fbxgm

Weidinger, L., Mellor, J., Rauh, M., Griffin, C., Uesato, J., Huang, P.-S., Cheng, M. . . . & Gabriel, I. (2021). Ethical and social risks of harm from language models. https://arxiv.org/pdf/2112.04359

Wellner, G. & Rothman, T. (2020). Feminist AI: Can we expect our AI systems to become feminist? *Philosophy & Technology*, *33*, 191-205,

Widder, D. G. (2024). Epistemic power in AI ethics labor: Legitimizing localized complaints. *FAccT '24*: *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency* (1295-1304). https://dl.acm.org/doi/10.1145/3630106.3658973

Wilson, C. (2022). Public engagement and AI: A values analysis of national strategies. *Government Information Quarterly*, *39*(1), 101652.

Wirtz, B. W., Weyerer, J. C., & Kehl, I. (2022). Governance of artificial intelligence: A risk and guideline-based integrative framework. *Government Information Quarterly*, *39*(4), October 2022, 101685.

Zuiderwijk, A., Chen, Y.-C., & Salem, F. (2021). Implications of the use of artificial intelligence in public governance: A systematic literature review and a research agenda. *Government Information Quarterly*, *38*(3), July 2021, 101577.

References for Chapter 2 (The Impacts of AI: A Theory and Analysis of AI Value Chains)

ACLU (2023, June 7). The fight to stop face recognition technology. Retrieved from https://www.aclu.org/news/topic/stopping-face-recognition-surveillance

Ada Lovelace Institute (2021). Regulate to innovate. Retrieved from https://www.adalovelaceinstitute.org/wp-content/uploads/2021/12/Regulate-to-innovate-Ada-report.pdf

Ahmed, N., Wahed, M., & Thompson, N. C. (2023). The growing influence of industry in AI research. *Science*, *375*(6635), 884-886.

AIAAIC (2023). AI, algorithmic and automation incidents and controversies. Retrieved from https://www.aiaaic.org/aiaaic-repository/ai-and-algorithmic-incidents-and-controversies

AI Incident Database (2023). Taxonomy: CSET. Retrieved from https://incidentdatabase.ai/taxonomy/cset/

Allee, V. (2008). Value network analysis and value conversion of tangible and intangible assets. *Journal of Intellectual Capital*, *9*(1), 5-24.

Alter, S. (2008). Service system fundamentals: Work system, value chain, and life cycle. *IBM Systems Journal*, *47*(1), 71-85.

Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016, May 23). Machine bias. *ProPublica*. Retrieved from https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

Appen (2021, July 1). Responsible AI across the value chain: Ethical approaches to AI from data to deployment–and beyond. Retrieved from https://appen.com/blog/responsible-ai-across-the-value-chain-ethical-approaches-to-ai-from-data-to-deployment-and-beyond/

Attard-Frost, B., De los Ríos, A., & Walters, D. R. (2023). The ethics of AI business practices: A review of 47 AI ethics guidelines. *AI and Ethics*, *3*, 389-406.

Aula, V. & Bowles, J. (2023). Stepping back from data and AI for Good - Current trends and ways forward. *Big Data & Society*, *10*(1), January-June.

Bales, R. A. & Stone, K. V. W. (2020). The invisible web at work: Artificial intelligence and electronic surveillance in the workplace. *Berkeley Journal of Employment and Labor Law*, *41*(1). UCLA School of Law, Public Law Research Paper No. 19-18. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3410655

Bartneck, C., Lütge, C., Wagner, A., & Welsh, S. (2020). Responsibility and liability in the case of AI systems. In Bartneck, Lütge, Wagner, & Welsh (Eds.) *An Introduction to Ethics in Robotics and AI* (39-44). Springer.

Basole, R. C. (2019). On the evolution of service ecosystems: A study of the emerging API economy. In Maglio, Kieliszewski, Spohrer, Lyons, Patricio, & Sawatani (Eds.). *Handbook of Service Science, Volume II* (479-495) Springer.

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? 🦜. *FAccT '21*. Retrieved from https://dl.acm.org/doi/pdf/10.1145/3442188.3445922

Birhane, A., Isaac, W., Prabhakaran, V., Díaz, M. J., Elish, M. C., Gabriel, I., & Mohamed, S. (2022a). Power to the people? Opportunities and challenges for participatory AI. *EAAMO '22: Equity and Access in Algorithms, Mechanisms, and Optimization*, 6, October 6-9 2022. Retrieved from https://dl.acm.org/doi/pdf/10.1145/3551624.3555290

Birhane, A., Ruane, E., Laurent, T., Brown, M. S., Flowers, J., Ventresque, A., & Dancy, C. L. (2022b). The forgotten margins of AI ethics. *Conference on Fairness, Accountability,* *and Transparency.* June 2022. 948-958. Retrieved from https://dl.acm.org/doi/10.1145/3531146.3533157

Birhane, A. (2021). Algorithmic injustice: A relational ethics approach. *Patterns*, *2*(2).

Broderick, R. (2023, May 31). AI can't replace humans yet – but if the WGA writers don't win, it might not matter. *Polygon*. Retrieved from https://www.polygon.com/23742770/ai-writers-strike-chat-gpt-explained

Brown, I. (2023, June 29). Expert explainer: Allocating accountability in AI supply chains. *Ada Lovelace Institute*. Retrieved from https://www.adalovelaceinstitute.org/resource/ai-supply-chains/

Brundage, M., Avin, S., Clark, J., Toner, J., Eckersley, P., Garfinkel, B., . . . & Amodei, D. (2018). The malicious use of artificial intelligence: Forecasting, prevention, and mitigation. Retrieved from https://maliciousaireport.com/

Buhman, C., Kekre, S., & Singhal, J. (2005). Interdisciplinary and interorganizational research: Establishing the science of enterprise networks. *Production and Operations Management*, *14*(4), 493-513.

Burr, C. & Leslie, D. (2023). Ethical assurance: A practical approach to the responsible design, development, and deployment of data-driven technologies. *AI and Ethics*, *3*, 73-98.

Butollo, F., Gereffi, G., Yang, C., & Krzywdzinski, M. (2022). Digital transformation and value chains: Introduction. *Global Networks: A Journal of Transnational Affairs*, *22*, 585-594.

Butollo, F. & Schneidemesser, L. (2022). Who runs the show in digitalized manufacturing? Data, digital platforms and the restructuring of global value chains. *Global Networks: A Journal of Transnational Affairs*, *22*, 595-614.

Chan-Olmsted, S. M. (2019). A review of artificial intelligence adoptions in the media industry. *International Journal on Media Management*, *21*(3-4), 193-215.

Chen, Y.-T. & Chiu, M.-C. (2015). A case-based method for service-oriented value chain and sustainable network design. *Advanced Engineering Informatics*, *29*(3), 269-294.

Cobbe, J., Veale, M., & Singh, J. (2023). Understanding accountability in algorithmic supply chains. *FAccT '23: Conference on Fairness, Accountability, and Transparency*. 1186-1197. Retrieved from https://dl.acm.org/doi/10.1145/3593013.3594073

Coe, N. M., Dicken, P., & Hess, M. (2008). Global production networks: Realizing the potential. *Journal of Economic Geography*, *8*(3), 271-295.

Coe, N. M. & Yeung, H. W. (2019). Global production networks: Mapping recent conceptual developments. *Journal of Economic Geography*, *19*(4), 775-801.

Crain, M. (2018). "The limits of transparency: Data brokers and commodification." *New Media & Society*, *20*(1), 88-104.

Crawford, K., & Joler, V. (2018). Anatomy of an AI system: The Amazon Echo as an anatomical map of human labor, data and planetary resources. Retrieved from http://www.anatomyof.ai

Crawford, K. (2021). *Atlas of AI: Power, politics, and the planetary costs of artificial intelligence*. New Haven, CT: Yale University Press.

Dastin, J. (2018, October 10). Amazon scraps secret AI recruiting tool that showed bias against women. *Reuters*. Retrieved from https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G

De Vynck, G. (2023, June 28). ChatGPT maker OpenAI faces a lawsuit over how it used people's data. *The Washington Post*. Retrieved from https://www.washingtonpost.com/technology/2023/06/28/openai-chatgpt-lawsuit-class-action/

Dyer, J. H. (2000). *Collaborative advantage: Winning through extended enterprise supplier networks*. Oxford University Press.

Dyer-Witheford, N., Kjøsen, A. M., & Steinhoff, J. (2019). *Inhuman power: Artificial intelligence and the future of capitalism*. Pluto Press.

Edvardsson, B., Skålén, P., & Tronvoll, B. (2015). Service systems as a foundation for resource integration and value co-creation. *Review of Marketing Research*, *9*, 79-126.

Eitel-Porter, R. (2021). Beyond the promise: Implementing ethical AI. *AI and Ethics*, *1*, 73-80.

Engler, A. C. & Renda, A. (2022). Reconciling the AI value chain with the EU's Artificial Intelligence Act. Centre for European Policy Studies. Retrieved from https://www.ceps.eu/ceps-publications/reconciling-the-ai-value-chain-with-the-eus-artificial-intelligence-act/

Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press.

European Commision (2018). Artificial intelligence for Europe. COM(2018) 237 final. Retrieved from https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52018DC0237&from=EN

European Commission (2021). Artificial Intelligence Act. COM(2021) 206 final. Retrieved from https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-laying-down-harmonised-rules-artificial-intelligence

European Commission (2022). AI Liability Directive. Retrieved from https://commission.europa.eu/system/files/2022-09/1_1_197605_prop_dir_ai_en.pdf

European Parliament (2020 June). The impact of the General Data Protection Regulation (GDPR) on artificial intelligence. European Parliamentary Research Service, Scientific Foresight Unit (STOA). Retrieved from https://www.europarl.europa.eu/RegData/etudes/STUD/2020/641530/EPRS_STU(2020)641530_EN.pdf

European Parliament (2023, June 14). Texts adopted - Artificial Intelligence Act - Wednesday, 14 June 2023 - Strasbourg. COM(2021)0206 – C9-0146/2021 – 2021/0106(COD). Retrieved from https://www.europarl.europa.eu/doceo/document/TA-9-2023-0236_EN.html

Fairwork (2023). AI principles. Retrieved from https://fair.work/en/fw/principles/ai/

Feller, A., Shunk, D., & Callarman, T. (2006). Value chains versus supply chains. *BP Trends*, March 2006. Retrieved from https://www.bptrends.com/publicationfiles/03-06-ART-ValueChains-SupplyChains-Feller.pdf

Ferrandis, C. M. (2022, October 21). Responsible AI licenses: A practical tool for implementing the OECD Principles for Trustworthy AI. OECD.AI Policy Observatory, Blog. Retrieved from https://oecd.ai/en/wonk/rails-licenses-trustworthy-ai

Ferrari, F. (2023). Neural production networks: AI's infrastructural geographies. *Environment and Planning F*, *2*(4), 459-476.

Fife, S. (2022, July 18). Appetite for AI: How data is transforming the food industry's value chain. Google Cloud Blog. Retrieved from https://cloud.google.com/blog/transform/agtech-how-ai-transforms-food-industry-value-chain

Frost, R. B., Cheng, M., & Lyons, K. (2019). A multilayer framework for service system analysis. In Maglio, Kieliszewski, Spohrer, Lyons, Patricio, & Sawatani (Eds.) *Handbook of Service Science, Volume II* (285-306) Springer.

Gans-Combe, C. (2022). Automated justice: Issues, benefits and risks in the use of artificial intelligence and its algorithms in access to justice and law enforcement. In O'Mathúna & Iphofen (Eds.) *Ethics, Integrity and Policymaking: The Value of the Case Study* (175-194). Springer.

Gebru, T. & Torres, É. P. [Nicolas Papernot] (2023, February 16). SaTML 2023 - Timnit Gebru - Eugenics and the promise of utopia through AGI [Video]. *YouTube*. Retrieved from https://www.youtube.com/watch?v=P7XT4TWLzJw

Gereffi, G., Humphrey, J., & Sturgeon, T. (2006). The governance of global value chains. *Review of International Political Economy, 12*(1), 78-104.

GitHub Copilot Litigation (2023). GitHub Copilot litigation. Retrieved from https://githubcopilotlitigation.com/

Government of Canada (2023a). Voluntary code of conduct on the responsible development and management of advanced generative AI systems. Retrieved from https://ised-isde.canada.ca/site/ised/en/voluntary-code-conduct-responsible-development-and-management-advanced-generative-ai-systems

Government of Canada (2023b). Guide on the use of generative AI. Retrieved from https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/responsible-use-ai/guide-use-generative-ai.html

GPAI (2021). Climate change and AI: Recommendations for government action. Global Partnership on AI, Climate Change AI, and Centre for AI & Climate. Retrieved from https://www.gpai.ai/projects/climate-change-and-ai.pdf

GPAI (2022). AI for fair work report. Global Partnership on AI. Retrieved from https://fair.work/wp-content/uploads/sites/17/2022/12/AI-for-fair-work-report-edited.pdf

Greene, D., Hoffman, A. L., & Stark, L. (2019). Better, nicer, clearer, fairer: A critical assessment of the movement for ethical artificial intelligence and machine learning. *Hawaii International Conference on System Sciences*, 2122-2131. Retrieved from https://scholarspace.manoa.hawaii.edu/server/api/core/bitstreams/849782a6-06bf-4ce8-9144-a93de4455d1c/content

Greshake, K., Abdelnabi, S., Mishra, S., Endres, C., Holz, T., & Fritz, M. (2023). Not what you've signed up for: Compromising real-world LLM-integrated applications with indirect prompt injection. arXiv:2302.12173v2. Retrieved from https://arxiv.org/pdf/2302.12173.pdf

Grote, T. & Berens, P. (2022). On the ethics of algorithmic decision-making in healthcare. *Journal of Medical Ethics*, *46*, 205-211.

Gwagwa, A., Kazim, E., & Hilliard, A. (2022). The role of the African value of Ubuntu in global AI inclusion discourse: A normative ethics perspective. *Patterns*, *3*(4).

Hagendorff, R. (2020). The ethics of AI ethics: An evaluation of guidelines. *Minds and Machines*, *30*, 99-120.

Härlin, T., Rova, G. B., Singla, A., Sokolov, O., & Sukharvesky, A. (2023, April 26). Exploring opportunities in the generative AI value chain. *QuantumBlack AI by McKinsey*. Retrieved from https://www.mckinsey.com/capabilities/quantumblack/our-insights/exploring-opportunities-in-the-generative-ai-value-chain

Hatmaker, T. (2022, May 9). Clearview AI banned from selling its facial recognition software to most US companies. *TechCrunch*. Retrieved from https://techcrunch.com/2022/05/09/clearview-settlement-bipa/

Henderson, J., Dicken, P., Hess, M., Coe, N., & Yeung, H. W. (2001). Global production networks and the analysis of economic development. SPA Working Paper 49. Spatial Policy Analysis, School of Geography, University of Manchester. Retrieved from https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=75866a2aa4d4e131117f368dffa7a1b2d35d0c1f

Hickok, M. & Maslej, N. (2023). A policy primer and roadmap on AI worker surveillance and productivity scoring tools. *AI and Ethics, 3*, 673-687. Retrieved from https://doi.org/10.1007/s43681-023-00275-8

Hind, S. & Seitz, T. (2022). Cynical technical practice: From AIs to APIs. *Convergence: The International Journal of Research into New Media Technologies*. Retrieved from https://journals.sagepub.com/doi/10.1177/13548565221133248

Hoijtink, M. & Hardeveld, A. P. (2022). Machine learning and the platformization of the military: A study of Google's machine learning platform TensorFlow. *International Political Sociology*, *16*(2), June 2022, olab036.

Howson, K., Johnston, H., Cole, M., Ferrari, F., Ustek-Spilda, F., & Graham, M. (2022a). Unpaid labour and territorial extraction in digital value networks. *Global Networks: A Journal of Transnational Affairs*, 2022, 1-23.

Howson, K. Ferrari, F., Ustek-Spilda, F., Salem, N., Johnston, H., Katta, S., . . . & Graham, M. (2022b). Driving the digital value network: Economic geographies of global platform capitalism. *Global Networks: A Journal of Transnational Affairs*, *22*, 631-648.

Humphrey, J. & Schmitz, H. (2000). Governance and upgrading: Linking industrial cluster to global value chain research. IDS Working Paper 120. Institute of Development Studies. Retrieved from https://www.marketlinks.org/sites/default/files/media/file/2020-10/Governance%20and%20Upgrading.pdf

Irani, L. (2015). The cultural work of microwork. *New Media & Society*, *17*(5), 720-739.

ISO (2022). Standards by ISO/IEC JTC 1/SC 42: Artificial intelligence. Retrieved from https://www.iso.org/committee/6794475/x/catalogue/p/0/u/1/w/0/d/0

Kak, A. & West, S. M. (2023, April 13). General purpose AI poses serious risks, should be excluded from the EU's AI Act: Policy brief. *AI Now Institute*. Retrieved from https://ainowinstitute.org/publication/gpai-is-high-risk-should-not-be-excluded-from-eu-ai-act

Kano, L., Tsang, E. W. K., & Yeung, H. W. (2020). Global value chains: A review of the multi-disciplinary literature. *Journal of International Business Studies*, *51*, 577-622.

Kim, S., Colicchia, C., & Menachof, D. (2018). Ethical sourcing: An analysis of the literature and implications for future research. *Journal of Business Ethics*, *152*, 1033-1052.

Krönke, C. (2019). Artificial intelligence and social media. In Wischmeyer & Rademacher (Eds.), *Regulating artificial intelligence* (145-173), Springer.

Lamdan, S. (2022). *Data cartels: The companies that control and monopolize our information*. Stanford University Press.

Langenkamp, M. & Yue, D. N. (2022). How open source machine learning software shapes AI. *AAAI/ACM Conference on AI, Ethics, and Society*. July 2022, 385-398. Retrieved from https://dl.acm.org/doi/10.1145/3514094.3534167

Lauer, D. (2021). You cannot have AI ethics without ethics. *AI and Ethics*.

Lee, K., Cooper, A. F., & Grimmelmann, J. (2023, September 14). Talkin' 'bout AI generation: Copyright and the generative-AI supply chain. arXiv:2309:08133.

Lewis, J. E., Abdilla, A., Arista, N., Baker, K., Benesiinaabandan, S., Brown, M., Cheung, M. . . . Whaanga, H. (2020). *Indigenous Protocol and Artificial Intelligence Position Paper*. Indigenous Protocol and Artificial Intelligence Working Group and the Canadian Institute for Advanced Research. Retrieved from https://www.indigenous-ai.net/position-paper/

Lewis, J. E., Arista, N., Pechawis, A., & Kite, S. (2018). Making kin with the machines. *Journal of Design and Science*.

Li, P., Yang, J., Islam, M. A., & Ren, S. (2023). Making AI less 'thirsty': Uncovering and addressing the secret water footprints of AI models. arXiv:2304.03271v1. Retrieved from https://arxiv.org/pdf/2304.03271.pdf

Lim, C., Kim, K.-H., Kim, M.-J., Heo, J.-Y., Kim, K.-J., & Maglio, P. P. (2018). From data to value: A nine-factor framework for data-based value creation in information-intensive services. *International Journal of Information Management*, *39*, 121-135.

Liu, S., Chen, X., & Chen, Y. (2022). Artificial intelligence and the evolution of global value chains in the context of intelligent manufacturing. *AIEE 2022: The 3rd International Conference on Artificial Intelligence in Electronics Engineering*, January 2022, 44-48.

Luccioni, A. S. & Hernandez-Garcia, A. (2023). Counting carbon: A survey of factors influencing the emissions of machine learning. arXiv:2302.08476v1. Retrieved from https://arxiv.org/pdf/2302.08476.pdf

Lush, R. F., Vargo, S. L., & Tanniru, M.(2010). Service, value networks and learning. *Journal of the Academy of Marketing Science*, *38*(19), 19-31.

MacKinnon, E. & King, J. (2022, January 11). Regulating AI through data privacy. *Stanford University Institute for Human-Centred Artificial Intelligence.* Retrieved from https://hai.stanford.edu/news/regulating-ai-through-data-privacy

Madianou, M. (2021). Nonhuman humanitarianism: When 'AI for Good' can be harmful. *Information, Communication & Society*, *24*(6), 850-868.

Mahoney, C. W. (2020). United States defence contractors and the future of military operatioins. *Defense & Security Analysis*, *36*(2), 180-200.

Mäntymäki, M., Minkkinen, N., Birkstedt, T., & Viljanen, M. (2022). Defining organizational AI governance. *AI and Ethics, 2*, 603-609.

Masiello, B. & Slater, D. (2023, June 16). Will open source AI shift power from 'Big Tech'? It depends. *Tech Policy Press*. Retrieved from https://techpolicy.press/will-open-source-ai-shift-power-from-big-tech-it-depends/

Mhlambi, S. (2020). Ubuntu as an ethical & human rights framework for artificial intelligence governance. *Carr Center for Humans Rights Policy, Harvard Kennedy School*. Retrieved from https://carrcenter.hks.harvard.edu/files/cchr/files/ccdp_2020-009_sabelo_b.pdf

Miceli, M., Posada, J., & Yang, T. (2022). Studying up machine learning data: Why talk about bias when we mean power? *Conference on Human-Computer Interaction.* Volume 6, Issue GROUP, Article 34.

Miceli, M., Yang, T., Garcia, A. A., Posada, J., Wang, S. M., Pohl, M., & Hanna, A. (2022). Documenting data production processes: A participatory approach for data work. *Conference on Human-Computer Interaction*. Volume 6, Issue CSCW2, Article No. 510.

Miceli, M. & Posada, J. (2022). The data-production dispositif. *Conference on Human-Computer Interaction*. Volume 6, Issue CSCW2, Article 460. Retrieved from https://dl.acm.org/doi/10.1145/3492853

Minister of Innovation, Science and Industry (2023, October 3). [Letter from the Minister to the Chair of the Canadian House of Commons Standing Committee on Industry and Technology]. Retrieved from https://www.ourcommons.ca/content/Committee/441/INDU/WebDoc/WD12751351/12751351/MinisterOfInnovationScienceAndIndustry-2023-11-28-Combined-e.pdf

Mitchell, M., Wu, S., Zaldivar, A., Barnes, P., Vasserman, L., Hutchinson, B. . . . & Gebru, T. (2019). Model cards for model reporting. *Conference on Fairness, Accountability, and Transparency*. January 2019, 220-229. Retrieved from https://dl.acm.org/doi/10.1145/3287560.3287596

Mökander, J. & Axente, M. (2023). Ethics-based auditing of automated decision-making systems: Intervention points and policy implications. *AI & Society*, *38*, 153-171.

Moore, J. (2019). AI for not bad. *Frontiers in Big Data*, *2* (32), Article 32. Retrieved from https://www.frontiersin.org/articles/10.3389/fdata.2019.00032/full

Morley, J., Kinsey, L., Elhalal, A., Garcia, F., Ziosi, M., & Floridi, L. (2023). Operationalising AI ethics: Barriers, enablers and next steps. *AI & Society*, *38*, 411-423.

Mulligan, D. K. & Bamberger, K. A. (2019). Procurement as policy: Administrative process for machine learning. *Berkeley Technology Law Journal*, *34*, 751-856. Retrieved from https://papers-ssrn-com.myaccess.library.utoronto.ca/sol3/papers.cfm?abstract_id=3464203

National Institute of Standards and Technology (2023). Artificial Intelligence Risk Management Framework (AI RMF0). Retrieved from https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf

Ochigame, R. (2019, December 20). The invention of 'ethical AI': How Big Tech manipulates academia to avoid regulation. *The Intercept*. Retrieved from https://theintercept.com/2019/12/20/mit-ethical-ai-artificial-intelligence/

OECD (2019). Proposal for the OECD AI Policy Observatory. DSTI/CDEP(2019)5. Retrieved from https://one.oecd.org/document/DSTI/CDEP(2019)5/en/pdf

OECD (2023, January 27). Credo AI Responsible AI Governance Platform. OECD.AI Policy Observatory, Catalogue of AI tools & metrics. Retrieved from https://oecd.ai/en/catalogue/tools/credo-ai-responsible-ai-governance-platform

Oosthuizen, K., Botha, E., Robertson, J., & Montecchi, M. (2020). Artificial intelligence in retail: The AI-enabled value chain. *Australasian Marketing Journal*, *29*(3), 264-273.

Palinkas, L. A., Horwitz, S. M., Green, C. A., Wisdom, J. P., Duan, N., & Hoagwood, K. (2015). Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. *Administration and Policy in Mental Health and Mental Health Services Research*, *42*(5), 533-544.

Parliament of Canada (2022). Digital Charter Implementation Act, 2022. Retrieved from https://www.parl.ca/legisinfo/en/bill/44-1/c-27

Pasquale, F. (2020). *New laws of robotics: Defending human expertise in the age of AI*. The Belknap Press of Harvard University Press.

Perrigo, B. (2022, May 27). An AI company scraped billions of photos for facial recognition. Regulators can't stop it. *Time*. Retrieved from https://time.com/6182177/clearview-ai-regulators-uk/

Perrigo, B. (2023, January 18). Exclusive: OpenAI used Kenyan workers on less than $2 per hour to make ChatGPT less toxic. *Time*. Retrieved from https://time.com/6247678/openai-chatgpt-kenya-workers/

Porter, M. E. (1985). *Competitive advantage: Creating and sustaining superior performance*. The Free Press.

Raji, I. D., Smart, A., White, R. N., Mitchell, M., Gebru, T., Hutchinson, B., . . . & Barnes, P. (2020). Closing the AI accountability gap: defining an end-to-end framework for internal algorithmic auditing. *Conference on Fairness, Accountability, and Transparency*. January 2020. 33-44. Retrieved from https://dl.acm.org/doi/10.1145/3351095.3372873

Rankin, D., Black, M., Bond, R., Wallace, J., Mulvenna, M., & Epelde, G. (2020). Reliability of supervised machine learning using synthetic data in health care: Model to preserve privacy for data sharing. *JMIR Medical Informatics*, *8*(7), e18910. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7400044/

Rességuier, A. & Rodrigues, R. (2020). AI ethics should not remain toothless! A call to bring back the teeth of ethics. *Big Data & Society*, July 2020, 1-5.

Robinson, R. S. (2014). Purposive sampling. In Michalos (Ed.), *Encyclopedia of Quality of Life and Wellbeing Research* (5243-5245). Springer.

Schneider, J., Abraham, R., Meske, C., & vom Brocke, J. (2023). Artificial intelligence governance for businesses. *Information Systems Management*, *40*(3), 229-249.

Shaw, G. & Arkan, C. (2019) The future computed: AI & manufacturing. Microsoft. Retrieved from https://news.microsoft.com/wp-content/uploads/prod/sites/393/2019/06/Microsoft_TheFutureComputed_AI_MFG_Final_Online.pdf

Shelby, R., Rismani, S., Henne, K., Moon, A., Rostamzadeh, N., Nicholas, P., Yilla-Akbari, N. . . . Virk, G. (2023). Identifying sociotechnical harms of algorithmic systems: Scoping a taxonomy for harm reduction. Retrieved from https://arxiv.org/pdf/2210.05791.pdf

Siltaloppi, J. & Vargo, S. L. (2018). Institutional change in service ecosystems. In Vargo & Lusch (Eds.), *The SAGE Handbook of Service Dominant Logic* (299-315). Sage.

Smuha, N. A. (2021). From a 'race to AI' to a 'race to AI regulation': Regulatory competition for artificial intelligence. *Law, Innovation and Technology*, *13*(1), 57-84.

Solaiman, I., Talat, Z., Agnew, W., Ahmad, L., Baker, D., Blodgett, S. L. . . . & Vassilev, A. (2023). Evaluating the social impact of generative AI systems in systems and society. arXiv:2306.05949v2. Retrieved from https://arxiv.org/pdf/2306.05949.pdf

Spash, C. L. (2012). New foundations for ecological economics. *Ecological Economics*, *77*, 36-47.

Stable Diffusion Litigation (2023). Stable Diffusion litigation. Retrieved from https://stablediffusionlitigation.com/

Stahl, B. C., Antoniou, J. Ryan, M., Macnish, K., & Jiya, T. (2022). Organisational responses to the ethical issues of artificial intelligence. *AI & Society*, *37*, 23-37.

Suwandi, I. (2019). *Value chains: The new economic imperialism*. Monthly Review Press.

Taddeo, M., McNeish, D., Blanchard, A., & Edgar, E. (2021). Ethical principles for artificial intelligence in national defence. *Philosophy & Technology*, *34*, 1707-1729.

Tiku, N. (2023, July 5). How elite schools like Stanford became fixated on the AI apocalypse. *The Washington Post*. Retrieved from https://www.washingtonpost.com/technology/2023/07/05/ai-apocalypse-college-students/

Torres, É. P. (2023, June 15). The acronym behind our wildest AI dreams and nightmares. *Truthdig*. Retrieved from https://www.truthdig.com/articles/the-acronym-behind-our-wildest-ai-dreams-and-nightmares/

Vargo, S. L. & Lusch, R. F. (2016). Institutions and axioms: An extension and update of service-dominant logic. *Journal of the Academy of Marketing Science*, *44*, 5-23.

Veale, M., Binns, R., & Edwards, L. (2018, October 15). Algorithms that remember: Model inversion attacks and data protection law. *Philosophical Transactions of the Royal Society A*, *376*(2133), 20180083. Retrieved from https://doi.org/10.1098/rsta.2018.0083

Vincent, J. (2023, February 6). Getty Images sues AI art generator Stable Diffusion in the US for copyright infringement. *The Verge*. Retrieved from https://www.theverge.com/2023/2/6/23587393/ai-art-copyright-lawsuit-getty-images-stable-diffusion

Wang, C., Chen, J., Yang, Y., Ma, X., & Liu, J. (2022). Poisoning attacks and countermeasures in intelligent networks: Status quo and prospects. *Digital Communications and Networks*, *8*(2), 225-234.

Webster, A. (2023, July 13). Actors say Hollywood studios want their AI replicas — for free, forever. *The Verge*. Retrieved from https://www.theverge.com/2023/7/13/23794224/sag-aftra-actors-strike-ai-image-rights

Weidinger, L., Mellor, J., Rauh, M., Griffin, C., Uesato, J., Huang, P.-S. . . . & Gabriel, I. (2021). Ethical and social risks of harm from language models. arXiv:2303.07529v1. Retrieved from https://arxiv.org/abs/2112.04359

West, S. M., Whittaker, M., & Crawford, K. (2019). Discriminating systems: Gender, race and power in AI. *AI Now Institute*. Retrieved from https://ainowinstitute.org/publication/discriminating-systems-gender-race-and-power-in-ai-2

Whittaker, M. (2021). The steep cost of capture. *Interactions*, *28*(2), 50-55. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4135581#

Widder, D. G. & Nafus, D. (2022). Dislocated accountabilities in the AI supply chain: Modularity and developers' notions of responsibility. arXiv:2209.09780v2. Retrieved from https://arxiv.org/pdf/2209.09780v2.pdf

Widder, D. G. & Nafus, D. (2023). Dislocated accountabilities in the 'AI supply chain': Modularity and developers' notions of responsibility. *Big Data & Society*, *10*(1), January-June 2023.

Widder, D. G., West, S., & Whittaker, M. (2023). Open (for business): Big tech, concentrated power, and the political economy of open AI. *SSRN*. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4543807

Widder, D. G., & Wong, R. Y. (2023). Thinking upstream: Ethics and policy opportunities in AI supply chains. arXiv:2303.07529v1. Retrieved from https://arxiv.org/pdf/2303.07529.pdf

Wolfewicz, A. (2022, November 16). The value chain of machine learning. Levity Blog. Retrieved from https://levity.ai/blog/machine-learning-value-chain

Woolley, S. (2018). The political economy of bots: Theory and method in the study of social automation. In Kiggins (Ed.), *The political economy of robots: Prospects for prosperity and peace in the automated 21st Century* (127-155). Palgrave Macmillan.

Zech, H. (2021). Liability for AI: Public policy considerations. *ERA Forum, 22*, 147-158.

References for Chapter 3 (AI Governance Initiatives in Canada)

Abdilla, A., Kelleher, M., Shaw, R., & Yunkaporta, T. (2021) Out of the black box: Indigenous protocols for AI. IP//AI Incubator, March 2021. Retrieved February 2, 2024 from https://static1.squarespace.com/static/5778a8e3e58c62bbf1a639ae/t/61808f1d034eda41942223a9/1635815199890/*Final+Unesco+Paper_Designed.pdf

Accenture & CIFAR (2020). Pan-Canadian AI Strategy impact assessment report. Retrieved February 2, 2024 from https://cifar.ca/wp-content/uploads/2020/11/Pan-Canadian-AI-Strategy-Impact-Assessment-Report.pdf

Adu-Daako, A. & Sieber, R. (2022). What is meaningful civic engagement. In A. Brandusescu & J. Reia (Eds.), *AI in the city: Building civic engagement and public trust*. (6-8). Centre for Interdisciplinary Research on Montreal, McGill University. Retrieved February 2, 2024 from https://www.mcgill.ca/centre-montreal/files/centre-montreal/artificial_intelligence_in_the_city.pdf

Ahmed, N., Wahed, M., & Thompson, N. C. (2023). The growing influence of industry in AI research. *Science*, *375*(6635), 884-886. Retrieved February 2, 2024 from https://www.science.org/doi/10.1126/science.ade2420

Arnstein, S. R. (1969). A ladder of citizen participation. *Journal of the American Institute of Planners*, *35*(4), 216-224.

Attard-Frost, B., De los Ríos, A., & Walters, D. R. (2022). The ethics of AI business practices: A review of 47 AI ethics guidelines. *AI and Ethics*, *3*, 389-406. Retrieved February 2, 2024 from https://link.springer.com/article/10.1007/s43681-022-00156-6

Attard-Frost, B. (2022). Comments on the 3rd review of the Directive on Automated Decision-making. Retrieved February 2, 2024 from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4238164

Auld, G., Casovan, A., Clarke, A., & Faveri, B. (2022). Governing AI through ethical standards: Learning from the experiences of other private governance initiatives. *Journal of European Public Policy*, *29*(11), 1822-1844.

Babashahiashtiani, K. (2021). The economic geography of Canada's AI ecosystem (Master's thesis, York University, Toronto). Retrieved February 2, 2024 from https://yorkspace.library.yorku.ca/xmlui/bitstream/handle/10315/39141/Babashahiashtiani_Kasra_2021_Masters.pdf

Bailey, J., Steeves, V., Burkell, J., Gosse, C., & Dunn, S. (2021). AI and technology-facilitated violence and abuse. In F. Martin-Bariteau & T. Scassa (Eds.), *Artificial intelligence and the law in Canada*. LexisNexis Canada. Retrieved February 2, 2024 from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3734663

Bakvis, H. & Skogstad, G. (2020). Canadian federalism: Performance, effectiveness, and legitimacy. In H. Bakvis & G. Skogstad (Eds.), *Canadian federalism: Performance, effectiveness, and legitimacy*. (4th ed., 14-55). University of Toronto Press.

Bao, G., Wang, X., Larsen, G. L., & Morgan, D. F. (2013). Beyond New Public Governance: A value-based global framework for performance management, governance, and leadership. *Administration & Society*, *45*(4), 443-467.

Benanav, A. (2020). *Automation and the future of work*. Verso.

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? 🦜. *FAccT '21*, March 3-10 2021. Retrieved February 2, 2024 from https://dl.acm.org/doi/pdf/10.1145/3442188.3445922

Birhane, A., Isaac, W., Prabhakaran, V., Díaz, M. J., Elish, M. C., Gabriel, I., & Mohamed, S. (2022). Power to the people? Opportunities and challenges for participatory AI. *EAAMO '22: Equity and Access in Algorithms, Mechanisms, and Optimization*, 6, October 6-9 2022. Retrieved February 2, 2024 from https://dl.acm.org/doi/pdf/10.1145/3551624.3555290

Birkstedt, T., Minkkinen, M., Tandon, A., & Mäntymäki, M. (2023). AI governance: Themes, knowledge gaps and future agendas. *Internet Research*, *33*(7), 133-167.

Bitar, O., Deshaies, B., & Hall, D. (2022). 3rd review of the Treasury Board Directive on Automated Decision-making. Retrieved February 2, 2024 from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4087546

Brandusescu, A. (2021). Artificial intelligence policy and funding in Canada: Public investments, private interests. Centre for Interdisciplinary Research on Montreal, McGill University. Retrieved February 2, 2024 from https://www.mcgill.ca/centre-montreal/files/centre-montreal/aipolicyandfunding_report_updated_mar5.pdf

Brandusescu, A. & Sieber, R. (2022). Comments on the third review of Canada's Directive on Automated Decision-making. Retrieved February 2, 2024 from https://aifortherestofus.ca/wp-content/uploads/2022/07/Comments-on-the-third-review-of-Canadas-Directive-on-ADM.pdf

Büthe, T., Djeffal, C., Lütge, C., Maasen, S., & von Ingersleben-Seip, N. (2022). Governing AI – attempting to herd cats? Introduction the special issue on the governance of artificial intelligence. *Journal of European Public Policy*, *29*(11), 1721-1752.

Cameron, D. & Simeon, R. (2002). Intergovernmental relations in Canada: The emergence of collaborative federalism. *Publius: The Journal of Federalism*, *32*(2), 49-72.

Camilleri, M. A. (2023). Artificial intelligence governance: Ethical considerations and implications for social responsibility. *Expert Systems*, 2023;e13406.

Cardoso, T. (2022, November 23). Parliamentary study hears federal FOI has become a 'dysfunctional system'. *The Globe and Mail*. Retrieved February 2, 2024 from https://www.theglobeandmail.com/canada/article-parliamentary-study-hears-federal-foi-has-become-a-dysfunctional/

Centre for Digital Rights (2022). Not fit for purpose – Canada deserves much better: Centre for Digital Rights' statement on Bill C-27. Retrieved February 2, 2024 from https://centrefordigitalrights.org/files/document/2022-11-13/257-013312.pdf

CIFAR (2020). AI futures: Policy lab facilitator guide. Retrieved February 2, 2024 from https://cifar.ca/wp-content/uploads/2020/11/ai-futures-policy-lab-facilitator-guide-en.pdf

Cihon, P., Schuett, J., & Baum, S. D. (2021). Corporate governance of artificial intelligence in the public interest. *Information*, *12*, Article 275.

Clarke, A. (2019). *Opening the Government of Canada: The federal bureaucracy in the digital age*. University of British Columbia Press.

Cole, M., Cant, C., Spilda, F. U., & Graham, M. (2022). Politics by automatic means? A critique of artificial intelligence ethics at work. *Frontiers in Artificial Intelligence*, *5*, Article 869114.

Cornwall, A. (2008). Unpacking 'participation': Models, meanings, and practices. *Community Development Journal*, *43*(3), 269-283.

Crawford, K. (2021). *Atlas of AI: Power, politics, and the planetary costs of artificial intelligence.* Yale University Press.

Daly, P. & Orct, B. (2022). Artificial intelligence accountability of public administration in Canada: Country report for the 2022 General Congress of the International Academy of Comparative Law. Ottawa Faculty of Law Working Paper No. 2022-30. Retrieved February 2, 2024 from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4266365

Deloitte (2019). Canada's AI imperative: Overcoming risks, building trust. Retrieved February 2, 2024 from https://www2.deloitte.com/content/dam/Deloitte/ca/Documents/deloitte-analytics/ca-overcoming-risks-building-trust-aoda-en.pdf

Digital Governance Standards Institute (2021). Ethical design and use of automated decisions systems. CAN/CIOSC 101:2019 (Reaffirmed 2021-10). Retrieved February 2, 2024 from https://dgc-cgn.org/standards/find-a-standard/standards-in-automated-decision-systems-ai/cisoc101/#gf_12

Djeffal, C., Siewert, M. B., & Wurster, S. (2022). Role of the state and responsibility in governing artificial intelligence: A comparative analysis of AI strategies. *Journal of European Public Policy*, *29*(11), 1799-1821.

Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C. Crick, T. . . . & Williams, M. D. (2021). Artificial intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management*, *57*, April 2021, 101994.

Dyer-Witheford, N., Kjøsen, A. M., & Steinhoff, J. (2019). *Inhuman power: Artificial intelligence and the future of capitalism*. Pluto Press.

Erdélyi, O. J. & Goldsmith, J. (2022). Regulating artificial intelligence: Proposal for a global solution. *Government Information Quarterly, 39*(4), October 2022, 101748.

European Commission (2018). Artificial intelligence for Europe. COM(2018) 237 final. Retrieved February 2, 2024 from https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52018DC0237&from=EN

European Commission (2021). Artificial Intelligence Act. COM(2021) 206 final. Retrieved February 2, 2024 from https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-laying-down-harmonised-rules-artificial-intelligence

European Commission (2023). The European AI Alliance. Retrieved February 2, 2024 from https://digital-strategy.ec.europa.eu/en/policies/european-ai-alliance

Fair Trade Federation (2023). Fair Trade Federation Code of Practice. Retrieved February 2, 2024 from https://www.fairtradefederation.org/ftf-code-of-practice

Fairwork (2023). AI principles. Retrieved February 2, 2024 from https://fair.work/en/fw/principles/ai/

Flamholtz, E. & Randle, Y. (2009). Understanding organizational change. In E. Flamholtz & Y. Randle (Eds.), *Leading strategic change: Bridging theory and practice*. (5-25). Cambridge University Press.

Frost, R B. (2020). Emerging national artificial intelligence innovation systems in Canada and China: Strategic governance and institutional evolution. *SPACES Online: Spatial Aspects Concerning Economic Structures*, *16*(1). Retrieved February 2, 2024 from https://www.spaces-online.com/include/SPACES_2020-01%20Frost.pdf

G7 2018 (2018, December 6). Final summary report: G7 Multistakeholder Conference on Artificial Intelligence. Retrieved February 2, 2024 from https://ised-isde.canada.ca/site/G7-multistakeholder-conference-artificial-intelligence/sites/default/files/attachments/EN_MSC-Final-Report.pdf

Gaon, A. & Stedman, I. (2019). A call to action: Moving forward with the governance of artificial intelligence in Canada. *Alberta Law Review*, *56*(4), 1137-1166.

Government of Canada (2021). Introduction to policy. Retrieved February 2, 2024 from https://www.canada.ca/en/heritage-information-network/services/digital-preservation/concepts-developing-policies/introduction-policy.html

Government of Canada (2024). Open Government Portal. Collection type: Algorithmic Impact Assessment. Retrieved February 2, 2024 from https://search.open.canada.ca/opendata/?collection=aia&page=1&sort=date_modified+desc

GPAI (2021). Climate change and AI: Recommendations for government action. Global Partnership on AI, Climate Change AI, and Centre for AI & Climate. Retrieved February 2, 2024 from https://www.gpai.ai/projects/climate-change-and-ai.pdf

GPAI (2022). AI for fair work report. Global Partnership on AI. Retrieved February 2, 2024 from https://fair.work/wp-content/uploads/sites/17/2022/12/AI-for-fair-work-report-edited.pdf

Gupta, A. (2021, September 18). The imperative for sustainable AI systems. *The Gradient*. Retrieved February 2, 2024 from https://thegradient.pub/sustainable-ai/

Harvey, D. (2005). *A brief history of neoliberalism*. Oxford University Press.

Hsieh, H.-F. & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, *15*(9), 1277-1288.

Ifill, E. (2022, July 4). The problems with the federal data-privacy bill will disproportionately hurt marginalized Canadians. *The Globe and Mail*. Retrieved February 2, 2024 from https://www.theglobeandmail.com/opinion/article-the-problems-with-the-federal-data-privacy-bill-will/

Information Commissioner of Canada (2002). Response to the report of the Access to Information Review Task Force: A special report to Parliament. Retrieved February 2, 2024 from https://www.oic-ci.gc.ca/sites/default/files/report-publications/IP4-1-2002E.pdf

Information Commissioner of Canada (2022). Annual report 2021-2022. Retrieved February 2, 2024 from https://www.oic-ci.gc.ca/sites/default/files/2022-06/AnnualReport-OIC-2021-2022-EN-Final.pdf

Innovation, Science and Economic Development Canada (2020). Commercialization Working Group final report February 2020. Retrieved February 2, 2024 from https://ised-isde.canada.ca/site/advisory-council-artificial-intelligence/en/commercialization-working-group/commercialization-working-group-final-report-february-2020

Innovation, Science and Economic Development Canada (2022). Evaluation of Innovation, Science and Economic Development Canada funding to CIFAR. Retrieved February 2, 2024 from https://ised-isde.canada.ca/site/audits-evaluations/en/evaluation/evaluation-innovation-science-and-economic-development-ised-canada-funding-cifar

Innovation, Science and Economic Development Canada (2023). Learning together for responsible artificial intelligence: Report of the Public Awareness Working Group.

Retrieved February 2, 2024 from https://ised-isde.canada.ca/site/advisory-council-artificial-intelligence/sites/default/files/attachments/2023/learning_together_for_responsible_artificial_intelligence_minisi_approved.pdf

Institute for the Future of Work (2023, March 28). Good Work Algorithmic Impact Assessment. Retrieved February 2, 2024 from https://www.ifow.org/publications/good-work-algorithmic-impact-assessment-an-approach-for-worker-involvement

Ipsos (2022, January 14). Canadians among least likely to believe artificial intelligence will make their lives better. Retrieved February 2, 2024 from https://www.ipsos.com/en-ca/news-polls/Canadians-Least-Likely-AI-Make-Lives-Better

Irwin, R. & White, T. H. (2019). Decolonising technological futures: A dialogical triptych between Te Haumoana White, Ruth Irwin, and Tegmark's artificial intelligence. *Futures*, *112*, 102431.

ISO (n.d.). Consumers and standards: Partnership for a better world. Retrieved February 2, 2024 from https://www.iso.org/sites/ConsumersStandards/1_standards.html

Janssen, M., Brous, P., Estevez, E., Barbosa, L. S., & Janowski, T. (2020). Data governance: Organizing data for trustworthy artificial intelligence. *Government Information Quarterly*, *37*(3), July 2020, 101493.

Karanicolas, M. (2019). To err is human, to audit divine: A critical assessment of Canada's AI Directive. *Journal of Parliamentary and Political Law*, *14*(1).

Keenan, P., Bickford, J. K., Doust, A., Tankersley, J. Johnson, C. McCaffrey, J., Dolfi, J., & Shah, G. (2013). Strategic initiative management: The PMO imperative. Boston Consulting Group. Retrieved February 2, 2024 from https://www.bcg.com/publications/2013/program_management_change_management_strategic_initiative_management_pmo_imperative

Kinder, T., Stenvall, J., Koskimies, E., Webb, H., & Janenova, S. (2023). Local public services and the ethical deployment of artificial intelligence. *Government Information Quarterly*, *40*(4), October 2023, 101865.

Krafft, P. M., Young, M., Katell, M., Lee, J. E., Narayan, S., Epstein, M. . . . & Barghouti, B. R. B. (2021). An action-oriented AI policy toolkit for technology audits by community advocates and activists. *FAccT '21*, March 3-10 2021. Retrieved February 2, 2024 from https://dl.acm.org/doi/pdf/10.1145/3442188.3445938

Kung, J., Boskovic, G., & Stix, C. (2020). Building an AI world: Report on national and regional AI strategies. Second Edition. CIFAR. Retrieved February 2, 2024 from https://cifar.ca/wp-content/uploads/2020/10/building-an-ai-world-second-edition.pdf

Kuziemski, M. & Misuraca, G. (2020). AI governance in the public sector: Three tales from the frontiers of automated decision-making in democratic settings. *Telecommunications Policy*, *44*(6), July 2020, 101976.

Law Commission of Ontario (2021). Regulating AI: Critical issues and choices. Retrieved February 2, 2024 from https://www.lco-cdo.org/wp-content/uploads/2021/04/LCO-Regulating-AI-Critical-Issues-and-Choices-Toronto-April-2021-1.pdf

Lehdonvirta, V. (2022). *Cloud empires: How digital platforms are overtaking the state and how we can regain control*. The MIT Press.

Lepage-Richer, T. & McKelvey, F. (2022). States of computing: On government organization and artificial intelligence in Canada. *Big Data & Society*, *9*(2), July-December 2022.

Lewis, J. E., Abdilla, A., Arista, N., Baker, K., Benesiinaabandan, S., Brown, M., Cheung, M. . . . Whaanga, H. (2020). Indigenous protocol and artificial intelligence position paper. The Initiative for Indigenous Futures and the Canadian Institute for Advanced Research (CIFAR). Retrieved February 2, 2024 from https://www.indigenous-ai.net/position-paper

Lewis, J. E., Arista, N., Pechawis, A., & Kite, S. (2018). Making kin with the machines. *Journal of Design and Science*.

Ligozat, A.-L., Lefevre, J., Bugeau, A., & Combaz, J. (2022). Unraveling the hidden environmental impacts of AI solutions for environment life cycle assessment of AI solutions. *Sustainability*, *14*(9), 5172.

Madan, R. & Ashok, M. (2023). AI adoption and diffusion in public administration: A systematic literature review and future research agenda. *Government Information Quarterly*, *40*(1), January 2023, 101774.

Mäntymäki, M., Minkkinen, M., Birkstedt, T., & Viljanen, M. (2022). Defining organizational AI governance. *AI and Ethics*, *2*, 603-609.

Mazzucato, M., Schaake, M., Krier, S., & Entsminger, J. (2022). Governing artificial intelligence in the public interest. UCL Institute for Innovation and Public Purpose, Working Paper Series (IIPP WP 2022-12). Retrieved February 2, 2024 from https://www.ucl.ac.uk/bartlett/public-purpose/sites/bartlett_public_purpose/files/governing_ai_in_public_interest.pdf

McKelvey, F., McPhail, B., & Rajabiun, R. (2022, February 2). AI accountability can't be left to the CRTC. *Policy Options*. Retrieved February 2, 2024 from https://policyoptions.irpp.org/magazines/february-2022/ai-accountability-crtc-oversight/

McKelvey, F. & MacDonald, M. (2019). Artificial intelligence policy innovations at the Canadian federal government. *Canadian Journal of Communication*, *44*(2), 43-50.

Miceli, M., Posada, J., & Yang, T. (2022). Studying up machine learning data: Why talk about bias when we mean power? *Proceedings of the ACM on Human-Computer Interaction*, 6, Article #34, 1-14. Retrieved February 2, 2024 from https://arxiv.org/pdf/2109.08131.pdf

Miceli, M. & Posada, J. (2022). The data-production dispositif. *Proceedings of the ACM on Human-Computer Interaction*, 6, CSCW 2022. Retrieved February 2, 2024 from https://arxiv.org/pdf/2205.11963.pdf

Mishra, S., Clark, J., & Perrault, C. R. (2020). Measurement in AI policy: Opportunities and challenges. Retrieved February 2, 2024 from http://arxiv.org/abs/2009.09071

Munn, L. (2023). The five tests: Designing and evaluating AI according to Indigenous Māori principles. *AI & Society*.

Munro, D. (2019). Governing AI: Navigating risks, rewards, and uncertainty. *Public Policy Forum*. Retrieved February 2, 2024 from https://ppforum.ca/wp-content/uploads/2019/01/Governing-AI-PPF-Jan2019-EN.pdf

National Artificial Intelligence Initiative Office (2023). National artificial intelligence initiative: Overseeing and implementing the United States national AI strategy. Retrieved February 2, 2024 from https://www.ai.gov/

Nitzberg, M. & Zysman, J. (2022). Algorithms, data, and platforms: The diverse challenges of governing AI. *Journal of European Public Policy*, *29*(11), 1753-1778.

Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic analysis: Striving to meet the trustworthiness criteria. *International Journal of Qualitative Methods*, *16*, 1-13.

OECD (2022a). OECD framework for the classification of AI systems. OECD Digital Economy Papers, No. 323. https://www.oecd-ilibrary.org/science-and-technology/oecd-framework-for-the-classification-of-ai-systems_cb6d9eca-en

OECD (2022b). National strategies, agendas and plans. OECD AI Policy Observatory. Retrieved February 2, 2024 from https://oecd.ai/en/dashboards/policy-instruments/National_strategies_agendas_and_plans

OECD (2022c). Measuring the environmental impact of artificial intelligence compute and applications: The AI footprint. OECD Digital Economy Papers #341. Retrieved February 2, 2024 from https://www.oecd.org/publications/measuring-the-environmental-impacts-of-artificial-intelligence-compute-and-applications-7babf571-en.htm

Office of the Prime Minister (2021, December 16). Minister of Innovation, Science and Industry mandate letter. Retrieved February 2, 2024 from https://pm.gc.ca/en/mandate-letters/2021/12/16/minister-innovation-science-and-industry-mandate-letter

Office of the Privacy Commissioner of Canada (2018). About the Contributions Program. Retrieved February 2, 2024 from https://priv.gc.ca/en/opc-actions-and-decisions/research/funding-for-privacy-research-and-knowledge-translation/about-the-contributions-program/

Office of the Privacy Commissioner of Canada (2020). A regulatory framework for AI: Recommendations for PIPEDA reform. Retrieved February 2, 2024 from https://www.priv.gc.ca/en/about-the-opc/what-we-do/consultations/completed-consultations/consultation-ai/reg-fw_202011/

Office of the Privacy Commissioner of Canada (2021a). Police use of facial recognition technology in Canada and the way forward. Retrieved February 2, 2024 from https://www.priv.gc.ca/en/opc-actions-and-decisions/ar_index/202021/sr_rcmp/

Office of the Privacy Commissioner of Canada (2021b). Submission of the Office of the Privacy Commissioner of Canada on Bill C-11, the Digital Charter Implementation Act, 2020. Retrieved February 2, 2024 from https://www.priv.gc.ca/en/opc-actions-and-decisions/submissions-to-consultations/sub_ethi_c11_2105/

Ouimette, M.-E. (2022, June 21). Canada's next industrial AI strategy needs to address adoption. *Policy Options*. Retrieved February 2, 2024 from https://policyoptions.irpp.org/magazines/june-2022/canada-ai-strategy-adoption/

Parliament of Canada (2020). Digital Charter Implementation Act, 2020. Retrieved February 2, 2024 from https://www.parl.ca/LegisInfo/en/bill/43-2/C-11

Parliament of Canada (2022). Digital Charter Implementation Act, 2022. Retrieved February 2, 2024 from https://www.parl.ca/legisinfo/en/bill/44-1/c-27

Pasquale, F. (2020). *New laws of robotics: Defending human expertise in the age of AI*. The Belknap Press of Harvard University Press.

Patterson, D., Gonzalez, J. Le, Q., Liang, C. Munguia, L.-M., Rothschild, D., So. D. . . . Dean, J. (2021). Carbon emissions and large neural network training. Retrieved March 20, 2023 from https://arxiv.org/abs/2104.10350

Radu, R. (2021). Steering the governance of artificial intelligence: National strategies in perspective. *Policy and Society*, *40*(2), 178-193.

Reeveley, D. (2021, October 26). Federal rules on AI too narrow and risk 'damaging public trust': Internal review. *The Logic*. Retrieved February 2, 2024 from https://thelogic.co/news/federal-rules-on-ai-too-narrow-and-risk-damaging-public-trust-internal-review/

Ricaurte, P. (2022, March 4). Artificial intelligence and the feminist decolonial imagination. *Bot Populi*. https://botpopuli.net/artificial-intelligence-and-the-feminist-decolonial-imagination/

Rice, F. (2021, June 11). Meaningful engagement for technology in the public sphere. *IT World Canada*. Retrieved February 2, 2024 from https://www.itworldcanada.com/blog/meaningful-engagement-for-technology-in-the-public-sphere/454605

Roberts, A. (2006). *Blacked out: Government secrecy in the Information Age*. Cambridge University Press.

Roberts, H., Hine, E., Taddeo, M., & Floridi, L. (2023). Global AI governance: Barriers and pathways forward. *SSRN*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4588040

Robertson, K., Khoo, C., & Song, Y. (2020). To surveil and predict: A human rights analysis of algorithmic policing in Canada. *Citizen Lab and University of Toronto Faculty of Law*. Retrieved February 2, 2024 from https://citizenlab.ca/wp-content/uploads/2020/09/To-Surveil-and-Predict.pdf

RSA (2018). Artificial intelligence: Real public engagement. Retrieved February 2, 2024 from https://www.thersa.org/globalassets/pdfs/reports/rsa_artificial-intelligence---real-public-engagement.pdf

Sambasivan, N., Kapania, S., Highfill, H., Akrong, D., Paritosh, P., & Arroyo, L. (2021). 'Everyone wants to do the model work, not the data work': Data cascades in high-stakes AI. *CHI '21*, May 8-13, 2021. Retrieved February 2, 2024 from https://storage.googleapis.com/pub-tools-public-publication-data/pdf/0d556e45afc54afeb2eb6b51a9bc1827b9961ff4.pdf

Saunders, M., Mann, R., & Smith, R. (2008). Implementing strategic initiatives: A framework of leading practices. *International Journal of Operations & Production Management*, *28*(11), 1095-1123.

Scassa, T. (2022a, August 2). Bill C-27 and a human rights-based approach to data protection. Retrieved February 2, 2024 from https://www.teresascassa.ca/index.php?option=com_k2&view=item&id=361:bill-c-27-and-a-human-rights-based-approach-to-data-protection&Itemid=80

Scassa, T. (2022b, May 17). Comments on the Third review of the Directive on Automated Decision-making. Retrieved February 2, 2024 from https://www.teresascassa.ca/index.php?option=com_k2&view=item&id=354:comments-on-the-third-review-of-canadas-directive-on-automated-decision-making&Itemid=80

Scassa, T. (2022c, August 8). Canada's proposed AI & Data Act - Purpose and application. Retrieved February 2, 2024 from https://www.teresascassa.ca/index.php?option=com_k2&view=item&id=362:canadas-proposed-ai--data-act-purpose-and-application&Itemid=80

Scassa, T. (2022d, August 29). Oversight and enforcement under Canada's proposed AI and Data Act. Retrieved February 2, 2024 from https://www.teresascassa.ca/index.php?option=com_k2&view=item&id=365:oversight-and-enforcement-under-canadas-proposed-ai-and-data-act&Itemid=80

Scassa, T. (2022e, October 11). Regulating AI in Canada - The federal government and the AIDA. Retrieved February 2, 2024 from https://www.teresascassa.ca/index.php?option=com_k2&view=item&id=366:regulating-ai-in-canada-the-federal-government-and-the-aida&Itemid=80

Scassa, T. (2022f, August 22). The unduly narrow scope for 'harm' and 'biased output' under the AIDA. Retrieved February 2, 2024 from https://www.teresascassa.ca/index.php?option=com_k2&view=item&id=364:the-unduly-narrow-scope-for-harm-and-biased-output-under-the-aida&Itemid=80

Scott, J. C. (1998). *Seeing like a state: How certain schemes to improve the human condition have failed*. Yale University Press.

Shared Services Canada (2022). Artificial Intelligence. Retrieved February 2, 2024 from https://www.canada.ca/en/shared-services/corporate/artificial-intelligence.html

Sieber, R. & Brandusescu, A. (2021). Civic empowerment in the development and deployment of AI systems: Final report of FAccT CRAFT workshop. *FAccT CRAFT*: *Critiquing and rethinking accountability, fairness and transparency workshop, June 30, 2021*. Retrieved February 2, 2024 from https://aifortherestofus.ca/wp-content/uploads/2021/10/Civic-Empowerment-Workshop-Report-FAccT_CRAFT.pdf

Sieber, R. & Johnson, P. A. (2015). Civic open data at a crossroads: Dominant models and current challenges. *Government Information Quarterly*, *32*(3), 308-215.

Skogstad, G. & Wilder, M. (2020). Federalism and Canada's economic union. In H. Bakvis & G. Skogstad (Eds.), *Canadian federalism: Performance, effectiveness, and legitimacy*. (4th ed., 195-221). University of Toronto Press.

Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research*, *104*, 333-339.

SQF Institute (2017). SQFI ethical sourcing. Edition 2.1. Retrieved February 2, 2024 from https://www.sqfi.com/wp-content/uploads/2019/02/Ethical-Sourcing-Code.pdf

Stahl, B. C., Antoniou, J., Ryan, M. Macnish, K., & Jiya, T. (2022). Organisational responses to ethical issues of artificial intelligence. *AI & Society*, *37*, 23-37.

Stevens, Y. & Brandusescu, A. (2021). Weak privacy, weak procurement: The state of facial recognition in Canada. *Centre for Media, Technology, and Democracy, McGill University, and Max Bell School of Public Policy*. Retrieved February 2, 2024 from https://static1.squarespace.com/static/5ea874746663b45e14a384a4/t/606c811b9aa6f40dbf84d858/1617723678314/Stevens_Brandusescu_FRT.pdf

Stevens, Y. & Solomun, S. (2021, February 17). Facing the realities of facial recognition: Recommendations for Canada's Privacy Act. *Cybersecure Policy Exchange and Centre for Media, Technology, and Democracy*. Retrieved February 2, 2024 from https://www.cybersecurepolicy.ca/frt-privacy-act

Strubell, E., Ganesh, A., & McCallum, A. (2019). Energy and policy considerations for deep learning in NLP. *57th Annual Meeting of the Association for Computational Linguistics*, July 2019. Retrieved February 2, 2024 from https://arxiv.org/pdf/1906.02243.pdf

Tallberg, J., Erman, E., Furendal, M., Geith, J., Klamberg, & Lundgren, M. (2023). The global governance of artificial intelligence: Next steps for empirical and normative research. *International Studies Review*, *25*(3), viad040.

Tessono, C., Stevens, Y., Malik, M. M., Solomun, S., Dwivedi, S., & Andrey, S. (2022). AI oversight, accountability and protecting human rights: Comments on Canada's proposed Artificial Intelligence and Data Act. *Cybersecure Policy Exchange, Center for Information Technology Policy at Princeton University, and Centre for Media,* *Technology and Democracy*. Retrieved February 2, 2024 from https://www.cybersecurepolicy.ca/aida

Toronto Police Services Board (2022). Use of new artificial intelligence technology policy – public consultation. Retrieved February 2, 2024 from https://tpsb.ca/ai

Towards Data Science (2021, December 3). Leading by example to improve civic life. Retrieved February 2, 2024 from https://towardsdatascience.com/leading-by-example-to-improve-civic-life-550d88a16b27

Treasury Board of Canada Secretariat (2022). 3rd review of the Directive on Automated Decision-making: Summary of key issues, policy recommendations, and proposed amendments. Retrieved February 2, 2024 from https://wiki.gccollab.ca/images/archive/f/f4/20230125001443!DADM_3rd_Review_-_Phase_2_Consultation_Deck_(EN).pdf

Vasileva, D. (2019). Initiative vs. project: The wrong approach to strategic planning. Kanbanize. Retrieved February 2, 2024 from https://kanbanize.com/blog/initiative-vs-project/#:~:text=They%20both%20represent%20a%20focused,indicating%20their%20rather%20operational%20character.

Villemure, R. (2022, November 4) House of Commons debates. Parliament of Canada. House of Commons. *Edited Hansard, 151*(125), 44th Parliament, 1st session. 9358-9361. Retrieved February 2, 2024 from https://www.ourcommons.ca/Content/House/441/Debates/125/HAN125-E.PDF

Wan, S. & Sieber, R. E. (2023). Artificial intelligence (AI) adoption in Canadian municipalities: In-house development versus outsourcing. *The 18th International Conference on Computational Urban Planning and Urban Management*. https://osf.io/fbxgm

Whittaker, M. (2021). The steep cost of capture. *Interactions*, *28*(2), 50-55. Retrieved February 2, 2024 from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4135581#

Williams, R. (2022, November 4). House of Commons debates. Parliament of Canada. House of Commons. *Edited Hansard, 151*(125), 44th Parliament, 1st session. 9352-9356. Retrieved February 2, 2024 from https://www.ourcommons.ca/Content/House/441/Debates/125/HAN125-E.PDF

Wilson, C. (2022). Public engagement and AI: A values analysis of national strategies. *Government Information Quarterly*, *39*(1), 101652.

Wilson, C. & van der Velden, M. (2022). Sustainable AI: An integrated model to guide public sector decision-making. *Technology in Society*, *68*, 101926.

Wirtz, B. W., Weyerer, J. C., & Kehl, I. (2022). Governance of artificial intelligence: A risk and guideline-based integrative framework. *Government Information Quarterly*, *39*(4), October 2022, 101685.

Witzel, M. (2022, August 11). A few questions about Canada's Artificial Intelligence and Data Act. *Centre for International Governance Innovation*. Retrieved February 2, 2024 from https://www.cigionline.org/articles/a-few-questions-about-canadas-artificial-intelligence-and-data-act/

Wu, C.-J., Raghavendra, R., Gupta, U., Acun, B., Ardalani, N., Maeng, K., Chang, G. . . . Hazelwood, K. (2021). Sustainable AI: Environmental implications, challenges and opportunities. *Proceedings of the 5th MLSys Conference*, 2022. Retrieved February 2, 2024 from https://proceedings.mlsys.org/paper_files/paper/2022/file/462211f67c7d858f663355eff93b745e-Paper.pdf

Zuiderwijk, A., Chen, Y-C., & Salem, F. (2021). Implications of the use of artificial intelligence in public governance: A systematic literature review and a research agenda. *Government Information Quarterly*, *38*(3), July 2021, 101577.

References for Chapter 4 (AI Governance Systems in Canada)

Aloisi, A., & De Stefano, V. (2023). Between risk mitigation and labour rights enforcement: Assessing the transatlantic race to govern AI-driven decision-making through a comparative lens. *European Labour Law Journal*, *14*(2), 238-307.

Ashok, M., Madan, R., Joha, A., & Sivarajah, U. (2022). Ethical framework for artificial intelligence and digital technologies. *International Journal of Information Management*, *62*, February 2022, 102433.

Attard-Frost, B., Brandusescu, A., & Lyons, K. (2024). The governance of artificial intelligence in Canada: Findings and opportunities from a review of 84 AI governance initiatives. *Government Information Quarterly*, *41*(2), June 2024, 101929.

Attard-Frost, B. (2023). Queering intelligence: A theory of intelligence as performance and a critique of individual and artificial intelligence. In M. Klipphahn-Karge, A.-K. Koster, & S. Morais dos Santos Bruss (Eds.), *Queer Reflections on AI: Uncertain Intelligences* (23-39). Routledge.

Attard-Frost, B. & Widder, D. G. (2023). The ethics of AI value chains. https://arxiv.org/ftp/arxiv/papers/2307/2307.16787.pdf

Birkstedt, T., Minkkinen, M., Tandon, A., & Mäntymäki, M. (2023). AI governance: Themes, knowledge gaps and future agendas. *Internet Research*, *33*(7), 133-167.

Blili-Hamelin, B., Hancox-Li, L., & Smart, A. (2024). Unsocial intelligence: A pluralistic, democratic, and participatory investigation of AGI discourse. *arXiv:2401.13142v2*. https://arxiv.org/pdf/2401.13142

Brandusescu, A. (2021). Artificial intelligence policy and funding in Canada: Public investments, private interests. Centre for Interdisciplinary Research on Montreal, McGill

University. https://www.mcgill.ca/centre-montreal/files/centre-montreal/aipolicyandfunding_report_updated_mar5.pdf

Brinkmann, F. (2019). Topical discourse structures: Using topic modeling in discourse analysis approaches. *Human IT*, *14*(3), 83-114.

Cachet-Rosset, G. & Klarsfeld, A. (2023). Diversity, equity, and inclusion in artificial intelligence: An evaluation of guidelines. *Applied Artificial Intelligence*, *37*(1), Article 2176618.

CIFAR (2024). The Pan-Canadian AI Strategy. https://cifar.ca/ai/

Cihon, P., Schuett, J., & Baum, S. D. (2021). Corporate governance of artificial intelligence in the public interest. *Information*, *12*, Article 275.

Crawford, K. (2021). *Atlas of AI: Power, politics, and the planetary costs of artificial intelligence.* Yale University Press.

Crawford, K. & Joler, V. (2018). Anatomy of an AI system: The Amazon Echo as an anatomical map of human labor, data and planetary resources. https://web.archive.org/web/20240103124839/https://anatomyof.ai/

Delgado, F., Yang, S., Madaio, M., & Yang, Q. (2023). The participatory turn in AI design: Theoretical foundations and the current state of practice. *EAAMO '23, October 30–November 1, 2023*. https://arxiv.org/pdf/2310.00907.pdf

Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C. Crick, T. . . . & Williams, M. D. (2021). Artificial intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management*, *57*, April 2021, 101994.

European Commission (2023). The European AI Alliance. https://digital-strategy.ec.europa.eu/en/policies/european-ai-alliance

Faveri, B. & Auld, G. (2023, December 10). Informing possible futures for the use of third-party audits in AI regulations. Carleton University, School of Public Policy and Administration. https://carleton.ca/rgi/wp-content/uploads/Informing-Possible-Futures_RGI_AI-Workshop-Background-Document_Dec-11-2023_Final.pdf

Figaredo, D. D. & Stoyanovich, J. (2023). Responsible AI literacy: A stakeholder-first approach. *Big Data & Society*, *10*(2), July 2023.

Frost, R. & Lyons, K. (2017). Service systems analysis methods and components: A systematic literature review. *Service Science*, *9*(3), 219-234.

Frost, R. B., Cheng, M., & Lyons, K. (2019). A multilayer framework for service system analysis. In P. P. Maglio, C. A. Kieliszewski, J. C. Spohrer, K. Lyons, L. Patricio, & Y. Sawatani (Eds.), *Handbook of Service Science, Volume II* (285-306). Springer.

Frost, R B. (2020). Emerging national artificial intelligence innovation systems in Canada and China: Strategic governance and institutional evolution. *SPACES Online: Spatial Aspects Concerning Economic Structures*, *16*(1). https://www.spaces-online.com/include/SPACES_2020-01%20Frost.pdf

Government of Canada (2023a). Algorithmic Impact Assessment tool. https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/responsible-use-ai/algorithmic-impact-assessment.html

Government of Canada (2023b). Directive on Automated Decision-making. https://www.tbs-sct.canada.ca/pol/doc-eng.aspx?id=32592

Government of Canada (2024). Guide on the use of generative AI. https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/responsible-use-ai/guide-use-generative-ai.html

Government of Ontario (2023). Ontario's Trustworthy Artificial Intelligence (AI) Framework. https://www.ontario.ca/page/ontarios-trustworthy-artificial-intelligence-ai-framework

Haugen, S., Hallstrom, L., Grant, P., Cha, J., & MacQuarrie, P. (2021). Policy responses to automation in Canada. *Journal of Rural and Community Development*, *16*(1), 68-85.

Häußermann, J. J., & Lütge, C. (2021). Community-in-the-loop: Towards pluralistic value creation in AI, or—why AI needs business ethics. *AI and Ethics*, *2*, 341-362.

Hsieh, H.-F. & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, *15*(9), 1277-1288.

Innovation, Science and Economic Development Canada (2022). Learning together for responsible artificial intelligence: Report of the Public Awareness Working Group. https://ised-isde.canada.ca/site/advisory-council-artificial-intelligence/sites/default/files/attachments/2023/learning_together_for_responsible_artificial_intelligence_minisi_approved_20230303.pdf

Innovation, Science and Economic Development Canada (2022). Pan-Canadian Artificial Intelligence Strategy. *Government of Canada*. https://ised-isde.canada.ca/site/ai-strategy/en

Innovation, Science and Economic Development Canada (2023, September). Voluntary Code of Conduct on the Responsible Development and Management of Advanced Generative AI Systems. *Government of Canada*. https://ised-isde.canada.ca/site/global-innovation-clusters/en

Innovation, Science and Economic Development Canada (2023). Global Innovation Clusters. *Government of Canada*. https://ised-isde.canada.ca/site/global-innovation-clusters/en

Innovation, Science and Economic Development Canada (2024). Mandate. *Government of Canada*. https://ised-isde.canada.ca/site/ised/en/about-us/our-organization/mandate

Karanicolas, M. (2019). To err is human, to audit divine: A critical assessment of Canada's AI Directive. *Journal of Parliamentary and Political Law*, *14*(1).

Kinder, T., Stenvall, J., Koskimies, E., Webb, H., & Janenova, S. (2023). Local public services and the ethical deployment of artificial intelligence. *Government Information Quarterly*, *40*(4), October 2023, 101865.

Kuhn, T. (2012). Negotiating the micro-macro divide: Thought leadership from organizational communication for theorizing organization. *Management Communication Quarterly*, *26*(4), 543-584.

Lepage-Richer, T. & McKelvey, F. (2022). States of computing: On government organization and artificial intelligence in Canada. *Big Data & Society*, *9*(2), July 2022.

Liebig, L., Güttel, L., Jobin, A., & Katzenbach, C. (2022). Subnational AI policy: Shaping AI in a multi-level governance system. *AI & Society*. https://link.springer.com/article/10.1007/s00146-022-01561-5

Lyons, K. & Tracy, S. (2013). Characterizing organizations as service systems. *Human Factors and Ergonomics in Manufacturing & Service Industries*, *23*(1), 19-27.

Madan, R. & Ashok, M. (2023). AI adoption and diffusion in public administration: A systematic literature review and future research agenda. *Government Information Quarterly*, *40*(1), January 2023, 101774.

Mäntylä, M. V., Graziotin, D., & Kuutila, M. (2018). The evolution of sentiment analysis–A review of research topics, venues, and top cited papers. *Computer Science Review*, *27*, 16-32.

Mäntymäki, M., Minkkinen, M., Birkstedt, T., & Viljanen, M. (2022). Defining organizational AI governance. *AI and Ethics*, *2*, 603-609.

Maragno, G., Tangi, L., Gastaldi, L., & Benedetti, M. (2023). Exploring the factors, affordances, and constraints outlining the implementation of artificial intelligence in public sector organizations. *International Journal of Information Management*, *73*, December 2023, 102686.

McKelvey, B. & Aldrich, H. (1983). Populations, natural selection, and applied organizational science. *Administrative Science Quarterly*, *28*(1), 101-128.

McKelvey, F. & MacDonald, M. (2019). Artificial intelligence policy innovations at the Canadian federal government. *Canadian Journal of Communication*, *44*(2), 43-50.

Merola, R. (2022). Inclusive growth in the era of automation and AI: How can taxation help? *Frontiers in Artificial Intelligence*, *5*, 867832.

Mila (2024). AI Policy Compass. https://mila.quebec/en/ai-society/ai-policy-compass/

Mishra, S., Clark, J., & Perrault, C. R. (2020). Measurement in AI policy: Opportunities and challenges. http://arxiv.org/abs/2009.09071

National Artificial Intelligence Initiative Act of 2020, 15 U.S.C. § 9411 (2020). https://uscode.house.gov/view.xhtml?path=/prelim@title15/chapter119&edition=prelim

Nitzberg, M. & Zysman, J. (2022). Algorithms, data, and platforms: The diverse challenges of governing AI. *Journal of European Public Policy*, *29*(11), 1753-1778.

Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic analysis: Striving to meet the trustworthiness criteria. *International Journal of Qualitative Methods*, *16*, 1-13.

OECD (2022). OECD framework for the classification of AI systems. OECD Digital Economy Papers, No. 323. https://www.oecd-ilibrary.org/science-and-technology/oecd-framework-for-the-classification-of-ai-systems_cb6d9eca-en

OECD (2024). OECD AI Principles overview. *OECD.AI Policy Observatory*. https://oecd.ai/en/ai-principles

Office of the Privacy Commissioner of Canada (2018). About the Contributions Program. https://priv.gc.ca/en/opc-actions-and-decisions/research/funding-for-privacy-research-and-knowledge-translation/about-the-contributions-program/

Office of the Privacy Commissioner of Canada (2022). Consultations. https://www.priv.gc.ca/en/about-the-opc/what-we-do/consultations/

Office of the Privacy Commissioner of Canada (2023, December 7). Principles for responsible, trustworthy and privacy-protective generative AI technologies. https://www.priv.gc.ca/en/privacy-topics/technology/artificial-intelligence/gd_principles_ai/

Ogunleye, I. (2022, July). AI's redress problem: Recommendations to improve consumer protection from artificial intelligence. *Center for Long-Term Cybersecurity, UC Berkeley*. https://cltc.berkeley.edu/wp-content/uploads/2022/08/AIs_Redress_Problem.pdf

O'Keefe, C., Cihon, P., Garfinkel, B., Flynn, C., Leung, J., & Dafoe, A. (2020). The windfall clause: Distributing the benefits of AI for the common good. *Future of Humanity Institute, University of Oxford*. https://www.fhi.ox.ac.uk/wp-content/uploads/Windfall-Clause-Report.pdf

Parliament of Canada (2022, October). Facial recognition and the growing power of artificial intelligence: Report of the Standing Committee on Access to Information, Privacy and Ethics. https://www.ourcommons.ca/DocumentViewer/en/44-1/ETHI/report-6/

Parliament of Canada (2023). Implications of artificial intelligence for the Canadian labour force. Standing Committee on Human Resources, Skills and Social Development and the Status of Persons with Disabilities. https://www.ourcommons.ca/committees/en/HUMA/StudyActivity?studyActivityId=12317899

Parliament of Canada (2024). Bill C-27 [Study of the Digital Charter Implementation Act conducted by the Standing Committee on Industry and Technology]. https://www.ourcommons.ca/Committees/en/INDU/StudyActivity?studyActivityId=12157763

Parliament of Canada (2024). Digital Charter Implementation Act, 2022. https://www.parl.ca/legisinfo/en/bill/44-1/c-27

Parsons, T. (1956). Suggestions for a sociological approach to the theory of organizations. *Administrative Science Quarterly*, *1*(1), 63-85.

Peretz-Andersson, E., Tabares, S., Mikalef, P., & Parida, V. (2024). Artificial intelligence implementation in manufacturing SMEs: A resource orchestration approach. *International Journal of Information Management*, *77*, August 2024, 102781.

Radu, R. (2021). Steering the governance of artificial intelligence: National strategies in perspective. *Policy and Society*, *40*(2), 178-193.

Ricaurte, P. (2022). Ethics for the majority world: AI and the question of violence at scale. *Media, Culture, & Society*, *44*(4), 726-745.

Roche, C., Wall, P. J., & Lewis, D. (2023). Ethics and diversity in artificial intelligence policies, strategies and initiatives. *AI and Ethics*, *3*, 1095-1115.

Scassa, T. (2021). Administrative law and the governance of automated decision making: A critical look at Canada's Directive on Automated Decision Making. *UBC Law Review*, *54*(1), Article 7.

Scassa, T. (2023). Regulating AI in Canada: A critical look at the proposed Artificial Intelligence and Data Act. *Canadian Bar Review*, *101*(1), 2023

Schmitt, L. (2022). Mapping global AI governance: A nascent regime in a fragmented landscape. *AI and Ethics*, *2*, 303-314.

Schwartz, R., Vassilev, A., Greene, K., Perine, L., Burt, A., & Hall, P. (2022). Towards a standard for identifying and managing bias in artificial intelligence. *NIST*. NIST Special Publication 1270. https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.1270.pdf

Shelby, R., Rismani, S., Henne, K., Moon, A., Rostamzadeh, N., Nicholas, P. . . . & Virk, G. (2023). Sociotechnical harms of algorithmic systems: Scoping a taxonomy for harm reduction. *ACM AIES 2023, August 9-11*. https://arxiv.org/pdf/2210.05791.pdf

Stahl, B. C. (2022). Responsible innovation ecosystems: Ethical implications of the application of the ecosystem concept to artificial intelligence. *International Journal of Information Management*, *62*, February 2022, 102411.

Standards Council of Canada (2023). AI and data governance. https://www.scc.ca/en/flagships/data-governance

Suchman, L. (2023). The uncontroversial 'thingness' of AI. *Big Data & Society*, *10*(2), July 2023.

Taeihagh, A. (2021). Governance of artificial intelligence. *Policy and Society*, *40*(2), 137-157.

Tallberg, J., Erman, E., Furendal, M., Geith, J., Klamberg, M., & Lundgren, M. (2023). The global governance of artificial intelligence: Next steps for empirical and normative research. *International Studies Review*, *25*(3), viad040.

Tessono, C., Stevens, Y., Malik, M. M., Solomun, S., Dwivedi, S., & Andrey, S. (2022, November). AI oversight, accountability and protecting human rights: Comments on Canada's proposed Artificial Intelligence and Data Act. *Cybersecure Policy Exchange*. https://www.cybersecurepolicy.ca/aida

The White House (2023, October 30). Executive order on the safe, secure, and trustworthy development and use of artificial intelligence. https://www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/

Towards Data Science (2021). Leading by example to improve civic life. https://towardsdatascience.com/leading-by-exampleto-improve-civic-life-550d88a16b27

United Kingdom Office for Artificial Intelligence (2023). A pro-innovation approach to AI regulation. *Gov.uk*. https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach/white-paper

Veale, M., Matus, K., & Gorwa, R. (2023). AI and global governance: Modalities, rationales, tensions. *Annual Review of Law and Social Science*, *19*, 255-275.

Vector Institute (2024). FastLane. https://vectorinstitute.ai/programs/ai-startup-and-scale-up-program/

Vibert, C. (2004). Population ecology perspective. In C. Vibert (Ed.), *Theories of Macro-organizational Behavior: A Handbook of Ideas and Explanations* (101-106). Routledge.

Wan, S. & Sieber, R. E. (2023). Artificial intelligence (AI) adoption in Canadian municipalities: In-house development versus outsourcing. *The 18th International Conference on Computational Urban Planning and Urban Management*. https://osf.io/fbxgm

West, S. M., Whittaker, M., & Crawford, K. (2019, April). Discriminating systems: Gender, race, and power in AI. *AI Now Institute*. https://ainowinstitute.org/wp-content/uploads/2023/04/discriminatingsystems.pdf

Wilson, C. (2022). Public engagement and AI: A values analysis of national strategies. *Government Information Quarterly*, *39*(1), 101652.

Ylä-Anttila, T., Eranti, V., & Kukkonen, A. (2021). Topic modeling for frame analysis: A study of media debates on climate change in India and USA. *Global Media and Communication*, *18*(1), 91-112.

Zuiderwijk, A., Chen, Y-C., & Salem, F. (2021). Implications of the use of artificial intelligence in public governance: A systematic literature review and a research agenda. *Government Information Quarterly*, *38*(3), July 2021, 101577.

References for Chapter 5 (Conclusion: A Future for Transfeminist AI Governance)

Access Now (2024, March 13). The EU AI Act: A failure for human rights, a victory for industry and law enforcement. Retrieved from https://www.accessnow.org/press-release/ai-act-failure-for-human-rights-victory-for-industry-and-law-enforcement/

Alfrink, K., Keller, I., Kortuem, G. & Doorn, N. (2023). Contestable AI by design: Towards a framework. *Minds and Machines, 33*, 613-639.

Alfrink, K., Keller, I., Semperena, M. Y., Bulygin, D., Kortuem, G. & Doorn, N. (2024). Envisioning contestability loops: Evaluating the agonistic arena as a generative metaphor for public AI. *She Ji: The Journal of Design, Economics, and Innovation*, *10*(1), 53-93.

Altman, S., Brockman, G., & Sutskever, I. (2023, May 22). Governance of superintelligence. *OpenAI*. https://openai.com/index/governance-of-superintelligence/

Amnesty International (2023, November 16). EU AI Act Trilogues: Status of fundamental rights recommendations. Retrieved from https://www.amnesty.eu/news/eu-ai-act-trilogues-status-of-fundamental-rights-recommendations/

Assembly of First Nations (2023, October). Bill C-27: The Digital Charter Act, 2023 and First Nations Rights. Brief by the Assembly of First Nations to the Parliamentary Standing Committee on Industry and Technology. Retrieved from https://www.ourcommons.ca/Content/Committee/441/INDU/Brief/BR12885140/br-external/AssemblyOfFirstNations-e.pdf

Attard-Frost, B. (2022). Comments on the 3rd review of the Directive on Automated Decision-making. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4238164

Attard-Frost, B. (2023a). Generative AI systems: Impacts on artists & creators and related gaps in the Artificial Intelligence and Data Act. Brief submitted to the Parliament of Canada

Standing Committee on Industry and Technology. Retrieved from https://www.ourcommons.ca/Content/Committee/441/INDU/Brief/BR12541028/br-external/AttardFrostBlair-e.pdf

Attard-Frost, B. (2023b). Comments on "Canadian Guardrails for Generative AI – Code of Practice". Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4562699

Attard-Frost, B. (2023c). AI countergovernance. *Midnight Sun*. https://www.midnightsunmag.ca/ai-countergovernance/

Attard-Frost, B. (2024). AI countergovernance: Opposing AI governance in theory & practice. [Slides from research talk delivered online for Essex Business School at University of Essex on March 20, 2024]. https://docs.google.com/presentation/d/1Bm2Hk4KpTp_2CGwEVZnJSsncA7IRrVBgV-EDvAk2wkw/edit?usp=sharing

Attard-Frost, B. & Hayes, H. A. (2025). Valuing value chains: The governance of AI value chains in Canada. Forthcoming chapter in N. Goodman & H. A. Hayes (eds.), *Regulating Digital*. University of Toronto Press. Preprint available on *SSRN* at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4624840

Balayn, A., Pi, Y., Widder, D. G., Alfrink, K., Yurrita, M., Upadhyay, S. . . . & Gadiraju, U. (2024). From stem to stern: Contestability along AI value chains. *CSCW Companion '24, 720-723*. https://arxiv.org/pdf/2408.01051

Barakat, H. (2024). Selective perspectives: A content analysis of The New York Times' reporting on artificial intelligence. *Computer Says Maybe*. https://www.saysmaybe.com/latest-work/selective-perspectives-nyt

Bonini, T. & Treré, E. (2024). *Algorithms of resistance: The everyday fight against platform power*. The MIT Press.

Brandusescu, A. (2021). Artificial intelligence policy and funding in Canada: Public investments, private interests. Centre for Interdisciplinary Research on Montreal, McGill University. Retrieved from https://www.mcgill.ca/centre-montreal/files/centre-montreal/aipolicyandfunding_report_updated_mar5.pdf

Brandusescu, A. & Sieber, R. (2022). Comments on the third review of Canada's Directive on Automated Decision-making. Retrieved from https://aifortherestofus.ca/wp-content/uploads/2022/07/Comments-on-the-third-review-of-Canadas-Directive-on-ADM.pdf

Castaldo, J. (2023, December 10). Federal consultations on AI regulations heavily skewed toward businesses, industry groups, say critics. *The Globe and Mail*. Retrieved from https://www.theglobeandmail.com/business/article-canada-ai-law/

Clement, A. (2023). Preliminary analysis of ISED's C-27 list of 300 stakeholder consultation meetings. *SSRN*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4658004

Corporate Europe Observatory (2023, November 24). Big tech lobbying is derailing the AI Act. Retrieved from https://corporateeurope.org/en/2023/11/big-tech-lobbying-derailing-ai-act

Crawford, K. (2021). *Atlas of AI: Power, politics, and the planetary costs of artificial intelligence.* Yale University Press.

Crawford, K. & Joler, V. (2018). Anatomy of an AI system: The Amazon Echo as an anatomical map of human labor, data and planetary resources. Retrieved from https://anatomyof.ai/

Crawford, K. & Joler, V. (2023). Calculating empires: A genealogy of technology and power since 1500. Retrieved from https://calculatingempires.net/

Dandurand, G., Blottière, M., Jorandon, G., Gertler, N., Wester, M., Chartier-Edwards, N., Roberge, J., & McKelvey, F. (2023). Training the news: Coverage of Canada's AI hype cycle (2012-2021). *Shaping AI*. Retrieved from https://espace.inrs.ca/id/eprint/13149/1/report_ShapingAI_verJ.pdf

DeVrio, A., Eslami, M., & Holsten, K. (2024). Building, shifting, & employing power: A taxonomy of responses from below to algorithmic harm. *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency (FAccT '24)*. 1093-1106. https://dl.acm.org/doi/pdf/10.1145/3630106.3658958

European Parliament (2024). Artificial Intelligence Act. P9_TA(2024)0138. Retrieved from https://www.europarl.europa.eu/doceo/document/TA-9-2024-0138_EN.pdf

Ganesh, M. I. & Moss, E. (2022). Resistance and refusal to algorithmic harms: Varieties of 'knowledge projects." *Media International Australia*, 183(1), 90-106.

Gebru, T. & Torres, É. (2024). The TESCREAL bundle: Eugenics and the promise of utopia through artificial general intelligence. *First Monday*, *29*(4), April 2024.

Government of Canada (2024). Guide on the use of generative artificial intelligence. Retrieved from https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/responsible-use-ai/guide-use-generative-ai.html

Henshall, W. (2024, April 30). There's an AI lobbying frenzy in Washington. Big tech is dominating. *Time*. Retrieved from https://time.com/6972134/ai-lobbying-tech-policy-surge/

Herrman, J. (2024, June 4). What ever happened to the AI apocalypse? *New York Magazine*. Retrieved from https://nymag.com/intelligencer/article/what-ever-happened-to-the-ai-apocalypse.html

Hine, E. (2024, September 18). Artificial intelligence laws in the US states are feeling the weight of corporate lobbying. *Nature Index*. Retrieved from https://www.nature.com/articles/d41586-024-02988-0

Innovation, Science and Economic Development Canada (2023a). Voluntary code of conduct on the responsible development and management of advanced generative AI systems. Retrieved from https://ised-isde.canada.ca/site/ised/en/voluntary-code-conduct-responsible-development-and-management-advanced-generative-ai-systems

Innovation, Science and Economic Development Canada (2023b). What we heard – Consultation on the development of a Canadian code of practice for generative artificial intelligence systems. Retrieved from https://ised-isde.canada.ca/site/ised/en/what-we-heard-consultation-development-canadian-code-practice-generative-artificial-intelligence

ISO (2024). ISO/IEC 42001:2023(en) Information technology – Artificial intelligence – Management system. Retrieved from https://www.iso.org/obp/ui/en/#iso:std:81230:en

Jiang, H., Brown, L., Cheng, J., Anonymous Artist, Khan, M., Gupta, A. . . . & Gebru, T. (2023). AI art and its impact on artists. *AIES '23: Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society*, 363-374. https://dl.acm.org/doi/10.1145/3600211.3604681

Johnson, K. (2023, September 14). Inside the Senate's private AI meeting with tech's billionaire elites. *Wired*. Retrieved from https://www.wired.com/story/senate-ai-forum-musk-zuckerberg/

Karadeglija, A. (2023, August 23). Federal government's generative-AI consultation 'odd' and 'secretive,' critics say. *National Post*. Retrieved from https://nationalpost.com/news/politics/liberals-generative-ai-consultation-odd-critics

McQuillan, D. (2022). *Resisting AI: An anti-fascist approach to artificial intelligence*. Bristol University Press.

Merica, D. (2024, September 8). How the AI industry uses a light lobbying touch to educate Congress from a corporate perspective. *AP*. Retrieved from https://apnews.com/article/artificial-intelligence-lobbying-congress-technology-regulation-d54fe64f3fd135fdca41bdaf528282b8

Minister of Innovation, Science and Industry (2023). [Letter to Chair of the House of Commons Standing Committee on Industry and Technology, November 28, 2023]. Retrieved from https://www.ourcommons.ca/content/Committee/441/INDU/WebDoc/WD12751351/12751351/MinisterOfInnovationScienceAndIndustry-2023-11-28-Combined-e.pdf

National Institute of Standards and Technology (2023, January). Artificial Intelligence Risk Management Framework (AI RMF1.0). Retrieved from https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf

OpenAI (2023). Planning for AGI and beyond. https://openai.com/index/planning-for-agi-and-beyond/

Parliament of Canada (2023). Bill C-27. https://www.parl.ca/DocumentViewer/en/44-1/bill/C-27/first-reading

Patten, D. (2023, November 12). Read full SAG-AFTRA deal summary released ahead of ratification vote; analysis of AI & streaming fund. *Deadline*. https://deadline.com/2023/11/sag-aftra-deal-summary-released-read-it-1235600852/

Prime Minister of Canada (2024, April 7). Securing Canada’s AI advantage. Retrieved from https://www.pm.gc.ca/en/news/news-releases/2024/04/07/securing-canadas-ai-advantage

Reia, J., Leach, R., & Li, S. (2025). Trans digital rights: Improving data visibility, privacy and belonging for gender-diverse communities. Digital Technology for Democracy Lab, University of Virginia (Forthcoming report).

Scassa, T. (2021). Administrative law and the governance of automated decision making: A critical look at Canada's Directive on Automated Decision Making. *UBC Law Review*, *54*(1), Article 7.

Scassa, T. (2023). Regulating AI in Canada: A critical look at the proposed Artificial Intelligence and Data Act. *Canadian Bar Review*, *101*(1), 2023

Tessono, C., Stevens, Y., Malik, M. M., Solomun, S., Dwivedi, S., & Andrey, S. (2022). AI oversight, accountability and protecting human rights: Comments on Canada's proposed Artificial Intelligence and Data Act. *Cybersecure Policy Exchange, Center for Information Technology Policy at Princeton University, and Centre for Media, Technology and Democracy*. Retrieved from https://www.cybersecurepolicy.ca/aida

The Dais & Centre for Media, Technology and Democracy (2023, November). Submission on the proposed Artificial Intelligence and Data Act. Retrieved from https://dais.ca/reports/submission-on-the-proposed-artificial-intelligence-and-data-act/

The White House (2022, October 4). Blueprint for an AI Bill of Rights: Making automated systems work for the American people. Retrieved from https://www.whitehouse.gov/ostp/ai-bill-of-rights/

The White House (2023, July 21). FACT SHEET: Biden-Harris Administration secures voluntary commitments from leading artificial intelligence companies to manage the risks posed by AI. Retrieved from https://www.wired.com/story/senate-ai-forum-musk-zuckerberg/

van der Drift, M. & Raha, N. (2020). Radical transfeminism: Trans as anti-static ethics escaping neoliberal encapsulation. In J. Cooke (Ed.), *The New Feminist Literary Studies* (13-24). Cambridge University Press.

Veale, M., Matus, K., & Gorwa, R. (2023). AI and global governance: Modalities, rationales, tensions. *Annual Review of Law and Social Science*, *19*, 255-275.

Writers Guild of America (2023). WGA Contract 2023. Retrieved from https://www.wgacontract2023.org/the-campaign/summary-of-the-2023-wga-mba

Appendices

Appendix 2A: Integrated Inventory of Ethical Concerns, Value Chain Actors, Resourcing Activities, & Sampled Sources

High-level ethical concern 1: Benefits of AI

Lower-level issues identified by Stahl et al. (2022)	Examples of related value chain actors	Examples of related resourcing activities	Related sources
Novel insights from data, efficiency improvement, economic benefits, environmental benefits, contribution to sustainable development goals, AI for Good	• Industry • Governments • Intergovernmental bodies • Civil society	• Analytics platform development & use • Adoption of industrial AI applications • Job creation & hiring • IP creation & licensing • Energy consumption optimization • Policy development & evaluation	• Aula & Bowles (2023) • Birhane (2021) • Cobbe, Veale, & Singh (2023) • Gansky & McDonald (2022) • Hind & Seitz (2022) • Madianou (2021) • Moore (2019) • Widder & Nafus (2023)

High-level ethical concern 2: Issues arising from machine learning

Lower-level issues identified by Stahl et al. (2022)	Examples of related value chain actors	Examples of related resourcing activities	Related sources
Control of data (misuse of personal data, lack of privacy, security, integrity)	• Data subjects, owners, users, & brokers • Data center operators • Model developers • Application users • Cloud services providers • Information security services providers • Red teamers, blue teamers, &	• Privacy policy development, enforcement, & application to ML systems • Informed consent and authorization for data collection, scraping, & use • Sale, purchase, brokerage, & ownership of data • Knowledge of ML vulnerabilities & vulnerability testing methods • Funding of ML research involving ethical concerns related to control of data	• European Parliament (2020) • MacKinnon & King (2022) • Veale, Binns, & Edwards (2018) • Crain (2018) • Lamdam (2022) • Fredrikson, Jha, & Ristenpart (2016) • Greshake et al. (2023)

	real adversaries • Research funding sources & funding recipients		• Wang et al. (2022) • Whittaker (2021) • Hatmaker (2022) • Perrigo (2022) • Lee, Cooper, & Grimmelmann (2023) • De Vynck (2023) • GitHub Copilot Litigation (2023) • Stable Diffusion Litigation (2023) • Vincent (2023)
Reliability (Accuracy of predictive recommendations, accuracy of non-individualized recommendations, lack of quality data, accuracy of data)	• Data providers • Model developers • Cloudwork platform providers • Application users & other application stakeholders • Research funding sources & funding recipients	• Development and implementation of ethical quality assurance practices for model training, testing, & management • Data annotation & verification and outsourcing of data work and model work • Funding of ML research involving ethical concerns related to reliability	• Angwin et al. (2016) • Bender et al. (2021) • Burr & Leslie (2023) • Eitel-Porter (2021) • Grote & Berens (2022) • Irani (2015) • Miceli, Posada, & Yang (2022) • Miceli & Posada (2022) • Mökander & Axente (2023) • Perrigo (2023) • Rankin et al. (2020)
Lack of transparency (Bias and discrimination, lack of accountability and liability)	• Data subjects • Model developers • Application users & other application stakeholders • Governments, courts, and regulatory bodies • Research funding sources &	• Incentivization & disclosure of funding for “responsible AI” and “AI ethics” research • Documentation, disclosure, and explanation of machine learning & automated decision-making processes and outcomes • Inclusion of stakeholder knowledge & perspectives • Distribution and enforcement of accountability and liability for harms amongst value chain actors	• ACLU (2023) • Ahmed, Wahed, & Thompson (2023) • Bartneck et al. (2020) • Birhane et al. (2022a, 2022b) • Broderick (2023) • Brown (2023) • Cobbe, Veale, & Singh (2023)

	funding recipients	• Community organizing & protest	• European Commission (2022) • Miceli et al. (2022) • Mitchell et al. (2019) • Ochigame (2019) • Raji et al. (2020) • Whittaker (2021) • Widder & Nafus (2023) • Zech (2021)

High-level ethical concern 3: Ethics of living in a digital world

Lower-level issues identified by Stahl et al. (2022)	**Examples of related value chain actors**	**Examples of related resourcing activities**	**Related sources**
Economic issues (Unemployment, concentration of economic power, ownership of data/IP)	• Workers • Unemployed & precariously employed people • Employment services providers • Education & training providers • Employers • Tech companies • Nations	• Use of AI applications in hiring, contracting, dismissal, & surveillance of workers • AI education & training • Distribution & redistribution of capital, profits, and other financial resources • Unionization & collective organizing • Distribution, open-sourcing, access to, & licensing of data, code, and other computational resources and IP	• Ahmed, Wahed, & Thompson (2023) • Bales & Stone (2020) • Broderick (2023) • Dastin (2018) • Dyer-Witheford, Kjøsen, & Steinhoff (2019) • Hickok & Maslej (2023) • Langenkamp & Yue (2022) • Masiello & Slater (2023) • Miceli, Posada, & Yang (2022) • Miceli & Posada (2022) • Perrigo (2023) • Webster (2023) • Widder, West, & Whittaker (2023)

Justice (Impact on justice systems, access to public services, impact on vulnerable groups, distribution, economic participation)	Justice system administrators • Public service administrators • People seeking public services or justice • Economically disadvantaged groups • Exploited & precarious workers • Employers	• Public and private funding of AI applications procured & used in justice systems and public services • Data preparation, design, and development of AI systems with impacts on justice & subsequent impacts of decision automation outcomes for justice • Inclusion of knowledges from vulnerable and marginalized groups in AI education, design, development, & governance processes • Redistribution of resources required to justly develop/use AI systems & just distribution of value co-created throughout AI system lifecycles • Macro-level social, political, and economic outcomes of widespread AI adoption	• Angwin et al. (2017) • Birhane et al. (2022a) • Dyer-Witheford, Kjøsen, & Steinhoff (2019) • Eubanks (2018) • Gans-Combe (2022) • Mulligan & Bamberger (2019) • Pasquale (2020) • Solaiman et al. (2023) • West, Whittaker, & Crawford (2019)
Human freedoms (Loss of freedom and individual autonomy, harm to physical integrity, impact on health, lack of access to and freedom of information, reduction of human contact)	• Exploited workers • Victims of material, psychological, & environmental harms • Government & corporate information owners • Companies, governments, researchers, and civil society actors seeking access to data and information	• Compensation for labor and resulting gains/losses to social, political, & economic freedoms and autonomy • Algorithmic discrimination resulting in reduction of social, political, & economic opportunities • Energy usage of AI models resulting in carbon emissions, water usage of data centers resulting in depletion of freshwater reserves • Creation, collection, and brokerage of access to proprietary data, information, and computational resources	• Ahmed, Wahed, & Thompson (2023) • Angwin et al. (2016) • Eubanks (2018) • GPAI (2022) • Miceli & Posada (2022) • Whittaker (2021)
Broader societal issues (Potential for military use, impact on	• Governments & police services • Military organizations & defense	• Military and police procurement, contracting, and use of AI applications • Mining activities & mineral extraction	• Crawford (2021) • Crawford & Joler (2018) • GPAI (2021) • Hoijtink &

environment, impact on democracy)	contractors • Cloud services providers • Data center operators • Hardware manufacturers • Social media platform providers & users	• Fuel required to ship & transport materials & hardware • Hardware manufacturing & assembly • Energy, mineral, & water consumption of AI systems • Disposal & recycling of e-waste at end of hardware lifecycle • Creation & reinforcement of filter bubbles based on algorithmic profiling of users	Hardeveld (2022) • Krönke (2019) • Li et al. (2023) • Luccioni & Hernandez-Garcia (2023) • Mahoney (2020) • Mulligan & Bamberger (2019) • Taddeo et al. (2021) • Woolley (2018)
Unknown issues (Unintended or unforeseeable adverse impacts, cost to innovation, potential for criminal and malicious use, prioritization of the “wrong” problems)	• Malicious actors • Security professionals • Governments & regulatory bodies • Research funders, research institutes, & researchers	• Unforeseen misuses/abuses of personal data & digital identities • Unforeseen consequences of social engineering, identity theft, ML model hacking, mis/disinformation, and malicious data brokerage operations • Implementation & enforcement of excessively strict or excessively permissive AI regulation • Excessive funding of misdirected AI research projects	• Ada Lovelace Institute (2021) • Brundage et al. (2018) • Smuha (2021) • Tiku (2023)

High-level ethical concern 4: Metaphysical issues

Lower-level issues identified by Stahl et al. (2022)	Examples of related value chain actors	Examples of related resourcing activities	Related sources
Machine consciousness, autonomous moral agents, super-intelligence, singularity, changes to human nature	• Governments • Intergovernmental bodies • Individual humans • Human organizations • Hypothetical conscious machines • Hypothetical artificial moral agents • Hypothetical superintelligent agents	• Data & knowledge assembled to develop hypothetical conscious machines & superintelligent agents • Consolidation of resources under the control of hypothetical conscious machines & superintelligent agents • Uneven distribution of hypothetical AI & technological capabilities resulting in divergent evolutionary trajectories	• Gebru & Torres (2023) • Torres (2023)

Appendix 3A: Inventory of Relevant Initiatives

#	Source	Initiative Title	Intervention Type	Governance Areas	Year of Origin	Discovery Method	Link
F01	Innovation, Science and Economic Development Canada	Pan-Canadian Artificial Intelligence Strategy	Strategic Plan	Data Governance, Technology Production & Use, Industry & Innovation, Social & Workforce, Public Administration	2017	External Knowledge	https://ised-isde.canada.ca/site/ai-strategy/en
F02	CIFAR	National AI Institutes	Program	AI Research	2017	External Knowledge	https://cifar.ca/ai/
F03	Natural Resources Canada	NRCan Digital Accelerator	Program	Industry & Innovation	2017	canada.ca advanced search	https://www.nrcan.gc.ca/digital-accelerator/current-artificial-intelligence-projects/22577
F04	Department of Finance Canada	Budget 2017 - Part 3 - Canada's Digital Future - Growing Canada's Advantage in Artificial Intelligence	Policy	AI Research, Industry & Innovation	2017	canada.ca advanced search	https://www.budget.canada.ca/2017/docs/plan/chap-01-en.html#Toc477707367
F05	National Research Council Canada	Advisory statement on human ethics in artificial intelligence and big data research	Ethics Statement	AI Research	2017	canada.ca advanced search	https://nrc.canada.ca/index.php/en/corporate/values-ethics/research-involving-human-participants/

							advisory-statement-human-ethics-artificial-intelligence-big-data-research-2017
F06	Standards Council of Canada	Mirror Committee to International Technical Committee on ISO/IEC JTC 1/SC 42 Artificial intelligence	Standard	Industry & Innovation, Technology Production & Use	2017	External Knowledge	https://www.scc.ca/en/standards/committees/mc-iso-iec-jtc-1-sc-42-artificial-intelligence
F07	Innovation, Science and Economic Development Canada	Global Innovation Clusters	Program	Industry & Innovation	2018	External Knowledge	https://www.ic.gc.ca/eic/site/093.nsf/eng/home
F08	CIFAR	National Program of Activities	Program	AI Education & Training, Industry & Innovation	2018	External Knowledge	https://cifar.ca/ai/national-program-of-activities/
F09	CIFAR	AI & Society	Program	AI Research, Technology Production & Use, Industry & Innovation, Data Governance, Social & Workforce, Public Administration	2018	External Knowledge	https://cifar.ca/ai/ai-society/ai-futures-policy-labs/#reports
F10	CIFAR	Canada CIFAR AI Chairs	Program	AI Research	2018	External Knowledge	https://cifar.ca/ai/canada-cifar-ai-chairs/
F11	Office of the	OPC	Program	Data	2018	canada.ca	https://priv.

	Privacy Commissioner of Canada	Contributions Program - Projects on facial recognition technology & Artificial Intelligence		Governance, Technology Production & Use, AI Research		advanced search	gc.ca/en/opc-actions-and-decisions/research/funding-for-privacy-research-and-knowledge-translation/
F12	Office of the Privacy Commissioner of Canada	OPC Consultations Program - Consultation on police use of facial recognition & Consultation on artificial intelligence	Program	Data Governance, Technology Production & Use	2018	canada.ca advanced search	https://www.priv.gc.ca/en/about-the-opc/what-we-do/consultations/
F13	Innovation, Science and Economic Development Canada	G7 Multistakeholder Conference on Artificial Intelligence	Strategic Plan	AI Research, Technology Production & Use, Industry & Innovation, Social & Workforce	2018	canada.ca advanced search	https://www.ic.gc.ca/eic/site/133.nsf/eng/home
F14	Global Affairs Canada	Charlevoix Common Vision for the Future of Artificial Intelligence	Ethics Statement	AI Research, AI Education & Training, Technology Production & Use, Industry & Innovation, Social & Workforce, Public Administration	2018	canada.ca advanced search	https://www.international.gc.ca/world-monde/international_relations-relations_internationales/g7/documents/2018-06-09-artificial-intelligence-artificielle.aspx?lang=eng

F15	Innovation, Science and Economic Development Canada	Canada-France Statement on Artificial Intelligence	Strategic Plan	AI Research, AI Education & Training, Technology Production & Use, Industry & Innovation, Social & Workforce, Public Administration	2018	canada.ca advanced search	https://www.international.gc.ca/world-monde/international_relations-relations_internationales/europe/2018-06-07-france_ai-ia_france.aspx?lang=eng
F16	Natural Resources Canada	Science and Policy Integration (SPI) Leadership Bootcamp - Artificial Intelligence Action Learning Team	Program	AI Education & Training, Public Administration	2018	canada.ca advanced search	https://www.nrcan.gc.ca/21639
F17	Public Service Commission of Canada	Integrated Intelligence Evaluation Report	Strategic Plan	Public Administration	2018	canada.ca advanced search	https://www.canada.ca/en/public-service-commission/services/publications/integrated-intelligence-evaluation-report.html
F18	Canada Revenue Agency	Corporate Business Plans (2018 to 2025)	Strategic Plan	Technology Production & Use, Data Governance, Digital Infrastructure, Public Administration	2018	External Knowledge	https://www.canada.ca/en/revenue-agency/corporate/about-canada-revenue-agency-cra/summary-corporate-business-

							plan.html
F19	Treasury Board of Canada Secretariat	Directive on Automated Decision-making	Policy	Technology Production & Use, Public Administration	2019	External Knowledge	https://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=32592
F20	Treasury Board of Canada Secretariat	Algorithmic Impact Assessment Tool	Policy	Technology Production & Use, Public Administration	2019	External Knowledge	https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/responsible-use-ai/algorithmic-impact-assessment.html
F21	Innovation, Science and Economic Development Canada - Advisory Council on Artificial Intelligence	Advisory Council on Artificial Intelligence - Annual Report 2020-21	Policy	Industry & Innovation	2019	External Knowledge	https://ised-isde.canada.ca/site/advisory-council-artificial-intelligence/en/annual-reports/annual-report-2020-21
F22	Innovation, Science and Economic Development Canada	Canada's Digital Charter	Strategic Plan	Data Governance, Technology Production & Use, Industry & Innovation, Social & Workforce, Public Administration	2019	External Knowledge	https://www.ic.gc.ca/eic/site/062.nsf/eng/h_00108.html
F23	National Research	Artificial Intelligence for	Program	AI Research, Industry &	2019	canada.ca advanced	https://nrc.canada.ca/en/

	Council Canada	Design Challenge Program		Innovation		search	research-development/research-collaboration/programs/artificial-intelligence-design-challenge-program
F24	Treasury Board of Canada Secretariat	List of Interested Artificial Intelligence (AI) Suppliers	Policy	Technology Production & Use, Public Administration	2019	External Knowledge	https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/responsible-use-ai/list-interested-artificial-intelligence-ai-suppliers.html
F25	National Research Council Canada	NRC Research Centre Programs - Artificial Intelligence	Program	AI Research	2019	canada.ca advanced search	https://nrc.canada.ca/index.php/en/research-development/research-collaboration/programs?f%5B0%5D=topic_programs%3A9553
F26	National Research Council Canada	NRC-Waterloo Collaboration on Artificial Intelligence, Internet of Things, and	Program	AI Research	2019	canada.ca advanced search	https://nrc.canada.ca/en/research-development/research-collaboration/collaborati

		Cybersecurity					on-centres/nrc-waterloo-collaboration-artificial-intelligence-internet-things-cybersecurity
F27	Innovation, Science and Economic Development Canada	Declaration of the International Panel on Artificial Intelligence	Strategic Plan	AI Research, AI Education & Training, Technology Production & Use, Industry & Innovation, Social & Workforce, Public Administration	2019	canada.ca advanced search	https://www.canada.ca/en/innovation-science-economic-development/news/2019/05/declaration-of-the-international-panel-on-artificial-intelligence.html
F28	Innovation, Science and Economic Development Canada - Canada Intellectual Property Office	Processing Artificial Intelligence: Highlighting the Canadian Patent Landscape	Policy	Industry & Innovation	2019	canada.ca advanced search	https://www.ic.gc.ca/eic/site/cipointernet-internetopic.nsf/eng/h_wr04776.html
F29	Parliament of Canada	Bill C-11 ("Digital Charter Implementation Act")	Policy	Data Governance, Technology Production & Use, Industry & Innovation	2020	External Knowledge	https://www.parl.ca/LegisInfo/en/bill/43-2/C-11
F30	Office of the Privacy Commissioner of Canada	A Regulatory Framework for AI: Recommendations for PIPEDA	Policy	Data Governance, Technology Production & Use	2020	External Knowledge	https://www.priv.gc.ca/en/about-the-opc/what-we-do/consultat

		Reform					ions/completed-consultations/consultation-ai/reg-fw_202011/
F31	Innovation, Science and Economic Development Canada	Canada-UK Artificial Intelligence Initiative	Program	AI Research	2020	canada.ca advanced search	https://www.canada.ca/en/social-sciences-humanities-research/news/2020/02/canada-uk-artificial-intelligence-initiative-projects-funded.html
F32	Global Affairs Canada	Freedom Online Coalition (FOC) Joint Statement on Artificial Intelligence and Human Rights	Ethics Statement	Technology Production & Use, Public Administration	2020	canada.ca advanced search	https://www.international.gc.ca/global-affairs-affaires-mondiales/news-nouvelles/2020/2020-11-05-internet-freedom-liberte-internet.aspx?lang=eng
F33	Innovation, Science and Economic Development Canada	Joint Statement from founding members of the Global Partnership on Artificial Intelligence	Strategic Plan	Technology Production & Use, Industry & Innovation, Social & Workforce	2020	canada.ca advanced search	https://www.canada.ca/en/innovation-science-economic-development/news/2020/06/joint-statement-from-founding-members-of-the-

							global-partnership-on-artificial-intelligence.html
F34	National Research Council Canada	Technical and Advisory Services - Artificial Intelligence	Program	AI Research, Industry & Innovation	2020	External Knowledge	https://nrc.canada.ca/en/research-development/products-services/technical-advisory-services/artificial-intelligence
F35	Innovation, Science and Economic Development Canada	International Centre of Expertise in Montréal for the Advancement of Artificial Intelligence	Program	AI Research, Industry & Innovation	2020	canada.ca advanced search	https://ceimia.org/en/
F36	Department of National Defence	Defence Capabilities Blueprint - Key Industrial Capabilities - Artificial Intelligence	Program	AI Research, Industry & Innovation, Public Administration	2020	canada.ca advanced search	http://dgpaapp.forces.gc.ca/en/defence-capabilities-blueprint/project-kic.asp?id=2
F37	Department of National Defence	Mobilizing Insights in Defence and Security (MINDS) - Defence Policy Challenges - Domains and Technology	Program	AI Research, Public Administration	2020	Followed link to "Strong, Secure, Engaged" national defence policy page from the Defence Capabilities Blueprint page	https://www.canada.ca/en/department-national-defence/programs/minds/defence-policy-challenges.html

F38	Innovation, Science and Economic Development Canada - Strategic Policy Sector	Public Awareness Working Group - Mandate	Policy	Industry & Innovation, Social & Social & Workforce	2020	External Knowledge	https://ised-isde.canada.ca/site/advisory-council-artificial-intelligence/en/public-awareness-working-group
F39	Information and Communications Technology Council	Maximizing Strengths and Spearheading Opportunity: Towards an Industrial Strategy for Canadian Artificial Intelligence	Policy	Industry & Innovation	2021	External Knowledge	https://www.ictc-ctic.ca/wp-content/uploads/2021/09/Maximizing-Strength-and-Spearheading-Opportunity.pdf
F40	Treasury Board of Canada Secretariat & Shared Services Canada	Responsible Use of Artificial Intelligence (AI) - Our Guiding Principles	Ethics Statement	Technology Production & Use	2021	External Knowledge	https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/responsible-use-ai.html#toc1
F41	Canada School of Public Service	Artificial Intelligence is Here Series	Program	AI Education & Training, Public Administration	2021	canada.ca advanced search	https://www.csps-efpc.gc.ca/events/artificial-intelligence-here-series/index-eng.aspx
F42	Canadian Trade	Canadian International	Program	Industry & Innovation	2021	canada.ca advanced	https://www.tradecommissioner.gc.c

	Commissioner Service	Innovation Program				search	a/funding-financement/ciip-pcii/index.aspx?lang=eng
F43	Standards Council of Canada	Canadian Data Governance Standardization Collaborative	Standard	Data Governance	2021	Link in ISED Advisory Council's 2020-2021 Annual Report	https://www.scc.ca/en/about-scc/publications/general/canadian-data-governance-standardization-roadmap
F44	National Research Council Canada	"3+2" Canada–Germany Collaborative Industrial Research and Development Program	Program	AI Research, Industry & Innovation	2021	canada.ca advanced search	https://nrc.canada.ca/en/stories/supporting-collaborative-projects-between-canada-germany-artificial-intelligence-value-added
F45	Innovation, Science and Economic Development Canada - Strategic Policy Sector	A Consultation on a Modern Copyright Framework for Artificial Intelligence and the Internet of Things	Policy	Industry & Innovation	2021	canada.ca advanced search	https://www.ic.gc.ca/eic/site/693.nsf/eng/00316.html
F46	Standards Council of Canada	Accreditation pilot for AI management systems	Standard	Technology Production & Use	2022	External Knowledge	https://www.scc.ca/en/news-events/news/2022/scc-launches-accreditation-pilot-for-ai-

							managemen t-systems
F47	Parliament of Canada	Bill C-27 ("Digital Charter Implementation Act, 2022")	Policy	Data Governance, Technology Production & Use, Industry & Innovation	2022	External Knowledge	https://www.parl.ca/DocumentViewer/en/44-1/bill/C-27/first-reading
F48	CIFAR	Pan-Canadian AI Compute Environment	Program	Digital Infrastructure	2022	External Knowledge	https://cifar.ca/ai/key-strategic-priorities/#advancing-ai-science
F49	CIFAR	AI for Health	Program	AI Research, Technology Production & Use, Industry & Innovation, Data Governance	2022	External Knowledge	https://cifar.ca/ai/key-strategic-priorities/#ai-for-health
F50	CIFAR	AI for Energy and the Environment	Program	AI Research, Technology Production & Use, Industry & Innovation	2022	External Knowledge	https://cifar.ca/ai/key-strategic-priorities/#ai-for-energy-and-the-environment
F51	CIFAR	AI Commercialization	Program	Technology Production & Use, Industry & Innovation	2022	External Knowledge	https://cifar.ca/ai/key-strategic-priorities/#ai-commercialization
F52	Parliament of Canada (House of Commons) - Standing Committee on Access to	Meetings: Use and Impact of Facial Recognition Technology (September 21 & 26, 2022),	Policy	Technology Production & Use, Public Administration	2022	External Knowledge	https://www.ourcommons.ca/Committees/en/ETHI/Meetings

	Information, Privacy and Ethics (ETHI)	Device Investigation Tools Used by the RCMP (September 28, 2022)					
F53	Shared Services Canada	Artificial Intelligence Program	Program	AI Education & Training, Technology Production & Use, Public Administration	2022	External Knowledge	https://www.canada.ca/en/shared-services/corporate/artificial-intelligence.html
ON 01	Government of Ontario	Artificial Intelligence (AI) Guidance	Ethics Statement	Technology Production & Use	2021	External Knowledge	https://www.ontario.ca/page/artificial-intelligence-ai-guidance
ON 02	Ontario Ministry of Government and Consumer Services	Strengthening Privacy Protection in Ontario	Policy	Data Governance	2021	ontario.ca search	https://www.ontario.ca/page/strengthening-privacy-protection-ontario
ON 03	Ontario Ministry of Economic Development, Job Creation and Trade	Digital Export Market Development Initiative	Program	Industry & Innovation	2020	ontario.ca search	https://www.ontario.ca/tradecalendar/search?from=2020-04-01&sort=asc&to=2023-03-19&query=artificial%20intelligence
ON 04	Government of Ontario - Expert Panel on Intellectual Property	Report: Intellectual Property in Ontario's Innovation Ecosystem	Policy	Industry & Innovation	2019	ontario.ca search	https://www.ontario.ca/document/report-intellectual-property-in-ontarios-innovation-

							ecosystem
ON 05	Ontario Ministry of the Attorney General	Putting Justice Within Reach: The Foundation for User-Focused Justice in Ontario	Strategic Plan	Public Administration	2017	ontario.ca search	https://www.ontario.ca/page/putting-justice-within-reach-plan-user-focused-justice-ontario
ON 06	Ontario Treasury Board Secretariat - Ontario Digital Service	Ontario's Digital and Data Strategy	Strategic Plan	AI Education & Training, Data Governance, Technology Production & Use, Industry & Innovation, Digital Infrastructure, Social & Workforce, Public Administration	2021	ontario.ca search	https://www.ontario.ca/page/building-digital-ontario
ON 07	Ontario Treasury Board Secretariat - Ontario Digital Service	Digital Service Standard	Standard	Data Governance, Technology Production & Use, Public Administration	2021	ontario.ca search	https://www.ontario.ca/page/digital-service-standard
ON 08	Government of Ontario	Digital and Data Directive	Policy	Data Governance, Technology Production & Use, Public Administration	2021	Followed link from the Digital Service Standard page	https://www.ontario.ca/page/ontarios-digital-and-data-directive-2021
ON 09	Ontario Ministry of Labour, Training, and Skills Development	The Future of Work in Ontario	Policy	Social & Workforce	2021	ontario.ca search	https://www.ontario.ca/document/future-work-ontario

ON 10	Ontario Ministry of Energy	Smart Grid Fund	Program	Industry & Innovation	2019	ontario.ca search	https://www.ontario.ca/document/projects-funded-smart-grid-fund
ON 11	Ontario Ministry of Transportation	Automated Vehicle Pilot Program	Program	Technology Production & Use, Industry & Innovation	2019	ontario.ca search	https://www.ontario.ca/page/automated-vehicle-pilot-program
ON 12	Ontario Ministry of Transportation	Cooperative Truck Platooning Pilot Program	Program	Technology Production & Use, Industry & Innovation	2022	ontario.ca search	https://www.ontario.ca/page/cooperative-truck-platooning-pilot-program
ON 13	Ontario Ministry of Economic Development, Job Creation and Trade	Driving Prosperity: The Future of Ontario's Automative Sector	Strategic Plan	Technology Production & Use, Industry & Innovation, Social & Workforce	2019	ontario.ca search	https://www.ontario.ca/page/driving-prosperity-future-ontarios-automotive-sector
ON 14	Ontario Treasury Board Secretariat - Ontario Digital Service	Digital and Data Innovation Fellowship Program	Program	Public Administration	2021	Link in Ontario's Digital and Data Strategy	https://www.ontario.ca/page/digital-and-data-innovation-fellowship-program
QC 01	Ministry of Economy and Innovation of Quebec	Forum IA Québec	Program	Industry & Innovation	2019	External Knowledge	https://forumia.quebec/en/
QC 02	Government of Quebec	International Centre of Expertise in Montréal for the Advancement of Artificial	Program	AI Research, Industry & Innovation	2020	canada.ca advanced search	https://ceimia.org/

		Intelligence					
QC 03	Government of Quebec	Permanent immigration pilot program for workers in the artificial intelligence, information technologies and visual effects sectors	Program	Industry & Innovation, Social & Workforce	2021	Google search: "site:quebec.ca "artificial intelligence""	https://www.quebec.ca/en/immigration/immigration-programs/artificial-intelligence
QC 04	Government of Quebec	Quebec Research and Innovation Strategy 2017-2022	Strategic Plan	AI Research, Industry & Innovation, Social & Workforce	2017	Google search: "site:quebec.ca "artificial intelligence""	https://numerique.banq.qc.ca/patrimoine/details/52327/2977697
QC 05	Government of Quebec	Stratégie d'intégration de l'intelligence artificielle dans l'administration publique 2021-2026	Strategic Plan	Technology Production & Use, Public Administration	2021	Google search: "site:quebec.ca "artificial intelligence""	https://cdn-contenu.quebec.ca/cdn-contenu/gouvernement/SCT/vitrine_numeriQc/strategie_IA/Strat_IA_2019_2023.pdf?1624995492
QC 06	Université de Montréal et al.	Montréal Declaration for a Responsible Development of Artificial Intelligence	Ethics Statement	AI Research, Industry & Innovation, Technology Production & Use, Data Governance, Digital Infrastructure, Social & Workforce	2018	Google search: "site:quebec.ca "artificial intelligence""	https://declarationmontreal-iaresponsable.com/
QC 07	Government of Quebec -	Le Québec et l'UNESCO	Policy	AI Research, Technology	2022	External Knowledge	https://cdn-contenu.que

	Ministère des Relations internationales et de la Francophonie (Direction des organisations et des forums internationaux)	une vision, une communauté, des priorités pour 2022-2023		Production & Use			bec.ca/cdn-contenu/adm/min/relations-internationales/publications-adm/autres-publications/BR-brochure-Quebec-UNESCO-2022-2023-MRIF.pdf?1663877152
QC 08	Government of Quebec - Ministère de l'Économie et de l'Innovation	Stratégie québécoise de recherche et d'investissement en innovation 2022-2027	Strategic Plan	AI Research, Industry & Innovation, Social & Workforce	2022	External Knowledge	https://cdn-contenu.quebec.ca/cdn-contenu/adm/min/economie/publications-adm/politique/PO_SQRI2_2022-2027_MEI.pdf?1658330226
QC 09	Government of Quebec	Plan Québécois pour la valorisation des minéraux critiques et stratégiques 2020-2025	Strategic Plan	Industry & Innovation, Technology Production & Use, Data Governance, Digital Infrastructure	2020	External Knowledge	https://cdn-contenu.quebec.ca/cdn-contenu/adm/min/energie-ressources-naturelles/publications-adm/plan-strategique/PL_valorisation_mineraux_critiques_strategiques.pdf?1618857094
QC	Government	Intelligence	Program	Industry &	2022	External	https://www

10	of Quebec	numérique en éducation		Innovation, Technology Production & Use, Data Governance, Digital Infrastructure, Social & Workforce, Public Administration		Knowledge	.quebec.ca/education/intelligence-numerique-education;
QC 11	Government of Quebec	Création d'un comité d'experts en matière de cybersécurité et de numérique	Policy	Industry & Innovation, Technology Production & Use, Data Governance, Public Administration	2022	External Knowledge	https://www.quebec.ca/nouvelles/actualites/details/creation-dun-comite-dexperts-en-matiere-de-cybersecurite-et-de-numerique-41740
QC 12	National Assembly of Quebec	Bill 64: An Act to modernize legislative provisions as regards the protection of personal information	Policy	Technology Production & Use, Industry & Innovation, Data Governance, Public Administration	2020	External Knowledge	https://www.assnat.qc.ca/en/travaux-parlementaires/projets-loi/projet-loi-64-42-1.html
AB 01	Government of Alberta	Major Innovation Fund	Program	Industry & Innovation	2018	alberta.ca search	https://www.alberta.ca/major-innovation-fund.aspx
AB 02	Government of Alberta	Alberta's 20-Year Strategic Capital Plan	Strategic Plan	Industry & Innovation	2021	alberta.ca search	https://www.alberta.ca/building-forward-albertas-20-year-strategic-capital-plan.aspx

AB 03	Government of Alberta	Micro-credential Pilot Program	Program	AI Education & Training, Social & Workforce	2021	alberta.ca search	https://www.alberta.ca/new-micro-credential-learning-opportunities.aspx
AB 04	Government of Alberta	Alberta Research and Innovation Framework	Strategic Plan	AI Research, Industry & Innovation	2017	alberta.ca search	https://open.alberta.ca/publications/9781460136126
AB 05	Government of Alberta	Alberta Technology and Innovation Strategy	Strategic Plan	AI Research, AI Education & Training, Industry & Innovation, Social & Workforce	2022	Link in Alberta Major Innovation Fund	https://www.alberta.ca/alberta-technology-and-innovation-strategy.aspx

Appendix 3B: Breakdown of Initiative Data

	Federal	Provincial	Total
Year of Origin			
2017	6	3	**9**
2018	12	2	**13**
2019	10	5	**16**
2020	10	4	**14**
2021	7	11	**18**
2022	8	6	**14**
Intervention Type			
Policy	13	7	**20**
Program	25	11	**36**
Strategic Plan	8	10	**18**
Standard	3	1	**4**
Ethics Statement	4	2	**6**
Governance Area			
AI Research	21	7	**28**
AI Education & Training	7	3	**10**
Technology Production & Use	26	14	**40**
Industry & Innovation	29	21	**50**
Data Governance	11	9	**20**
Digital Infrastructure	2	4	**6**
Social & Workforce	9	10	**19**
Public Administration	18	9	**27**

Appendix 4A: Dataset

This dataset contains the full inventory of 610 topics that were aggregated from across the 20 interviews we conducted during the primary data analysis phase of our study (February 2023 – July 2023). Counts of the frequency with which topics were perceived by participants across the 20 interviews are included for analytical dimensions 1, 3, and 6-11. Topics in analytical dimensions 1, 3, and 6-11 contain counts of the frequency with which aggregate topics emerged across each of the 20 interviews. Topics in analytical dimensions 2, 4, and 5 contain categories instead of frequency counts: the topics in these dimensions represent every unique actor, resource, and network that emerged over the course of the 20 interviews instead of aggregate topics.

Abbreviations used in column titles:

LEAD: Interviews with leaders of public sector AI governance initiatives.

SME-PS: Interviews with subject matter experts employed in the private sector.

SME-CS: Interviews with subject matters experts employed in the academic or civil sectors.

ID#	Analytical Dimensions	Topics	Categories	Count (TOTAL)	Count (LEAD)	Count (SME-PS)	Count (SME-CS)
1.1	**1. Interview Context**	Regulatory development	N/A	**11**	1	6	4
1.2		Organizational AI governance & enablement	N/A	**8**	1	7	
1.3		Standards development	N/A	**6**	2	1	3
1.4		Governance gaps & needs assessments	N/A	**4**	2	2	
1.5		Data governance mechanisms	N/A	**4**	1	1	2
1.6		Regulatory compliance	N/A	**4**		4	
1.7		National AI governance frameworks	N/A	**4**	1	2	1
1.8		Ecosystem development	N/A	**3**	3		
1.9		Adaptability of governance frameworks	N/A	**3**	1	2	
1.10		Procurement guidance	N/A	**3**	2		1

1.11		Incentive structures	N/A	**3**		1	2
1.12		Scalability	N/A	**2**	2		
1.13		Global leadership	N/A	**2**	2		
1.14		International alignment	N/A	**2**	2		
1.15		Guardrails on AI	N/A	**2**	2		
1.16		Public trust	N/A	**2**	2		
1.17		Public accountability & oversight	N/A	**2**	2		
1.18		Risk management	N/A	**2**	2		
1.19		Proactive governance	N/A	**2**	1	1	
1.20		Sectoral guidance	N/A	**2**		2	
1.21		Fundamental, applied, & policy research	N/A	**2**		1	1
1.22		AI supply chains	N/A	**2**		1	1
1.23		Financing mechanisms	N/A	**2**	1	1	
1.24		Long-term planning	N/A	**1**	1		
1.25		Talent development	N/A	**1**	1		
1.26		Procedural fairness	N/A	**1**	1		
1.27		AI registry development	N/A	**1**	1		
1.28		Interpretive clarity	N/A	**1**	1		
1.29		Protection of public interests	N/A	**1**	1		
1.30		Persuasiveness of policy	N/A	**1**		1	
1.31		AI maturity	N/A	**1**		1	
1.32		Public-private AI development partnerships	N/A	**1**		1	
1.33		Private AI governance services	N/A	**1**			1
1.34		Government priorities & values	N/A	**1**			1
2.1	**2. Actors**	AI-providers	Type of socio-technical actor	N/A	N/A	N/A	N/A
2.2		AI-users	Type of socio-technical actor	N/A	N/A	N/A	N/A
2.3		AI-applying actors	Type of socio-technical actor	N/A	N/A	N/A	N/A
2.4		AI-impacted actors	Type of socio-technical actor	N/A	N/A	N/A	N/A
2.5		AI rule-makers	Type of socio-technical actor	N/A	N/A	N/A	N/A

2.6		AI rule-takers	Type of socio-technical actor	N/A	N/A	N/A	N/A
2.7		Vulnerable groups	Type of socio-technical actor	N/A	N/A	N/A	N/A
2.8		AI developers	Type of socio-technical actor	N/A	N/A	N/A	N/A
2.9		AI users	Type of socio-technical actor	N/A	N/A	N/A	N/A
2.10		End users	Type of socio-technical actor	N/A	N/A	N/A	N/A
2.11		Beneficiaries of AI	Type of socio-technical actor	N/A	N/A	N/A	N/A
2.12		Public AI governance services providers	Type of socio-technical actor	N/A	N/A	N/A	N/A
2.13		Private AI governance services providers	Type of socio-technical actor	N/A	N/A	N/A	N/A
2.14		Actors involved in AI governance initiatives	Type of socio-technical actor	N/A	N/A	N/A	N/A
2.15		Actors affected by AI governance initiatives	Type of socio-technical actor	N/A	N/A	N/A	N/A
2.16		Privacy professionals	Type of socio-technical actor	N/A	N/A	N/A	N/A
2.17		Legal professionals	Type of socio-technical actor	N/A	N/A	N/A	N/A
2.18		AI researchers	Type of socio-technical actor	N/A	N/A	N/A	N/A
2.19		AI developers	Type of socio-technical actor	N/A	N/A	N/A	N/A
2.20		Activists	Type of socio-technical actor	N/A	N/A	N/A	N/A
2.21		Lobbyists	Type of socio-technical actor	N/A	N/A	N/A	N/A
2.22		Policymakers	Type of socio-technical actor	N/A	N/A	N/A	N/A
2.23		Policy advisors	Type of socio-technical actor	N/A	N/A	N/A	N/A
2.24		Tech industry	Type of organizational actor	N/A	N/A	N/A	N/A
2.25		Academic sector	Type of organizational actor	N/A	N/A	N/A	N/A
2.26		Civil sector	Type of organizational actor	N/A	N/A	N/A	N/A
2.27		Canadian tech industry	Type of organizational actor	N/A	N/A	N/A	N/A
2.28		Canadian manufacturing industry	Type of organizational actor	N/A	N/A	N/A	N/A
2.29		Canadian service industries	Type of organizational actor	N/A	N/A	N/A	N/A
2.30		Public sector	Type of organizational actor	N/A	N/A	N/A	N/A

2.31		Private sector	Type of organizational actor	N/A	N/A	N/A	N/A
2.32		Nonprofit sector	Type of organizational actor	N/A	N/A	N/A	N/A
2.33		Internal experts	Type of organizational actor	N/A	N/A	N/A	N/A
2.34		External experts	Type of organizational actor	N/A	N/A	N/A	N/A
2.35		Government departments	Type of organizational actor	N/A	N/A	N/A	N/A
2.36		Federal governments	Type of organizational actor	N/A	N/A	N/A	N/A
2.37		Provincial governments	Type of organizational actor	N/A	N/A	N/A	N/A
2.38		Territorial governments	Type of organizational actor	N/A	N/A	N/A	N/A
2.39		Regional governments	Type of organizational actor	N/A	N/A	N/A	N/A
2.40		Municipal governments	Type of organizational actor	N/A	N/A	N/A	N/A
2.41		National AI Institutes	Type of organizational actor	N/A	N/A	N/A	N/A
2.42		Industry partners	Type of organizational actor	N/A	N/A	N/A	N/A
2.43		AI startups	Type of organizational actor	N/A	N/A	N/A	N/A
2.44		AI SMEs	Type of organizational actor	N/A	N/A	N/A	N/A
2.45		AI enterprises	Type of organizational actor	N/A	N/A	N/A	N/A
2.46		Large enterprises	Type of organizational actor	N/A	N/A	N/A	N/A
2.47		Residents	Type of organizational actor	N/A	N/A	N/A	N/A
2.48		Citizens	Type of organizational actor	N/A	N/A	N/A	N/A
2.49		Clients	Type of organizational actor	N/A	N/A	N/A	N/A
2.50		Research institutes	Type of organizational actor	N/A	N/A	N/A	N/A
2.51		Policy shops	Type of organizational actor	N/A	N/A	N/A	N/A
2.52		Privacy commissioners	Type of organizational actor	N/A	N/A	N/A	N/A
2.53		Indigenous peoples	Type of organizational actor	N/A	N/A	N/A	N/A
2.54		Indigenous organizations	Type of organizational actor	N/A	N/A	N/A	N/A
2.55		Academic partners	Type of organizational actor	N/A	N/A	N/A	N/A

2.56		Civil society partners	Type of organizational actor	N/A	N/A	N/A	N/A
2.57		International organizations	Type of organizational actor	N/A	N/A	N/A	N/A
2.58		Communities of practice	Type of organizational actor	N/A	N/A	N/A	N/A
2.59		Legal counsels	Type of organizational actor	N/A	N/A	N/A	N/A
2.60		Publics	Type of organizational actor	N/A	N/A	N/A	N/A
2.61		Intergovernmental councils	Type of organizational actor	N/A	N/A	N/A	N/A
2.62		Large consulting firms	Type of organizational actor	N/A	N/A	N/A	N/A
2.63		Small consulting firms	Type of organizational actor	N/A	N/A	N/A	N/A
2.64		Professional associations	Type of organizational actor	N/A	N/A	N/A	N/A
2.65		Labour unions	Type of organizational actor	N/A	N/A	N/A	N/A
2.66		Advocacy organizations	Type of organizational actor	N/A	N/A	N/A	N/A
2.67		Educational institutions	Type of organizational actor	N/A	N/A	N/A	N/A
2.68		CIFAR	Instance	N/A	N/A	N/A	N/A
2.69		Vector Institute	Instance	N/A	N/A	N/A	N/A
2.70		Mila	Instance	N/A	N/A	N/A	N/A
2.71		Amii	Instance	N/A	N/A	N/A	N/A
2.72		Global Affairs Canada (GAC)	Instance	N/A	N/A	N/A	N/A
2.73		Treasury Board of Canada Secretariat (TBS)	Instance	N/A	N/A	N/A	N/A
2.74		Government of Canada (GC)	Instance	N/A	N/A	N/A	N/A
2.75		Government of Ontario	Instance	N/A	N/A	N/A	N/A
2.76		Government of Quebec	Instance	N/A	N/A	N/A	N/A
2.77		Government of Alberta	Instance	N/A	N/A	N/A	N/A
2.78		Global Innovation Clusters	Instance	N/A	N/A	N/A	N/A
2.79		Standards Council of Canada (SCC)	Instance	N/A	N/A	N/A	N/A
2.80		CIO Strategy Council	Instance	N/A	N/A	N/A	N/A
2.81		US National Institute of Standards and Technology (NIST)	Instance	N/A	N/A	N/A	N/A
2.82		Canada Revenue Agency (CRA)	Instance	N/A	N/A	N/A	N/A

2.83		Innovation, Science and Economic Development Canada (ISED)	Instance	N/A	N/A	N/A	N/A
2.84		Employment and Social Development Canada (ESDC)	Instance	N/A	N/A	N/A	N/A
2.85		Immigration, Refugees, and Citizenship Canada (IRCC)	Instance	N/A	N/A	N/A	N/A
2.86		City of London (Ontario)	Instance	N/A	N/A	N/A	N/A
2.87		City of Toronto	Instance	N/A	N/A	N/A	N/A
2.88		Toronto Police Services	Instance	N/A	N/A	N/A	N/A
2.89		Clearview AI	Instance	N/A	N/A	N/A	N/A
2.90		City of Montreal	Instance	N/A	N/A	N/A	N/A
2.91		The Digital Nations	Instance	N/A	N/A	N/A	N/A
2.92		United States	Instance	N/A	N/A	N/A	N/A
2.93		United Kingdom	Instance	N/A	N/A	N/A	N/A
2.94		OECD	Instance	N/A	N/A	N/A	N/A
2.95		UNESCO	Instance	N/A	N/A	N/A	N/A
2.96		Council of Europe	Instance	N/A	N/A	N/A	N/A
2.97		European Union (EU)	Instance	N/A	N/A	N/A	N/A
2.98		NATO	Instance	N/A	N/A	N/A	N/A
2.99		Canadians	Instance	N/A	N/A	N/A	N/A
2.100		Ontarians	Instance	N/A	N/A	N/A	N/A
2.101		AI & Data Governance Standardization Collaborative	Instance	N/A	N/A	N/A	N/A
2.102		Global Partnership on AI (GPAI)	Instance	N/A	N/A	N/A	N/A
2.103		Responsible AI Institute (RAII)	Instance	N/A	N/A	N/A	N/A
2.104		AI & Data Commissioner	Instance	N/A	N/A	N/A	N/A
2.105		Parliament of Canada	Instance	N/A	N/A	N/A	N/A
2.106		Standing Committee on Industry & Technology (INDU)	Instance	N/A	N/A	N/A	N/A
2.107		Standing Committee on Access to Information, Privacy, & Ethics (ETHI)	Instance	N/A	N/A	N/A	N/A
2.108		Privacy Commissioner of Canada	Instance	N/A	N/A	N/A	N/A
2.109		Information and Privacy Commissioner of Ontario	Instance	N/A	N/A	N/A	N/A
2.110		Government of British Columbia	Instance	N/A	N/A	N/A	N/A
2.111		Advisory Council on AI	Instance	N/A	N/A	N/A	N/A

2.112		Cities for Digital Rights	Instance	N/A	N/A	N/A	N/A
2.113		Element AI	Instance	N/A	N/A	N/A	N/A
2.114		Armilla AI	Instance	N/A	N/A	N/A	N/A
2.115		EY	Instance	N/A	N/A	N/A	N/A
2.116		Deloitte	Instance	N/A	N/A	N/A	N/A
2.117		KPMG	Instance	N/A	N/A	N/A	N/A
2.118		Google	Instance	N/A	N/A	N/A	N/A
2.119		Microsoft	Instance	N/A	N/A	N/A	N/A
2.120		Amazon	Instance	N/A	N/A	N/A	N/A
3.1	**3. Benefits, Risks, & Harms**	Public mistrust	N/A	**7**	4	2	1
3.2		Risk mitigation	N/A	**6**	4	1	1
3.3		Improved efficiency	N/A	**6**	4	1	1
3.4		Improved service quality	N/A	**6**	4	1	1
3.5		Epistemic risks of misinterpreting "AI"	N/A	**6**	1	3	2
3.6		Financial & economic gains	N/A	**5**	1	3	1
3.7		Harms to worker wellbeing & workplace quality	N/A	**5**	1	4	
3.8		Unfair decision outcomes	N/A	**5**	3	1	1
3.9		Tensions between responsible AI & profitable AI	N/A	**4**		3	1
3.10		Rights-based harms	N/A	**4**	2	1	1
3.11		Reinforcement of structural injustices	N/A	**4**	2	1	1
3.12		Application-specific performance improvements	N/A	**3**	1	1	1
3.13		Safety provided by good governance frameworks	N/A	**3**		2	1
3.14		Harms caused by institutional incapacities	N/A	**3**	1		2
3.15		Narrow scope of impact assessment & regulation	N/A	**3**		1	2
3.16		Workforce development	N/A	**2**	1	1	
3.17		Creation of AI talent	N/A	**2**	1	1	
3.18		Strong public accountability & oversight	N/A	**2**	2		
3.19		Robustness of governance frameworks	N/A	**2**	1	1	

3.20		Protection of public interests	N/A	**2**	1	1	
3.21		Corporate social responsibility & benefits to brand	N/A	**2**		1	1
3.22		Reputational risks	N/A	**2**	2		
3.23		Physical harms	N/A	**2**	1	1	
3.24		Unfair & inequitable regulatory enforcement	N/A	**2**	1		1
3.25		Automation bias	N/A	**2**	2		
3.26		Systemic political & economic harms	N/A	**2**	1		1
3.27		Negligence of edge cases	N/A	**2**	1	1	
3.28		Privacy violations	N/A	**2**	2		
3.29		Harms to innovation caused by strict regulation	N/A	**2**		1	1
3.30		Situatedness of harms	N/A	**2**		2	
3.31		Costs of compliance	N/A	**2**			2
3.32		Misalignments with public interests & values	N/A	**2**		1	1
3.33		Labor displacement	N/A	**2**		2	
3.34		Widening of socio-economic divides	N/A	**2**		1	1
3.35		Exclusion of marginalized groups from governance activities	N/A	**2**		2	
3.36		Distribution of harms & accountabilities in AI supply chains	N/A	**2**		2	
3.37		Industry consolidation & protectionism	N/A	**2**		1	1
3.38		Investment in Canada	N/A	**1**	1		
3.39		Pre-competitive collaboration	N/A	**1**	1		
3.40		Explainability of decisions	N/A	**1**	1		
3.41		Client comfort	N/A	**1**	1		
3.42		Improved global leadership	N/A	**1**	1		
3.43		Improved public trust	N/A	**1**	1		
3.44		Cost savings	N/A	**1**	1		
3.45		Fair automated decisions	N/A	**1**	1		
3.46		Shared understandings of problems & solutions	N/A	**1**	1		
3.47		Timeliness of service	N/A	**1**	1		
3.48		Flexibility of standards	N/A	**1**			1

3.49		Agile governance frameworks	N/A	**1**			1
3.50		Fulfillment of government mandates	N/A	**1**		1	
3.51		Improved value alignments between organizations & applications	N/A	**1**		1	
3.52		Private sector benefits from regulatory & standards leadership	N/A	**1**			1
3.53		Improved business & national competitiveness	N/A	**1**		1	
3.54		Freedom to innovate	N/A	**1**		1	
3.55		Improved scalability of AI businesses & systems	N/A	**1**		1	
3.56		Upstream benefits in AI supply chains (hardware & infrastructure)	N/A	**1**		1	
3.57		Operational risks	N/A	**1**	1		
3.58		Competition for scarce resources	N/A	**1**	1		
3.59		Disruptions to AI ecosystem equilibrium	N/A	**1**	1		
3.60		Psychological harms	N/A	**1**	1		
3.61		Complexification of legacy risks	N/A	**1**	1		
3.62		Legal risks	N/A	**1**	1		
3.63		Public perception of tokenization in consultation processes	N/A	**1**	1		
3.64		Intensification of mass surveillance	N/A	**1**	1		
3.65		Environmental harms	N/A	**1**		1	
3.66		Risks of regulatory experimentation	N/A	**1**			1
3.67		Disproportionate costs to small & medium enterprises	N/A	**1**			1
3.68		Harms caused by AI vs. harms caused by AI governance	N/A	**1**			1
3.69		Harms caused by outsourcing practices (extragovernmental & transnational)	N/A	**1**			1
3.70		Normalization of harmful governance frameworks	N/A	**1**		1	
3.71		IP theft involved in generative AI training	N/A	**1**		1	
3.72		Risks of deferring benefits of AI & non-innovation	N/A	**1**		1	
4.1	**4. Resources**	Political approvals	Cultural resources	N/A	N/A	N/A	N/A
4.2		AI-aware organizational cultures	Cultural resources	N/A	N/A	N/A	N/A
4.3		Positive public perceptions of AI & AI governance	Cultural resources	N/A	N/A	N/A	N/A

4.4		Political buy-in from upper management	Cultural resources	N/A	N/A	N/A	N/A
4.5		AI governance cultures	Cultural resources	N/A	N/A	N/A	N/A
4.6		Ethical AI cultures	Cultural resources	N/A	N/A	N/A	N/A
4.7		Training data	Data & computational resources	N/A	N/A	N/A	N/A
4.8		Testing data	Data & computational resources	N/A	N/A	N/A	N/A
4.9		High quality data	Data & computational resources	N/A	N/A	N/A	N/A
4.10		Machine learning models	Data & computational resources	N/A	N/A	N/A	N/A
4.11		Compute resources	Data & computational resources	N/A	N/A	N/A	N/A
4.12		Cloud services	Data & computational resources	N/A	N/A	N/A	N/A
4.13		Local machines	Data & computational resources	N/A	N/A	N/A	N/A
4.14		Domain-specific training data	Data & computational resources	N/A	N/A	N/A	N/A
4.15		Cleaned data	Data & computational resources	N/A	N/A	N/A	N/A
4.16		Labelled data	Data & computational resources	N/A	N/A	N/A	N/A
4.17		Reliable data pipelines	Data & computational resources	N/A	N/A	N/A	N/A
4.18		Predictive models	Data & computational resources	N/A	N/A	N/A	N/A
4.19		Generative models	Data & computational resources	N/A	N/A	N/A	N/A
4.20		Productionized models	Data & computational resources	N/A	N/A	N/A	N/A
4.21		Digital infrastructures	Data & computational resources	N/A	N/A	N/A	N/A
4.22		Technology stacks	Data & computational resources	N/A	N/A	N/A	N/A
4.23		AI supply chains	Data & computational resources	N/A	N/A	N/A	N/A
4.24		Finances	Financial resources	N/A	N/A	N/A	N/A
4.25		Budgets	Financial resources	N/A	N/A	N/A	N/A
4.26		Public funding	Financial resources	N/A	N/A	N/A	N/A
4.27		Private funding	Financial resources	N/A	N/A	N/A	N/A
4.28		Funding mechanisms	Financial resources	N/A	N/A	N/A	N/A
4.29		Capital	Financial resources	N/A	N/A	N/A	N/A

4.30		Public investments	Financial resources	N/A	N/A	N/A	N/A
4.31		Private investments	Financial resources	N/A	N/A	N/A	N/A
4.32		Profits	Financial resources	N/A	N/A	N/A	N/A
4.33		Commercialized models	Financial resources	N/A	N/A	N/A	N/A
4.34		Funding for compliance programs	Financial resources	N/A	N/A	N/A	N/A
4.35		Business models	Financial resources	N/A	N/A	N/A	N/A
4.36		Funding for participating in policy & standards development	Financial resources	N/A	N/A	N/A	N/A
4.37		Funding for marginalized groups	Financial resources	N/A	N/A	N/A	N/A
4.38		Public funding of private development & governance practices	Financial resources	N/A	N/A	N/A	N/A
4.39		Funding for civil society organizations	Financial resources	N/A	N/A	N/A	N/A
4.40		Funding for AI development & governance in SMEs	Financial resources	N/A	N/A	N/A	N/A
4.41		Funding for AI development & governance in large enterprises	Financial resources	N/A	N/A	N/A	N/A
4.42		Knowledge of AI development & use	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.43		Knowledge of AI ethics	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.44		Knowledge of critical analysis methods	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.45		Research expertise	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.46		Ethics boards	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.47		Review boards	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.48		Educational resources	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.49		Training programs	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.50		Talent	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.51		Technical expertise	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.52		Business expertise	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.53		External expertise	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.54		Consulting contracts	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.55		Human resources	Knowledge & cognitive resources	N/A	N/A	N/A	N/A

4.56		Networks of experts	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.57		Multidisciplinary knowledge	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.58		Interdisciplinary knowledge	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.59		Knowledge of external stakeholders	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.60		Technical knowledge	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.61		Legal knowledge	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.62		Policy knowledge	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.63		Explanatory webpages	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.64		Databases of records & documentation	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.65		Bilateral engagements	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.66		Vendors	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.67		Procurement relationships	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.68		Cybersecurity expertise	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.69		Privacy expertise	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.70		IT management expertise	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.71		Public awareness of AI impacts	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.72		AI literacies	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.73		Data literacies	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.74		Digital literacies	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.75		Advisory & decision-making bodies	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.76		Sectoral expertise	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.77		Knowledge of sectoral needs	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.78		Awareness of AI use cases	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.79		Data preparation expertise	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.80		Data science expertise	Knowledge & cognitive resources	N/A	N/A	N/A	N/A

4.81		Academic partnerships	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.82		Deployment expertise	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.83		Applied AI research expertise	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.84		AI commercialization expertise	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.85		Knowledge of privacy governance	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.86		Knowledge of risk governance	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.87		Knowledge of data governance	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.88		Policy expertise	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.89		Legal expertise	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.90		Knowledge of regulatory environment	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.91		Chief AI Officers	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.92		Corporate executives	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.93		External expertise	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.94		Internal expertise	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.95		Agile software development teams	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.96		Agile policy development teams	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.97		Paper-based knowledge resources	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.98		Digitization expertise	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.99		Communication tools & supports	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.100		Communities of practice	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.101		Hubs of expertise	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.102		Working groups	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.103		Education & training for marginalized groups	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.104		Education & training for organizations	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.105		Translated knowledge between departments/organizations	Knowledge & cognitive resources	N/A	N/A	N/A	N/A

4.106		Up-to-date technical expertise	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.107		Awareness of sector-specific AI impacts	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.108		Knowledge of AI implementation best practices	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.109		Implementation expertise	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.110		AI ethics expertise	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.111		Thought leadership	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.112		Policy teams	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.113		Knowledge of human rights	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.114		Professional bodies of knowledge	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.115		Clear KPIs & metrics for AI system performance	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.116		Dashboards for managing AI systems	Knowledge & cognitive resources	N/A	N/A	N/A	N/A
4.117		Memoranda of Understanding (MOUs)	Policy & legal resources	N/A	N/A	N/A	N/A
4.118		Data governance protocols	Policy & legal resources	N/A	N/A	N/A	N/A
4.119		Strategies	Policy & legal resources	N/A	N/A	N/A	N/A
4.120		Guidance documents	Policy & legal resources	N/A	N/A	N/A	N/A
4.121		Policies	Policy & legal resources	N/A	N/A	N/A	N/A
4.122		Directives	Policy & legal resources	N/A	N/A	N/A	N/A
4.123		Compliance supports	Policy & legal resources	N/A	N/A	N/A	N/A
4.124		Administrative authorizations	Policy & legal resources	N/A	N/A	N/A	N/A
4.125		Procurement policies	Policy & legal resources	N/A	N/A	N/A	N/A
4.126		Policy models & best practices	Policy & legal resources	N/A	N/A	N/A	N/A
4.127		Technical standards	Policy & legal resources	N/A	N/A	N/A	N/A
4.128		Industry standards	Policy & legal resources	N/A	N/A	N/A	N/A
4.129		Auditing & assessment frameworks	Policy & legal resources	N/A	N/A	N/A	N/A
4.130		Algorithmic impact assessments	Policy & legal resources	N/A	N/A	N/A	N/A

4.131		Privacy impact assessments	Policy & legal resources	N/A	N/A	N/A	N/A
4.132		Data governance frameworks	Policy & legal resources	N/A	N/A	N/A	N/A
4.133		Legal teams	Policy & legal resources	N/A	N/A	N/A	N/A
4.134		Market assessments	Policy & legal resources	N/A	N/A	N/A	N/A
4.135		Corporate policies	Policy & legal resources	N/A	N/A	N/A	N/A
4.136		Personalized policy solutions	Policy & legal resources	N/A	N/A	N/A	N/A
4.137		AI laws	Policy & legal resources	N/A	N/A	N/A	N/A
4.138		AI regulations	Policy & legal resources	N/A	N/A	N/A	N/A
4.139		Compliance management programs	Policy & legal resources	N/A	N/A	N/A	N/A
4.140		Independent audits	Policy & legal resources	N/A	N/A	N/A	N/A
4.141		Advocacy groups	Policy & legal resources	N/A	N/A	N/A	N/A
4.142		Human rights	Policy & legal resources	N/A	N/A	N/A	N/A
5.1	**5. Networks**	CIFAR programs network	Governance networks	N/A	N/A	N/A	N/A
5.2		ISED policymaking network	Governance networks	N/A	N/A	N/A	N/A
5.3		TBS policy network	Governance networks	N/A	N/A	N/A	N/A
5.4		SCC standards network	Governance networks	N/A	N/A	N/A	N/A
5.5		Bill C-27 legislative network	Governance networks	N/A	N/A	N/A	N/A
5.6		National AI institutes networks	Governance networks	N/A	N/A	N/A	N/A
5.7		Centralized vs. decentralized governance networks	Governance networks	N/A	N/A	N/A	N/A
5.8		Edge governance networks	Governance networks	N/A	N/A	N/A	N/A
5.9		Formalized governance networks	Governance networks	N/A	N/A	N/A	N/A
5.10		Informal/semi-formal governance networks	Governance networks	N/A	N/A	N/A	N/A
5.11		Sectoral governance networks	Governance networks	N/A	N/A	N/A	N/A
5.12		Organizational governance networks	Governance networks	N/A	N/A	N/A	N/A
5.13		Legal services networks	Governance networks	N/A	N/A	N/A	N/A

5.14		Policymaking networks	Governance networks	N/A	N/A	N/A	N/A
5.15		Regulatory compliance networks	Governance networks	N/A	N/A	N/A	N/A
5.16		Emerging networks of AI governance services providers	Governance networks	N/A	N/A	N/A	N/A
5.17		Stakeholder consultation networks	Governance networks	N/A	N/A	N/A	N/A
5.18		Public consultation networks	Governance networks	N/A	N/A	N/A	N/A
5.19		Private consultation networks	Governance networks	N/A	N/A	N/A	N/A
5.20		Standards development networks	Governance networks	N/A	N/A	N/A	N/A
5.21		Civic advocacy networks	Governance networks	N/A	N/A	N/A	N/A
5.22		Advisory Council network	Governance networks	N/A	N/A	N/A	N/A
5.23		Academic & research networks	Resource integration networks	N/A	N/A	N/A	N/A
5.24		Intra-departmental knowledge sharing networks	Resource integration networks	N/A	N/A	N/A	N/A
5.25		Inter-departmental knowledge sharing networks	Resource integration networks	N/A	N/A	N/A	N/A
5.26		Inter-governmental knowledge sharing networks	Resource integration networks	N/A	N/A	N/A	N/A
5.27		Inter-sectoral knowledge sharing networks	Resource integration networks	N/A	N/A	N/A	N/A
5.28		International knowledge sharing networks	Resource integration networks	N/A	N/A	N/A	N/A
5.29		Multinational/bilateral knowledge sharing networks	Resource integration networks	N/A	N/A	N/A	N/A
5.30		Networks & platforms for sharing code and data	Resource integration networks	N/A	N/A	N/A	N/A
5.31		Networks of experts	Resource integration networks	N/A	N/A	N/A	N/A
5.32		Networks of AI practitioners	Resource integration networks	N/A	N/A	N/A	N/A
5.33		Social media networks & platforms	Resource integration networks	N/A	N/A	N/A	N/A
5.34		Social networking supports	Resource integration networks	N/A	N/A	N/A	N/A
5.35		Physical spaces for networking events	Resource integration networks	N/A	N/A	N/A	N/A
5.36		Fundraising & funding networks	Resource integration networks	N/A	N/A	N/A	N/A
5.37		Networks of privacy professionals	Resource integration networks	N/A	N/A	N/A	N/A
5.38		AI education & training networks	Resource integration networks	N/A	N/A	N/A	N/A

5.39		Influencer & thought leader networks	Resource integration networks	N/A	N/A	N/A	N/A
5.40		Element AI network	Resource integration networks	N/A	N/A	N/A	N/A
6.1	**6. Evaluations**	Benefits to public realized & risks mitigated	N/A	**6**	3	2	1
6.2		AI systems are aligned with organizational values	N/A	**6**		5	1
6.3		Reduction of stakeholder frictions	N/A	**5**	1	2	2
6.4		Trust of public & other stakeholders secured	N/A	**4**	2	1	1
6.5		International harmonization on AI regulations & standards	N/A	**4**	1	1	2
6.6		Shared meanings established	N/A	**4**	1		3
6.7		Standardized quality/compliance measures achieved	N/A	**3**	1	2	
6.8		High awareness of AI impacts & governance activities	N/A	**3**	1	1	1
6.9		Positive qualitative societal outcomes	N/A	**3**	1	2	
6.10		AI is aligned with national values	N/A	**3**	1	2	
6.11		Involvement of vulnerable groups in AI governance process	N/A	**3**	1	2	
6.12		Governance mechanisms balance generalizability & context-sensitivity	N/A	**3**		1	2
6.13		Interoperable AI markets & systems	N/A	**3**	1	1	1
6.14		Quality of AI talent & workforce	N/A	**2**	1	1	
6.15		Quality & scale of ecosystem collaborations & partnerships	N/A	**2**	1		1
6.16		Canadian impact on global AI governance practices	N/A	**2**	1	1	
6.17		Strategic & operational goals achieved	N/A	**2**	1	1	
6.18		Inter-departmental & inter-sectoral harmonization	N/A	**2**	1		1
6.19		Ease of interpreting & applying policy tools	N/A	**2**	1		1
6.20		Fairness & transparency in organizational decision-making	N/A	**2**	2		
6.21		Guardrails effectively prevent harms	N/A	**2**	2		
6.22		Compliance with relevant privacy laws	N/A	**2**	1	1	
6.23		Effective knowledge-sharing between stakeholders	N/A	**2**	1	1	
6.24		Implementation of effective socio-technical & ethical standards	N/A	**2**	1	1	

6.25		Freedom to innovate without restrictions	N/A	**2**		1	1
6.26		Valid KPIs/quality measures identified	N/A	**2**		1	1
6.27		Strengthening of organization's reputation & brand	N/A	**2**		1	1
6.28		Reduction of operational risks	N/A	**1**	1		
6.29		Creation of public value	N/A	**1**	1		
6.30		Policies are easy to implement & enforce	N/A	**1**	1		
6.31		Improvements to business & service outcomes	N/A	**1**	1		
6.32		AI is adopted only when necessary (resistance against techno-determinism)	N/A	**1**	1		
6.33		Documentation, disclosure, & recourse for actual harms	N/A	**1**	1		
6.34		Public participation is proactive	N/A	**1**	1		
6.35		Compliance with emerging AI laws & regulations	N/A	**1**		1	
6.36		Balance between business objectives & compliance achieved	N/A	**1**		1	
6.37		Effective organizational management practices	N/A	**1**		1	
6.38		International markets are stable & accessible	N/A	**1**			1
6.39		Competitive nations & businesses	N/A	**1**		1	
6.40		First to market in a product/service category	N/A	**1**		1	
6.41		Organizational interests represented in policy process & outcomes	N/A	**1**		1	
6.42		Efficient public service delivery	N/A	**1**			1
6.43		Profitable AI companies	N/A	**1**			1
6.44		Technological competency successfully performed in social settings	N/A	**1**			1
6.45		AI systems are effectively monitored & evaluated throughout the lifecycle	N/A	**1**		1	
6.46		Cultivation of effective AI governance cultures	N/A	**1**		1	
6.47		Buy-in from senior leadership on organizational AI governance	N/A	**1**		1	
6.48		Governance frameworks are scalable & transferable	N/A	**1**		1	
7.1	**7. Logics**	Maximize benefits & mitigate risks of AI development & use	N/A	**8**	5	3	

7.2		Align Canadian AI governance with international partners	N/A	**6**	2		4
7.3		Achieve balance between public & private interests in AI outcomes	N/A	**6**	1	4	1
7.4		Maximize profit & shareholder value from AI adoption	N/A	**6**		4	2
7.5		Align AI systems with organizational values	N/A	**5**	1	4	
7.6		Facilitate cross-organizational & cross-sectoral collaboration	N/A	**4**	3		1
7.7		Strengthen public/consumer trust in AI applications	N/A	**4**	1	3	
7.8		Ensure AI systems are compliant with relevant laws & standards	N/A	**4**	1	2	1
7.9		Ensure benefits realized & harms prevented for vulnerable groups	N/A	**4**	1	3	
7.10		Improve quality, efficiency, & effectiveness of public service delivery	N/A	**3**	3		
7.11		Ensure AI development & use contributes to economic development	N/A	**3**	1		2
7.12		Align AI systems with Canadian values	N/A	**3**	1	1	1
7.13		Cultivate a trustworthy brand image	N/A	**3**		3	
7.14		Enable greater access to international markets	N/A	**3**		1	2
7.15		Create AI systems that are interoperable across markets & jurisdictions	N/A	**3**		1	2
7.16		Create shared AI ontologies & interpretations	N/A	**3**	2		1
7.17		Ensure diverse interests are represented in policies	N/A	**3**		3	
7.18		Ensure FAT+ principles reflected in AI development & use	N/A	**2**	1	1	
7.19		Align governance frameworks with other departments & jurisdictions	N/A	**2**	1	1	
7.20		Ensure privacy risks are mitigated across AI application areas	N/A	**2**	1		1
7.21		Develop greater AI skills & experience	N/A	**2**		1	1
7.22		Protect democratic institutions from harmful AI impacts	N/A	**2**		1	1
7.23		Effectively use AI applications to full organizational goals & mandates	N/A	**2**	1	1	
7.24		Achieve a competitive advantage	N/A	**2**		1	1
7.25		Set an example for how to practice responsible AI	N/A	**2**	1	1	

7.26		Maintain & strengthen Canada's global leadership in AI	N/A	**1**	1		
7.27		Maintain & strengthen Canada's AI ecosystem	N/A	**1**	1		
7.28		Cultivate cultures of responsible AI governance	N/A	**1**	1		
7.29		Resist hype-based AI implementations	N/A	**1**	1		
7.30		Effectively manage risks of AI procurement processes	N/A	**1**	1		
7.31		Cultivate in-house AI governance capacities	N/A	**1**	1		
7.32		Facilitate effective knowledge-sharing relationships	N/A	**1**	1		
7.33		Facilitate the creation & application of auditing frameworks	N/A	**1**	1		
7.34		Innovate and experiment with minimal constraints	N/A	**1**		1	
7.35		Avoid tensions with international partners & markets	N/A	**1**			1
7.36		Avoid addressing technical debt	N/A	**1**			1
7.37		Improve technological capabilities & capacities	N/A	**1**		1	
7.38		Prevent AI harms throughout supply chains	N/A	**1**		1	
7.39		Mitigate catastrophic/existential risks of AI	N/A	**1**		1	
8.1	**8. Bounds**	Knowledge/expertise resource limitations within organizations	N/A	**10**	5	3	2
8.2		Gaps in stakeholder literacies & awareness	N/A	**8**	2	4	2
8.3		Knowledge/expertise resource limitations within ecosystem	N/A	**6**	2	3	1
8.4		Gaps in knowledge of stakeholder needs & requirements	N/A	**6**	1	4	1
8.5		Financial limitations within organizations	N/A	**5**	4	1	
8.6		Administrative incapacities within organizations	N/A	**4**	3		1
8.7		Interpretive barriers & ontological uncertainties	N/A	**4**	2	1	1
8.8		Gaps in interdisciplinary knowledge & expertise	N/A	**3**		3	
8.9		Epistemic disconnect from AI application contexts	N/A	**3**		2	1
8.10		Knowledge/expertise resource limitations within sectors	N/A	**2**		2	
8.11		Technical debts	N/A	**2**		1	1

8.12		Cultural limitations in data sourcing & documentation practices	N/A	**2**		1	1
8.13		Cultural limitations of in-house capacity-building vs. outsourcing	N/A	**1**	1		
8.14		Administrative incapacities within ecosystem	N/A	**1**	1		
8.15		Cultural limitations of information-seeking & knowledge-sharing behaviors	N/A	**1**	1		
8.16		Limitations in civil society & academia resources	N/A	**1**	1		
8.17		Democracy/agility trade-offs	N/A	**1**	1		
8.18		Cultural limitations of public comfort with AI	N/A	**1**			1
8.19		Unclear mechanisms for public participation in governance activities	N/A	**1**		1	
8.20		AI information overload & limitations on sensemaking	N/A	**1**		1	
8.21		Opacity of automated processes	N/A	**1**			1
8.22		Automation bias	N/A	**1**			1
8.23		Ethical limitations of business models	N/A	**1**		1	
8.24		Unpredictability of AI futures	N/A	**1**		1	
8.25		Tensions with other organizations	N/A	**1**		1	
9.1	**9. Rules**	Requirements set by AI-related laws & regulations	N/A	**9**	2	6	1
9.2		Impact of organizational norms & culture on AI policies	N/A	**7**	2	4	1
9.3		Limitations on scope & jurisdiction of AI policies	N/A	**7**	3	3	1
9.4		Requirements set by voluntary standards & guidance	N/A	**4**	2		2
9.5		Privatization of audit, compliance, & regulatory services	N/A	**4**	1	1	2
9.6		Norms of & imbalances in political & economic power	N/A	**4**		3	1
9.7		Avoidance of conflicts of interest in policy co-design	N/A	**3**	2	1	
9.8		Emerging norm of accountability & enforceability gaps	N/A	**3**		3	
9.9		Cross-jurisdictional legal frameworks & agreements	N/A	**2**	1		1
9.10		Separation of private & public sector AI regulations	N/A	**2**	2		
9.11		Legal limitations on data sharing & integration practices	N/A	**2**	1		1

9.12		Organizational policies & norms as potential barriers to collaboration	N/A	**2**	1	1	
9.13		Self-regulation of private sector	N/A	**2**		1	1
9.14		Emerging norm of agile AI governance	N/A	**2**		1	1
9.15		Misalignments between worker values & organizational norms	N/A	**1**		1	
9.16		Norms of stakeholder inclusion/exclusion from governance activities	N/A	**1**		1	
9.17		Competing/conflicting mandates of government institutions	N/A	**1**		1	
9.18		Emerging norm of AI policy work becoming professionalized	N/A	**1**			1
9.19		Legal & moral responsibilities to prevent harm	N/A	**1**		1	
9.20		Norms of technological futurism	N/A	**1**			1
9.21		Norms of supply chain governance	N/A	**1**		1	
10.1	**10. Ecosystem**	International alignment on AI governance frameworks	N/A	**13**	5	4	4
10.2		Cross-initiative feedback loops & learning	N/A	**11**	8	1	2
10.3		AI Brussels effect	N/A	**7**	3	1	3
10.4		Transferable & open-source models for organizational AI governance	N/A	**6**	4	1	1
10.5		Canadian access to international AI markets	N/A	**5**	1	1	3
10.6		Institutional clustering & scope of co-governance activities	N/A	**5**	4	1	
10.7		Coordination between AI strategies & AI policies	N/A	**4**	2	2	
10.8		Implementation of organizational & sectoral AI governance frameworks	N/A	**4**	2	1	1
10.9		Harmonization of economic development & technology regulation goals	N/A	**3**	2	1	
10.10		Influence of existing privacy governance frameworks	N/A	**3**	2		1
10.11		Power of private & public investments to steer the ecosystem	N/A	**3**	2	1	
10.12		Sector-specific strategies & requirements for AI adoption	N/A	**2**	2		
10.13		Influence of existing impact assessment & audit frameworks	N/A	**2**	1	1	
10.14		Sector-specific gaps in AI awareness & governance capabilities	N/A	**2**		2	

10.15		Environmental uncertainties & risk of systemic benefit/harm misperceptions	N/A	**2**		2	
10.16		Widening of geopolitical & digital divisions	N/A	**2**		1	1
10.17		Pre-regulatory persuasiveness of soft law instruments	N/A	**2**		1	1
10.18		Integration of standards & regulations	N/A	**2**	1		1
10.19		Exclusion of Indigenous communities from governance activities	N/A	**1**		1	
10.20		Exclusion of small & medium enterprises from governance activities	N/A	**1**		1	
10.21		Epistemic & political dependencies on national thought leaders	N/A	**1**		1	
10.22		Proliferation of military AI & potential for malicious use	N/A	**1**		1	
10.23		Implementation of ethical practices in public-private AI partnerships	N/A	**1**		1	
10.24		Design & implementation of regulatory sandboxes	N/A	**1**			1
10.25		Cultivation of trust within ecosystem	N/A	**1**		1	
11.1	**11. Improvements**	Stronger shared understandings of AI systems & best practices for AI governance	N/A	**8**	4	2	2
11.2		Create more opportunities for public participation in AI governance	N/A	**7**	3	2	2
11.3		Stronger strategic coherence & coordination	N/A	**6**	3	1	2
11.4		More strongly integrated ecosystem across departments, governments, & sectors	N/A	**6**		3	3
11.5		More diversity in AI governance activities	N/A	**6**	2	2	2
11.6		Implement stronger participatory design & governance practices	N/A	**6**	4	1	1
11.7		Supply of knowledge & expertise (AI policy experts)	N/A	**4**	2	1	1
11.8		Improved awareness of AI risks, impacts, & needs of impacted stakeholders	N/A	**4**	2	1	1
11.9		More agile & adaptable governance frameworks	N/A	**4**	1	1	2
11.10		Stronger public accountability & oversight in governance activities	N/A	**4**	2	1	1
11.11		Cultivation of robust AI governance cultures	N/A	**3**	2	1	
11.12		More interdisciplinarity in AI governance practices	N/A	**3**	2		1

11.13		Scalability & resilience of AI ecosystem	N/A	**3**	3		
11.14		Improved enforceability of policies & regulations	N/A	**3**	1	1	1
11.15		Resistance against techno-determinism	N/A	**3**	2	1	
11.16		Cultivation of AI literacies & stronger AI education capacities	N/A	**3**	1	2	
11.17		Stronger AI policy capacities in public sector institutions	N/A	**3**	1	1	1
11.18		More financial resources allocated to AI governance activities	N/A	**3**	1	2	
11.19		Reduce barriers to public participation	N/A	**3**	1		2
11.20		Encourage more experimentation in designing AI governance frameworks	N/A	**3**		3	
11.21		Supply of knowledge & expertise (Technical experts)	N/A	**2**		2	
11.22		Stronger long-term planning	N/A	**2**	1	1	
11.23		Improved scalability & resilience of AI governance frameworks	N/A	**2**	1		1
11.24		Resistance against AI hype	N/A	**2**	1	1	
11.25		Secure greater public trust in AI systems & AI governance	N/A	**2**	1	1	
11.26		Stronger public protection measures in emerging regulations	N/A	**2**	1		1
11.27		Resolve technical debts in digital infrastructures	N/A	**2**		1	1
11.28		Greater support for helping domestic AI companies to maturity	N/A	**2**		2	
11.29		More clear & realistic compliance expectations for AI companies	N/A	**2**		1	1
11.30		Prioritize the creation of sectoral AI governance frameworks	N/A	**2**		1	1
11.31		Supply of knowledge & expertise (AI business experts)	N/A	**1**	1		
11.32		Better access to global markets	N/A	**1**	1		
11.33		Creation of more open data initiatives in public sector	N/A	**1**	1		
11.34		Resistance against authenticity crisis caused by generative AI	N/A	**1**	1		
11.35		Stronger risk profiling & impact assessment frameworks	N/A	**1**	1		
11.36		More effective balance between innovation and regulation	N/A	**1**	1		
11.37		More proactive governance initiatives	N/A	**1**	1		

11.38		Greater support for operational & implementation needs of AI companies	N/A	**1**		1	
11.39		Prioritize the creation of AI audit markets	N/A	**1**			1
11.40		Implement regulatory sandboxes	N/A	**1**		1	
11.41		Stronger access to justice/recourse in the event of AI harms	N/A	**1**			1
11.42		Stronger participatory design & governance practices	N/A	**1**			1
11.43		Stronger models & supports for companies wishing to refuse AI adoption	N/A	**1**		1	
11.44		Implement more comprehensive frameworks for governing AI value chains	N/A	**1**		1	
11.45		Implement more comprehensive frameworks for organizational AI & data governance	N/A	**1**		1	

Appendix 4B: Service System Analysis Framework & AIG System Process Model

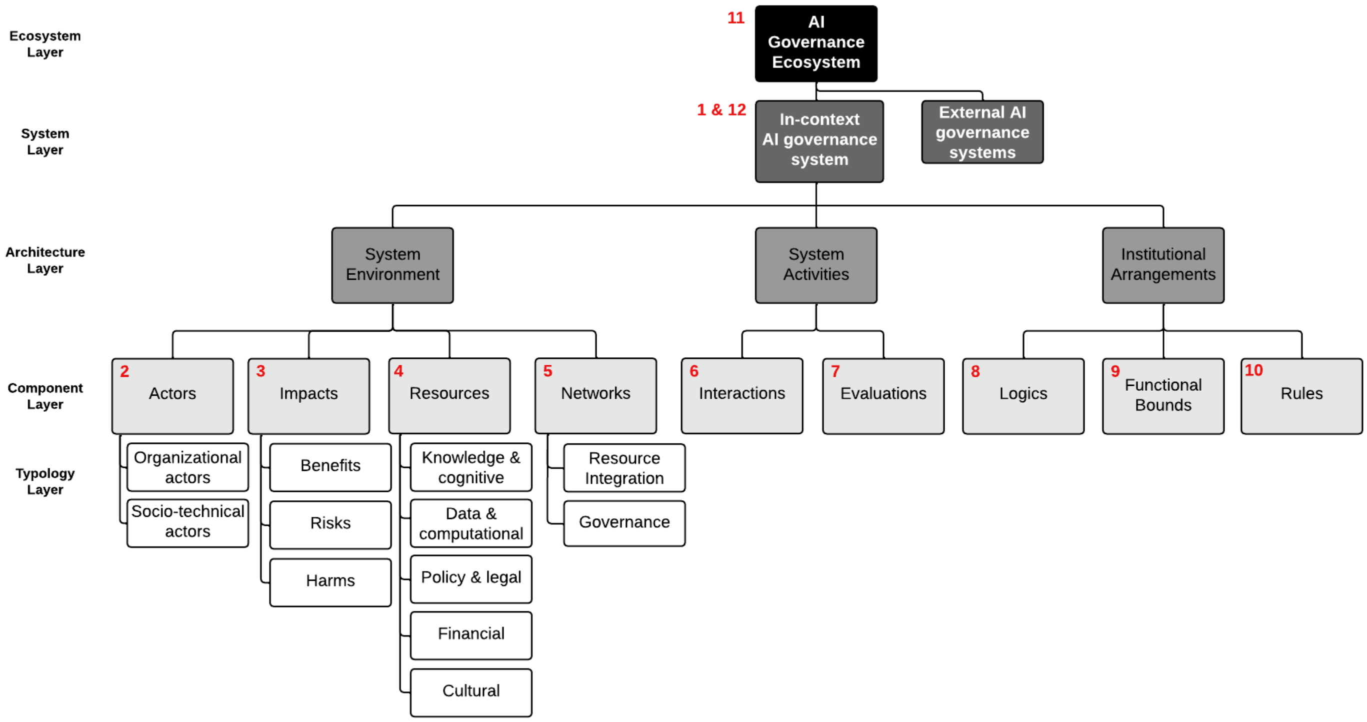

Figure 4B1: A diagram of the service system analysis framework we adapted from the original framework developed by Frost, Cheng, and Lyons (2019). The 12 analytical dimensions that we applied to structure our data collection and analysis process are grounded in components of the framework, highlighted above in 12 red numbers

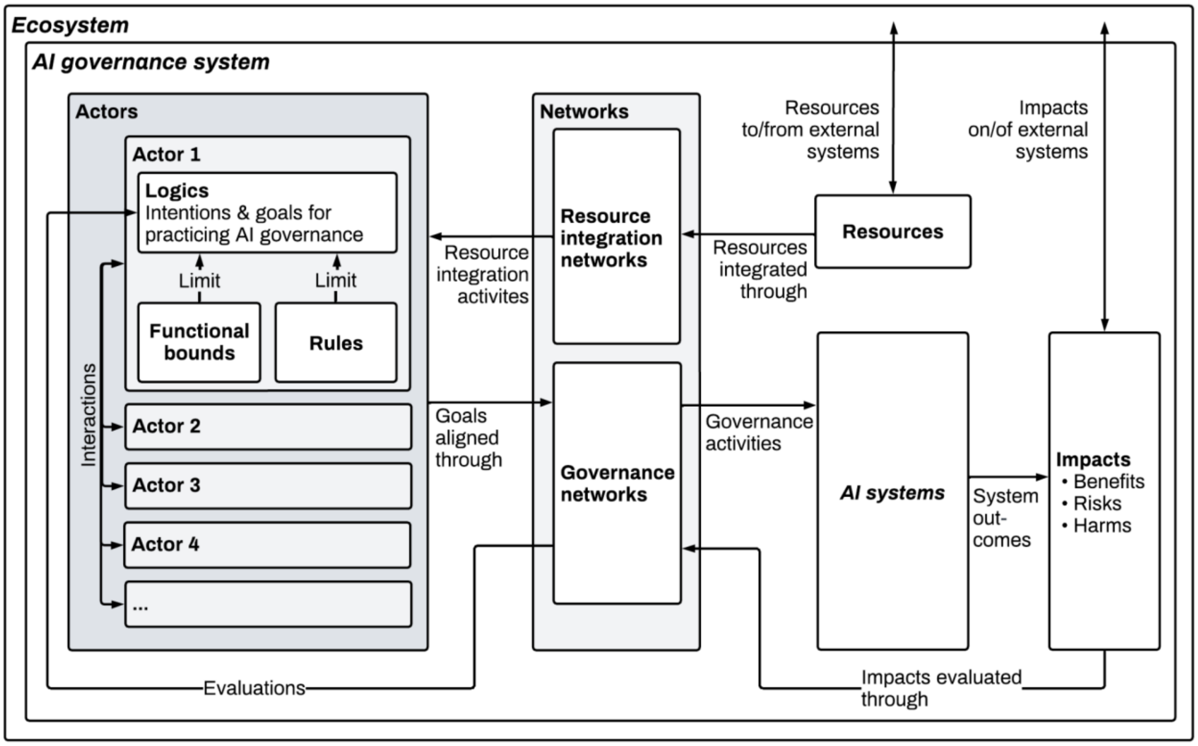

Figure 4B2: Process model of the AIG system structures and activities contained in our framework.

Appendix 4C: Questionnaire for Initial Interviews

Questions for first round of interviews with *initiative leaders* in phase 2 of the study (data collection):

(1) *Work Context*: Tell me about [the name of the AI governance initiative(s) the participant leads]. What are the objectives of the initiative, and why are those particular objectives significant to you?

(2) *Actors*: Which organizations, groups, industries, communities, or other types of social actors are involved in or affected by this initiative?

(3) *Benefits, Risks, & Harms*: What benefits, risks, or harms might each of the actors realize as a consequence of the initiative?

(4) *Resources*: What resources do those actors require in order to be involved in the initiative?

(5) *Networks*: What types of networks do those actors require to interact with one another?

(6) *Interactions*: What are the most significant interactions that occur among those actors? Are any interactions of particular importance to the success of the initiative?

(7) *Evaluations*: How do the actors evaluate those interactions? What qualities or properties might constitute a successful interaction from each actor's perspective?

(8) *Logics*: Why have those actors decided to get involved in this initiative? What are their intentions in doing so?

(9) *Functional Bounds*: Do any of the actors in this initiative face any cognitive, informational, physical, or any other functional limitations?

(10) *Rules*: Are there any significant legal, ethical, or rule-based limitations on their actions?

(11) *Ecosystem*: Are there any other AI governance initiatives in Canada or outside of Canada that you're aware of that are significant to the success of your initiative?

(12) *Improvement*: What are some opportunities for improving this initiative?

Questions for first round of interviews with *subject matter experts* in phase 2 of the study (data collection):

(1) *Work Context*: What AI governance initiatives in Canada are you aware of? Are there any that are of particular interest to you?

(2) *Actors*: Which organizations, groups, industries, communities, or other types of social actors are involved in or affected by those initiatives?

(3) *Benefits, Risks, & Harms*: What benefits, risks, or harms might each of the actors realize as a consequence of those initiatives?

(4) *Resources*: What resources do those actors require in order to be involved in those initiatives?

(5) *Networks*: What types of networks do those actors require to interact with one another?

(6) *Interactions*: What are the most significant interactions that occur among those actors? Are any interactions of particular importance to the success of the initiatives?

(7) *Evaluations*: How do the actors evaluate those interactions? What qualities or properties might constitute a successful interaction from each actor's perspective?

(8) *Logics*: Why have those actors decided to get involved in the initiatives? What are their intentions in doing so?

(9) *Functional Bounds*: Do any of the actors in the initiatives face any cognitive, informational, physical, or any other functional limitations?

(10) *Rules*: Are there any significant legal, ethical, or rule-based limitations on their actions?

(12) *Ecosystem*: Are there any other AI governance initiatives in Canada or outside of Canada that you're aware of that are significant to the success of the initiatives?

(13) *Improvement*: What are some opportunities for improving the initiatives?

Appendix 4D: Questionnaire for Follow-up Interviews

Questions for follow-up interviews with initiative leaders and subject matter experts in phase 4 of the study (quality assurance):

(1) *General Discussion*: Did any particular issues or aspects of the preliminary report's findings stick out to you or strike your interest?

(2) *Accuracy*: Based on your own knowledge of these issues, are these findings a reasonably complete representation of Canada's system of AI governance?

(3) *Completeness*: Based on your own knowledge of these issues, are these findings an accurate representation of Canada's system of AI governance?

Appendix 4E: Inclusion Criteria for Participation in Study

We selected leaders of public sector AI governance initiatives to recruit as participants in accordance with the following inclusion criteria:

- *Active leadership:* At the time of the interview, the participant has a significant and active role in leading the design and/or implementation of a federal, provincial, or municipal AI governance initiative in Canada (e.g., the participant has been actively involved in conceptualizing or managing the initiative, not just analyzing or reporting on its outcomes).
- *Initiative impact:* The initiative has garnered international, national, and/or local attention due to the activities, goals, and/or outcomes of the initiative.
- *Knowledge breadth & depth*: Due to their knowledge of and expertise in AI governance practices, the participant may also be capable of commenting on the broader state of AI governance activities in Canada as they relate to the goals of their own initiative as well as beyond the goals of their own initiative.

We selected subject matter experts to recruit as participants in accordance with the following inclusion criteria:

- *Knowledge breadth & depth*: As a result of their experience carrying out research, professional practice, and/or civic engagement, the participant has a significant breadth and depth of knowledge of AI governance activities in Canada.
- *Public voice*: The participant has demonstrated their subject matter expertise by publishing written work and/or by commenting publicly about AI governance activities in Canada.

Appendix 4F: Strategic Objectives & Potential Interventions for Strengthening Canada's AIGS

Strategic Objective 1: Implement new collaboration & coordination mechanisms. In describing opportunities for improving the overall state of Canada's AIGS, participants frequently noted challenges of coordination, resource integration, and value misalignment between actors and across scales of activity. Participants suggested opportunities for *creating more opportunities for public participation in AI governance* (ID#11.2), *stronger strategic coherence & coordination* (ID#11.3), a *more strongly integrated ecosystem across departments, governments, & sectors* (ID#11.4), *improved awareness of AI risks, impacts, & needs of impacted stakeholders* (ID#11.8), and to *implement stronger participatory design & governance practices* (ID#11.6). These challenges confirm findings from a previous study of 84 Canadian AIG initiatives (Attard-Frost, Brandusescu, & Lyons, 2023), in which the authors observe "an opportunity for policymakers and public servants to cultivate a more unified national approach to AI governance by designing initiatives that are intended to integrate resources and perspectives from a more diverse range of stakeholders, reduce barriers to collaboration and coordination, resolve conflicts, and build greater legitimacy" (p. 29). The authors of that study recommend that policymakers and public servants address these challenges by implementing new governance mechanisms such as pan-Canadian AIG collaboratives or formalized AIG communities of practice. In light of the perceptions of our participants in our study, we reiterate that recommendation here.

We also note that such mechanisms are becoming standard practice among the AIG systems of some of Canada's closest international partners. The United States coordinates AIG activities between departments, agencies, governments, sectors, and civil society through several mechanisms, including an all-of-government executive order on AI (The White House, 2023) and the National AI Initiative, established in 2020 with a mandate to coordinate activities across departments and sectors (National Artificial Intelligence Initiative Act, 2020). The European AI Alliance (European Commission, 2024) is an initiative launched by the European Commission in 2018 to facilitate collaboration on AIG between governments, sectors, and civil society. The United Kingdom's proposed regulatory framework for AI (2023) calls for the creation of new "central functions" within the federal government. These central functions will be mandated with supporting the implementation, monitoring, and evaluation of the UK regulatory framework, but also mandated with strengthening coordination, collaboration, education, and awareness across departments, agencies, sectors, and civil society. Canada should make stronger efforts to follow the example of its international partners by implementing new national collaboratives and formal communities of practice, working groups, new central functions and agencies with coordination mandates, and other mechanisms for enabling a diverse, inclusive, and equitable approach to pan-Canadian coordination and collaboration on AIG issues. With their central role in developing and integrating the knowledge resources of many actors from across Canada through the Pan-Canadian AI Strategy, CIFAR is a well-positioned organization to take an initial lead on the planning of new cross-sectoral coordination mechanisms. Within the public sector, the existing expertise of Treasury Board of Canada Secretariat (TBS) in integrating knowledge resources from across departments and governments to guide the federal public service on AIG issues makes TBS well-positioned to take an initial lead on the planning of new interdepartmental and intergovernmental coordination mechanisms.

Strategic Objective 2: Create guidance for designing & implementing participatory AIG initiatives. Our participants in this study identified several opportunities to improve public participation and trust in AI strategy initiatives and AI policy initiatives. In describing opportunities for improving the *design* of AIG initiatives, participants often perceived a need to include a more diverse range of AI risks, impacts, and impacted stakeholders in the scoping and planning of AI strategies and policies. Participants described these opportunities with reference to key national AI strategies and policies (most significantly, the Pan-Canadian AI Strategy and the Artificial Intelligence and Data Act), as well as with reference to the internal AI strategies and policies of specific public sector and private sector organizations. In the *implementation* of AIG initiatives, participants expressed concerns about burdensome compliance requirements and the enforceability of laws, regulations, guidelines, and other policies for AI. Participants described administrative, financial, and knowledge barriers to effective regulatory enforcement, and similar barriers to effective compliance management in small and medium-sized enterprises. Participants also suggested that AI strategies and policies could be made more agile, adaptable, and inclusive by enabling impacted stakeholders to continue contributing their knowledge of application-specific and sector-specific risks, of advances in AI innovation and AI impacts, and of shifts in the strategic environment throughout the implementation process.

In addition to design and implementation challenges, some participants also perceived a lack of sufficient innovation and experimentation in Canada's AIG initiatives. Participants stated that Canada would benefit from new public sector and cross-sectoral open data initiatives, initiatives to improve the reliability and resilience of existing technical infrastructures, initiatives to support the scale-up and maturation of domestic AI companies, pre-market regulatory sandboxes for high-risk AI applications, and initiatives to develop and test new mechanisms for public accountability and oversight in AI systems.

To begin acting upon these opportunities for improvement, policymakers, public servants, and other practitioners of AIG in Canada should develop and implement context-sensitive guidance on principles and best practices for participatory governance of AI systems. Voluntary guidance documents have become a popular instrument for supporting organizations in remaining agile, adaptable, and effective amidst the rapid, widespread adoption of generative AI. In Canada's federal government, ISED, TBS, and the OPC have all collaborated with their respective partner organizations to create guidance on the design, development, deployment, use, and management of generative AI systems. Government institutions should also collaborate to develop guidance that can support public sector and private sector organizations in implementing more diverse and inclusive stakeholder participation practices throughout the design, development, deployment, use, management, and governance of AI systems. The coordination mechanisms we outline in our first strategic objective would support government institutions in collaboratively developing the guidance, and would also support public sector and private sector organizations across Canada in implementing the guidance.

Against the backdrop of a global "participatory turn" in AI design (Delgado et al., 2023), participatory AIG practices–such as policy co-design, continuous collaborative decision-making, public oversight bodies, awareness-building and capacity-building initiatives, and removing barriers to and incentivizing participation of marginalized stakeholders–are a strategic necessity for maintaining Canada's global leadership in AI. Participation gaps were implicated in many of the opportunities for improvement perceived by our participants, most notably *create more opportunities for public participation in AI governance* (ID#11.2) and *implement stronger participatory design & governance practices* (ID#11.6). Corroborating those perceptions, a 2023 report by ISED's AI Public Awareness Working Group recommends that the government strengthen public participation in AI throughout Canada by launching "sustained and government-led public information campaigns" and by recognizing "the value and necessity of engagement with citizens" in AI policy initiatives (p. 12). The report describes the results of a survey of 1,222 Canadians that was commissioned by the Working Group:

People in Canada demonstrate a strong desire to be consulted prior to the elaboration of

> public policies on AI: partly because of a fear of loss of control in the face of increasing automation in administrative processes; and partly because of the principle that those impacted by public policy should have a voice [...] these public consultations should aim to include the widest diversity of participants and groups: not just experts or people who already have knowledge, but all people impacted by AI deployment. (p. 36)

Context-sensitive guidance on participatory governance will enable Canadian organizations to create AI systems and governance frameworks that are more responsive to the contexts and needs of impacted stakeholders, and more adaptable to changes in those contexts and needs. Encouraging public sector, private sector, and civil society organizations across Canada to implement more participatory approaches to AIG will provide stronger civic counterbalances against imbalances of political and economic power in the design and implementation of AIG initiatives. Guidance on participatory AIG can also benefit new initiatives for open data, regulatory sandboxes, technical infrastructure improvement, and business scale-up and maturation by providing organizations with recommended principles and practices for collaborating with a diverse set of impacted stakeholders throughout the design and implementation of those initiatives.

Strategic Objective 3: Expand access to key resources needed for effective AIG practices. In describing opportunities for improving the overall state of Canada's AIGS, participants frequently noted that the development of Canada's AIGS is limited by many knowledge resource constraints. These constraints include the limited availability of AI policy experts, technical experts, and business scale-up experts throughout Canada; the supply of stakeholder knowledge of AI risks and impacts within stakeholder-specific AI application contexts to policymakers and other AIG practitioners; and the supply of training and education on AI issues to workers, decision-makers, publics, and other impacted stakeholders. Financial constraints were also perceived as limiting the ability of many small and medium-sized enterprises, public sector organizations, civil society organizations, and marginalized individuals and groups to meaningfully participate in AIG initiatives, particularly standardization and policy co-design initiatives. Participants also noted the limited availability of audit and governance frameworks and other policy resources that can be easily transferred across organizations and sectors, a challenge which is only compounded by the low availability of AI policy experts, gaps in context-specific stakeholder knowledge and training, and the still-nascent AI governance cultures of most Canadian organizations.

To address these resource constraints, policymakers, public servants, and other practitioners of AIG will need to act upon our first and second strategic objectives: implement new coordination mechanisms and new guidance for participatory AIG initiatives. If new coordination mechanisms and participatory AIG practices are effectively implemented, they will strengthen cross-organizational and cross-sectoral policy development and transfer, knowledge of context-specific AI impacts and impacted stakeholders, and capacities for training, education, and regulation. Alongside stronger coordination and participation practices, Canada will also require a national strategy to improve public sector, private sector, and civil society access to specialized expertise, to fund and invest in AIG practices, to build stronger AIG capacities within organizations, and to support organizations in cultivating AIG cultures that are responsible, robust, and scalable.

Such a strategy could be implemented through a series of new strategic programs that are intended to expand the supply of and equitable access to AIG knowledge resources across Canada. While some Canadian institutions have already implemented training programs to develop a domestic supply of AI policy expertise and scale-up expertise–such as Mila’s AI Policy Compass program (2024) and the Vector Institute’s FastLane program (2024)–a more comprehensive approach to these challenges is still needed. Specialized AIG training and education programs for governments and sectors across Canada, global AIG talent recruitment programs, and public funding and investment programs for building AIG capacities should be implemented to reduce constraints on AIG activities and to expand access to key AIG resources throughout Canada. Needs assessments should be conducted to identify vulnerable groups facing high AI risks and severe AIG resource constraints, and to develop targeted programs and other means of supporting access to the type of resources that they require to secure greater benefit from AI while also mitigating risks and harms. Access supports could include prioritized public funding and investment in the needed resources, specialized AI training and education programs that are sensitive to the unique needs and contexts of diverse stakeholders (Figaredo & Stoyanovich, 2023), and low- or no-cost access to AI strategy experts, legal and policy experts, or advocacy organizations for under-resourced individuals, groups, or organizations. At the federal level, these programs could be layered onto the existing funding, recruitment, talent development, training, and business support functions of ISED and CIFAR's Pan-Canadian AI Strategy, onto the OPC's existing research, consultation, and awareness-building functions, or onto existing internal AI training programs within federal institutions. Programs could also be managed by the new AI & Data Commissioner established by the Artificial Intelligence and Data Act, or by other new agencies with coordination mandates as discussed in our first strategic objective.

Strategic Objective 4: Advance diversity, equity, & inclusion in AIG activities. Participants frequently perceived opportunities to advance more diverse, inclusive, and equitable governance practices across Canada's AIGS, such as *more diversity in AI governance activities* (ID#11.5), *stronger public accountability and oversight in governance activities* (ID#11.10), *more interdisciplinarity in AI governance practices* (ID#11.12), and *reduce barriers to public* *participation* (ID#11.19). Challenges of diversity, equity, and inclusion in AIG practices have also been widely observed beyond the Canadian context (Cachet-Rosset & Klarsfeld, 2023; Roche, Wall, & Lewis, 2023; West, Whittaker, & Crawford, 2019). In the lifecycle of AI systems, challenges of diversity, equity, and inclusion feed into one another: excluding a diverse set of AI impacts and impacted stakeholders from the design of AI systems and governance frameworks can cascade into downstream biases in development and implementation, ultimately resulting in system outcomes that disproportionately harm the excluded stakeholders and reinforce structural inequities (Schwartz et al., 2022). These inequitable outcomes can include many forms of systemic discrimination and representational harm (Shelby et al., 2023), as well as allocative harms, such as inequitable distributions of resources, benefits, and responsibilities across the value chains of AI systems (Attard-Frost & Widder, 2023) and structural barriers to accessing justice, recourse, or remedy for harmful AI outcomes (Ogunleye, 2022).

Acting upon the three strategic objectives we previously discussed–implementing stronger coordination mechanisms to integrate a diversity of actors, stronger participatory AIG practices, and programs to expand and ensure equitable access to key AIG resources–would support the advancement of diversity, equity, and inclusion in Canadian AIG activities. Additionally, the 2023 report of ISED's AI Public Awareness Working Group provides guidance on implementing a national "Equity, Diversity, Inclusion, and Accessibility strategy . . . to ensure the financial, accessibility, and outreach resources are available to enable representation of the diversity of peoples in Canada" (p. 12). A strategic initiative such as this could serve as a platform for a variety of resource-providing and capacity-building programs to support the advancement of diversity, equity, and inclusion in Canada's AIGS. A national survey and impact assessment of the effects and potential effects of AI systems on equity-deserving groups across Canada could provide additional support. Community-centric accountability processes

(Häußermann & Lütge, 2022) and interventions in social and economic policy (Haugen et al., 2021; Merola, 2022; O'Keefe et al., 2020) should also be developed by governments and socially responsible businesses to provide equity-deserving groups with greater guarantees of beneficial outcomes from AI systems, as well as guarantees of access to justice when faced with harmful outcomes.