ETHICS, PRIVACY AND BEYOND: A HABERMASIAN ANALYSIS OF ARTIFICIAL INTELLIGENCE DISCOURSES IN CANADIAN HEALTHCARE

A Research Project Submitted to the
College of Graduate and Postdoctoral Studies
In Partial Fulfillment of the Requirements
For the Degree of Master of Arts in Sociology
In the Department of Sociology
University of Saskatchewan
Saskatoon

By
Cristina Diego-Chavez

© Copyright Cristina Diego-Chavez, April 2024. All rights reserved
Unless otherwise noted, copyright of the material in this thesis belongs to the author
Permission To Use

In presenting this research project in partial fulfilment of the requirements for a Master of Arts in Sociology from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this project in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department of Sociology or the Dean of the College of Arts and Sciences. It is understood that any copying or publication or use of this project or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my research project.

Requests for permission to copy or to make other use of material in this research project in whole or part should be addressed to:

Head of the Department of Sociology
9 Campus Drive
University of Saskatchewan
Saskatoon, SK S7N 5A5

OR

Dean College of Graduate and Postdoctoral Studies
University of Saskatchewan
116 Thorvaldson Building, 110 Science Place
Saskatoon, Saskatchewan S7N 5C9, Canada
ABSTRACT
This paper examines the challenges of Artificial Intelligence (AI) in Canadian healthcare by employing Jürgen Habermas's three-dimensional theory of rationality (Cognitive-Instrumental, Aesthetic-Expressive and Moral-Practical) as a framework for analysis. Through a review of both peer-reviewed and grey literature from 2018 to 2023, the research reveals an interconnection among ethical, privacy, technological, and legal concerns. This reflects the validity and applicability of Habermas's three-dimensional theory of rationality as a useful framework to assess the multifaceted discourses on AI in healthcare. Ethics and privacy are identified as central to navigating AI's implications in healthcare, suggesting their foundational role in shaping policies and practices.
TABLE OF CONTENTS

PERMISSION TO USE ................................................................. II
ABSTRACT ........................................................................ III

TABLE OF CONTENTS ........................................................... IV

LIST OF FIGURES AND TABLES ............................................... V

CHAPTER ONE: INTRODUCTION ........................................... 1
  1.1 BACKGROUND AND CONTEXT ....................................... 1
  1.2 RESEARCH QUESTION AND OBJECTIVE ....................... 2

CHAPTER TWO: LITERATURE REVIEW .................................... 3
  2.1 JÜRGEN HABERMAS: RATIONALITY AND MODERNIZATION ....... 3
  2.2 THE THREE-DIMENSIONAL RATIONALITY APPROACH ............ 3
    2.2.1 Cognitive-Instrumental Rationality ................................ 4
    2.2.2 Aesthetic-Expressive Rationality .................................. 4
    2.2.3 Moral-Practical Rationality......................................... 4
  2.3 ARTIFICIAL INTELLIGENCE IN HEALTHCARE ...................... 5
    2.3.1 Data Management: Ethics and Privacy in AI Healthcare Integration .... 6
    2.3.2 Canadian Healthcare Context and Technological Transformation ....... 8
    2.3.3 Privacy and Data Use Regulations in Canada.................... 8

CHAPTER THREE: RESEARCH METHODS .................................... 10
  3.1 DATA COLLECTION - INCLUSION CRITERIA ....................... 10
  3.2 LITERATURE IDENTIFICATION AND QUALITY ASSESSMENT .... 10
  3.3 DATA EXTRACTION AND ANALYSIS ................................... 11

CHAPTER FOUR: RESULTS ....................................................... 13
  4.1 DESCRIPTIVE ANALYTICS ............................................. 13
  4.2 COMMON STRUCTURE OF THE THREE RATIONALITIES ACROSS ACADEMIC AND GREY LITERATURE ........................................ 14
    4.2.1. Cognitive-Instrumental Rationality Discussions .................. 16
    4.2.2. Aesthetic-Expressive Rationality Discussions .................... 18
    4.2.3 Moral-Practical Rationality Discussions ............................ 22
  4.3 MAIN DIFFERENCES ACROSS GREY AND ACADEMIC LITERATURE .... 25

CHAPTER FIVE: DISCUSSION AND CONCLUSIONS ...................... 27
REFERENCES ......................................................................... 31

APPENDICES ....................................................................... 38
  APPENDIX A: TABLES OF SEARCH QUERIES AND SEARCH RESULTS ........... 38
  APPENDIX B: REFERENCES AND TERRITORIAL DIVISION CLARIFICATION TABLE .... 42
  APPENDIX C: REFERENCES BASED ON RATIONALITY AND SUBTOPICS TABLE .... 43
LIST OF FIGURES AND TABLES

Figure 1 - Diagram of the Convergence of the Three Forms of Rationality – Types of Issues Raised (15)
Table 1 - Differences Between the Three Forms of Rationality – Types of Issues Raised (25)
CHAPTER ONE: INTRODUCTION

1.1 Background and Context

Artificial Intelligence (AI) refers to algorithms that learn from data and experiences to make decisions and perform tasks. These systems adapt and respond based on the information they accumulate, simulating human cognitive processes to address complex challenges (Craglia et al., 2018). AI is a transformative and disruptive technology that has been widely deployed and used in multiple settings, including healthcare. Its applications in healthcare include diagnostics, image analysis and interpretation, administrative tasks, treatment precision optimization and clinical drug trials (Iqbal, Cortés Jaimes, Makineni et al., 2023; Davenport & Kalakota 2019; Marr 2018).

While AI seeks to enhance healthcare and improve patient outcomes, it faces several types of issues throughout the development, implementation, oversight, and regulatory processes. These challenges require a careful balance between using AI's capabilities and adhering to ethical and privacy standards in healthcare settings.

The integration of AI into healthcare systems and practices raises important questions about the current literature on this topic, especially in the context of Canadian healthcare. To explore these issues, we can apply Habermas's three-dimensional theory of rationality as a framework to categorize the key concerns.

According to Habermas, the increasing use of rationality as a guiding principle for decision-making in societies is an indicator of modernization, which manifests in three separate subsystems of human knowledge and action. These subsystems include the objective-physical world, governed by cognitive-instrumental rationality; the subjective-individual world, guided by aesthetic-expressive rationality; and the intersubjective-social world, driven by moral-practical rationality. These classifications highlight how different forms of rationality shape our interactions and interpretations across the physical, individual, and societal spheres, reflecting the complexities of modern social structures (Dickinson 2021, 2020; Johnson 2006).

The three-dimensional concept of rationality serves as a framework to understand how a variety of emerging issues of implementing AI in Canadian healthcare can be aligned with instrumental, expressive and moral rationalities.

The data collected for the literature mapping included both peer-reviewed articles and grey literature about the use of AI in healthcare, concretely on issues linked to privacy and ethical
aspects of applying this technology in Canada from 2018 to 2023. The search strategy integrated the use of academic databases for retrieving peer-reviewed literature: Medline, Scopus, Semantic Scholar and Web of Science. To access the grey literature, the web search engine Google was used to retrieve journals, magazines, newspapers and institutional and governmental sources.

1.2 Research Question and Objective
The objective of this paper is to categorize the discourses present in the literature selected and reviewed on AI in Canadian healthcare, using Habermas’s three worlds/three rationalities model. To meet this purpose, the central research question states: Is the three-dimensional rationality concept of Jürgen Habermas evident in the literature regarding the challenges of implementing Artificial Intelligence in Canadian Healthcare? If so, how is this concept represented across the literature selected?

The following research is divided into five chapters. Chapter 1 provides the research introduction, including the background, objectives and research question. Chapter 2 provides a review of the core concepts of modernization theory from Habermas's and the three-dimensional theory of rationality, elaborating on Cognitive-Instrumental, Aesthetic-Expressive and Moral-Practical Rationalities. This theoretical framework is used to map the literature. Additionally, this chapter outlines existing literature regarding AI in healthcare, focusing on the topics of data ethics and data privacy. Data aspects are crucial for understanding AI functioning and intersect with different problems and discussions. Finally, this chapter covers the aspects of the Canadian healthcare context in regards to technological transformation and data privacy and use regulations. Chapter 3 is dedicated to discussing the research methodology employed in this study. Chapter 4 systematically presents the findings derived from the conducted research. Finally, Chapter 5 covers both the discussion and conclusion, providing a summary of four key findings that cover the most relevant insights of this project.
CHAPTER TWO: LITERATURE REVIEW

2.1 Jürgen Habermas: Rationality and Modernization

To Habermas, modernization is a process that involves the historical shift from traditional societies to modern societies. In these modern societies, rationality is the core element of transformation because it enables subjects to continue communicative action through rational argumentation of knowledge claims (Habermas 1984).

Since rationality is central to modernization, Habermas proposes a three-dimensional concept of rationalization that hypothesizes three structurally and functionally differentiated subsystems of knowledge and action: Cognitive-Instrumental, Aesthetic-Expressive and Moral-Practical (Dickinson 2021, 2020; Johnson 2006). This concept will be further elaborated in the next section.

Habermas’ theory usually considers modernization since the development of Western society, from the end of the medieval period to the late 20th century. The process of modernization includes several developments. An example of this is the exponential growth in knowledge, more specifically in the natural sciences after the 17th century, displacing epistemic religious knowledge with natural science and reason, where religious values of the Middle Ages increasingly became subject to criticism, demanding rational justifications of the knowledge claims made. This massive increase in technically useful knowledge promoted the wake of rationalization in societies, which provides context for understanding the emergence of anti-modernization counter-movements as responses to the challenges and changes introduced by Western modernization (Edgar 2006; Johnson 2006; Zapf 2004).

Habermas emphasizes the importance of understanding the consequences of an imbalanced form of rationality. While the instrumental rationalization process can enhance the production of specialized knowledge, overemphasizing aspects of instrumental rationality in the spheres of inter-subjective social interactions and subjective-personality systems can lead to adverse effects (Dickinson 2020; Johnson 2006).

2.2 The Three-Dimensional Rationality Approach

According to Johnson (2006), Habermas states that the rationalization process that results in modernization generates three spheres of knowledge and action: cognitive-instrumental, aesthetic-expressive and moral-practical. Specific types of knowledge claims and criteria of
validity characterize each sphere, or “world” of knowledge-action. These “worlds” are known as objective-physical, subjective-personal and intersubjective-social, and they correlate with the three spheres of discourse: theoretical, therapeutic and moral-practical.

2.2.1 Cognitive- Instrumental Rationality
This type of rationality seeks to assess the validity of claims made in relation to the physical world of objects. It is concerned with efficiency, effectiveness, and goal-achieving actions to solve problems or make decisions. This dimension of rationality is often associated with scientific and technological development and the application of technical knowledge to solve problems. Discourses and discussions related to this type of rationality can include an emphasis on outcomes achieved for a specific purpose, prioritizing efficiency and effectiveness (in terms of cost-effectiveness, time efficiency or source utilization), or reference to the application of scientific, technical or specialized knowledge to solve problems (Dickinson 2021, 2020; Johnson 2006). In the context of AI in healthcare, this could include the use of data or algorithms to optimize health procedures.

2.2.2 Aesthetic- Expressive Rationality
The aesthetic-expressive rationality, which focuses on self-expression and self-knowledge (Dickinson 2021, 2020; Johnson 2006), emphasizes the development of individual identity and aesthetic sensibilities (Ingram 1991). It engages with topics such as lifestyle choices, subjective experiences, and the expression of individuality, stating that feelings and experiences constitute valid forms of knowledge (Dickinson 2021, 2020). Within the context of AI implementation in healthcare, this perspective could manifest in discussions surrounding patient autonomy, exemplified by practices such as informed consent procedures.

2.2.3 Moral- Practical Rationality
According to Habermas (1973), the growing complexity of possible results requires an accelerated change of social norms when traditional societies are modernized. The moral-practical knowledge sphere provides a normative foundation for mutual understanding to make just decisions collectively. This aims to establish a social order grounded in specific moral rules, principles, and legitimate laws. Discourses that involve moral-practical rationality often include
normative claims that offer frameworks on ethical behaviour, decisions, or policies. This involves moral and ethical critiques that consider the impact of decisions on the well-being of individuals, communities, and society (Dickinson 2020; Johnson 2006; Finlayson, 2006). In healthcare settings seeking to apply AI, this can be translated to discussions about legal regulations to protect patients’ privacy or ethics about data management. These dimensions of rationality are interconnected and interact in complex ways. Habermas argues that a comprehensive grasp of rationality must account for all three dimensions to understand different claims and discourses (Dickinson 2021, 2020).

2.3 Artificial Intelligence in Healthcare

The following sections explore the transformative impact of AI by describing its application in healthcare, including machine learning, deep learning and natural language processing. Additionally, the ethical and privacy challenges linked to AI integration are addressed, with a special focus on data management.

In general terms, Artificial Intelligence is a generic name for any machine or algorithm capable of learning based on training and experience gained, taking specific actions and proposing decisions (Craglia et al., 2018). Even though AI amplifies human intelligence and improves human efficiency, it is not meant to replace the essential elements of human interaction in medicine (Bajwa, Munir, Nori & Williams 2021).

AI is transforming healthcare by enhancing diagnostics and clinical research. For instance, AI models are being developed for breast cancer detection in the UK and South Korea. In the US, it is used for identifying skin cancer and predicting heart diseases. These developments showcase AI's role in boosting healthcare precision and efficiency, with potential applications in robotic surgery and virtual nursing assistance (Alowais, Alghamdi, Alsuebany et al., 2023; IBM 2023; Bajwa et al., 2021; Hewlett Packard Enterprise n.d.).

Before being employed in healthcare, AI technologies must be trained with clinical data, such as demographics, medical records, equipment outputs, assessments, and test images. This training helps AI identify patterns and correlations between patient characteristics and outcomes (Jiang, Jiang, Zhi et al., 2017).
Some of the most common forms of AI integrated into healthcare include:

- **Machine learning (ML)**: This branch of AI is focused on training algorithms using data sets to create models capable of performing specific tasks, such as predicting outcomes after analyzing a vast amount of data and identifying patterns (Davenport and Kalakota 2019; Jiang, Jiang, Zhi et al., 2017).

- **Deep learning (DL)**. DL is a neural network with many layers, which enables it to analyze greater volumes of data and perform more complex tasks considering non-linear relationships (Davenport and Kalakota 2019; Jiang, Jiang, Zhi et al., 2017).

- **Natural language processing (NLP)**: Refers to using ML to understand human language, whether verbal or written. Much clinical data is in unstructured text form, like physical exams and lab reports. NLP extracts valuable information from this text to aid clinical decisions and analysis (Davenport and Kalakota 2019; Jiang et al., 2017).

- **Robotic process automation (RPA)**: AI can be applied in computer programs to automate administrative and clinical workflows to improve patient experience and daily functions in healthcare facilities. This is frequently defined as RPA (Davenport and Kalakota 2019).

### 2.3.1 Data Management: Ethics and Privacy in AI Healthcare Integration

While the incorporation of AI into healthcare highlights its significant potential, the technology predominantly remains in the experimental phase. There is a noticeable gap in comprehensive research to implement and replicate AI across various clinical settings (Antoniou, Muhammad 2021; Geis, Brady, Wu, et al., 2019).

Within a broad spectrum of AI application concerns, issues of data management, including data ethics and data privacy stand out as particularly significant, as data forms the foundational basis for the functioning of AI systems. Data management is central to a wide variety of challenges and discussions including technical aspects, ethical concerns, and regulatory considerations, which collectively influence the development and effective utilization of AI technologies. These concerns not only encompass the collection, storage, and processing of patient data but also extend to the implications of data use on patient autonomy, confidentiality, and trust in the healthcare system (Soumit 2022).
The ethics of data involves a wide range of moral principles and considerations related to all stages of data management. Even though legal regulations are key, this dimension of data management goes beyond legal requirements to consider what is normatively appropriate and inappropriate, fair and unfair, or harmful and beneficial in the handling of data (Soumit 2022; Geis, Brady, Wu, et al., 2019).

The concerns related to data ethics include questions about consent (how and whether it is obtained), transparency (how data usage is communicated to individuals), equity (ensuring data practices do not discriminate against certain groups and lack representativeness), and accountability (who is responsible for the outcomes of data usage) (Soumit 2022; Gerke, Minssen & Cohen 2020; Challen, Denny, Pitt, et al., 2019; Brady, Wu, et al., 2019).

Accountability for mistakes made by AI is currently ambiguous, which limits assigning ownership of its decision-making process and outcomes (Tigard 2020). Moreover, Murdoch, Jandura, & Caulfield (2022) discuss the financial benefits of private companies gathering data in healthcare through AI systems. Health information is economically valuable for certain organizations, especially because it supports the profitable use of AI technologies. These organizations benefit financially by using data-sharing agreements to access large amounts of patient health information held by the government. Likewise, Jaremko et al. (2019) contribute to the discussion of data and economy by problematizing the issue of data commoditization (treating data as a basic good that can be bought, sold, and traded). Data has the potential to generate insights and drive innovation, making it a valuable asset in today's digital economy.

However, the commoditization of data also poses significant privacy concerns, as personal and sensitive information is often collected, and may be shared or sold without individuals' explicit consent.

Data privacy specifically concerns the rights of individuals to control information about themselves. It involves legal and regulatory frameworks that dictate how personal information is collected, stored, shared, and used. This includes implementing technical and organizational measures to protect the access and handling of personal data through compliance with laws and regulations (Soumit 2022; Prakash, Balaji, Joshi, et al., 2022). For instance, The Artificial Intelligence and Data Act (AIDA) in Canada is not yet approved but is in the legislative process. This proposed legislation aims to establish a framework for the responsible design, development,
and deployment of AI systems. It seeks to ensure these systems are safe, non-discriminatory and accountable (Government of Canada 2023).

2.3.2 Canadian Healthcare Context and Technological Transformation

Canada's healthcare system is decentralized and varies across provinces and territories in terms of policies, funding, and priorities, lacking a singular national system (Dickinson & Bolaria 2002; Romanow 2002). In 1984, The Canada Health Act introduced foundational principles like universality, accessibility, comprehensiveness, and public administration, initially tied to federal Medicare funding but eventually embodying the core values of Canada's healthcare approach (Romanow 2002; Nancarrow Clarke 1990). Given modern expectations, Romanow (2002) advocates for adding Accountability as a sixth principle to the Canada Health Act, emphasizing the need for the healthcare system to adapt and reflect current realities, ensuring greater accountability in its management.

The Romanow Report, issued by the Commission on the Future of Health Care in Canada in 2002, identifies ten crucial areas for improvement to sustain the Canadian healthcare system. One of these areas elaborates on incorporating specific technologies into the healthcare system, like the integration of Electronic Health Records (EHRs). This transition would stand out as technological innovation, but Romanow advocates for the careful integration of EHRs. While it may support access and ownership of personal health information, it is also expected to require strict measures to safeguard privacy. He acknowledges the social and ethical challenges posed by new health technologies and calls for re-evaluations, including amendments to the Criminal Code to protect personal health information and prevent its abuse (Romanow 2002).

Nonetheless, despite Canada’s role as a pioneer and international leader in AI, adoption of AI in the health sector has lagged behind other countries and other industries (Healthcare Excellence Canada 2021).

2.3.3 Privacy and Data Use Regulations in Canada

In Canada, privacy laws are regulated both at the federal and provincial levels, with specific policies addressing the handling of personal health information. In this context, several privacy regulations apply to particular provinces and territories, and others are applied nationwide. For instance, a key federal privacy law is the Privacy Act. As in its last amendment in 2023, it
applies to the federal government and its agencies, stating how to collect, use, and disclose personal information (Government of Canada 2023). In the Canadian private sector, The Personal Information Protection and Electronic Documents Act (PIPEDA) governs how organizations manage personal information in the course of commercial business (Office of the Privacy Commissioner of Canada 2021). Additionally, independent regulations allow provinces to oversee health information management autonomously. Some examples of this include The Health Information Act (HIA) in Alberta (Government of Alberta 2024), The Personal Health Information Protection Act (PHIPA) in Ontario (Government of Ontario 2023), and The Personal Health Information Act (PHIA) in Nova Scotia (Government of Nova Scotia 2021).

In the research field, accessing data for training purposes without violating the privacy of protected health information presents a significant challenge in Canada. Currently, only personnel working within hospitals or hospital networks have access to these data. However, external stakeholders, including researchers, developers, and policymakers, frequently lack the necessary access to these valuable datasets (Digital Health Canada 2022).

In the Canadian context, privacy legislation encounters difficulties due to the lack of standardization both between provinces and within them, as identified by Murdoch, Jandura, & Caulfield (2022). Furthermore, some provinces, like Alberta, have no specific laws regulating the use of AI, so they use the current privacy laws when AI systems process health information (HI) or personal information (PI) (Office of the Information and Privacy Commissioner of Alberta 2023). These privacy laws, while not explicitly designed for AI and data management applications, have been enacted to address related issues.

Despite these disparities, efforts towards aligning provincial policies and calls for harmonization between federal and provincial privacy laws suggest a move towards uniformity across the nation's legal healthcare framework (Prychitko & Siddall 2022; Ontario Chamber of Commerce (n.d)).
CHAPTER THREE: RESEARCH METHODS

3.1 Data Collection - Inclusion Criteria
The research included two types of literature: peer-reviewed articles and grey literature publications, which must provide information on the use of AI in healthcare and issues linked to privacy/ethics discussions or contributions. Only publications written in English from January 2018 to December 2023 were included. Incorporating both academic and grey literature into the research provided a more thorough and up-to-date overview of the topic. This approach enriched the paper by introducing a wider range of perspectives, ensuring that the data collection reflects the most recent debates surrounding AI in Canadian healthcare and within specific provinces and territories.

3.2 Literature Identification and Quality Assessment
To access academic publications, the databases consulted were Medline, Scopus, Semantic Scholar and Web of Science, four frequently used databases by multidisciplinary researchers. In recent years, technological advancements have changed at a rapid pace, especially in the field of AI. Because of this, the publication limit date will range from 2018 to 2023 (articles published in the last five years).

Using Boolean operators, combinations of the keywords were introduced on the research engines (i.e., “Artificial Intelligence” AND “Healthcare,” “Ethics/ Clinical Ethics” OR “Privacy/Confidentiality.”) Different features on each database were used to identify potential articles. For example, Web of Science allows to select fields within the filters such as Sociology, Social Issues, Healthcare Services or Medical Ethics. If the database allowed filtering by country, this feature was used (being Canada marked as a filter). Otherwise, "Canada" was added to the search terms to narrow down the results.

The preliminary screening process for each article focused on the title, abstract, and keywords. The origin and affiliation of the article was double-checked in the screening process as well. If the content seemed to discuss elements of the research topic, the article was selected for further evaluation.

To retrieve grey literature, the web search engine Google was used. The criteria for inclusion, such as the publication date limit, language, use of keywords and verifying the origin/context of the publication, were aligned with those established for the academic literature search ("AI-
Artificial Intelligence,” “Healthcare,” “Privacy,” “Ethics,” and “Canada”). While the academic search focused on Canada without a specific scope for each province or territory, the grey literature search aimed to complement this gap by adding provinces in the search queries (i.e., “AI in healthcare Saskatchewan privacy ethics”, “Artificial Intelligence in health “Ontario” “privacy” OR “ethics”).

The search primarily focused on domains such as .gov, .org, .edu, and .ca for more authoritative sources but assessed the eligibility of certain .com sources since they contributed to the objectives of the research. The results were reviewed based on the relevance order that Google presented. Given the extensive volume of accessible data, this study systematically examined the first 10 publications identified for each search query of Canadian provinces and territories, namely Ontario, Quebec, Nova Scotia, New Brunswick, Manitoba, British Columbia, Prince Edward Island, Saskatchewan, Alberta, Newfoundland and Labrador, Northwest Territories, Yukon, and Nunavut, to ensure a comprehensive and balanced review. The preliminary screening process for each article focused on the title and website domain. If these elements seemed to elaborate on the research topic, the article was further examined.

The quality assessment process for both types of literature involved a multi-faceted approach, ensuring the selection of relevant, up-to-date, and multidisciplinary research sources.

3.3 Data Extraction and Analysis

From each research article, information was extracted based on the following topics that belong to each type of rationality. Depending on the type of issues within the discourses of the article, it could fall into one or more rationalities: (1) Cognitive-Instrumental, (2) Aesthetic-Expressive and (3) Moral-Practical. For example, concerns regarding the protection of data privacy and the potential of breaches when using AI for data management can be interpreted under the framework of rationalities 1 and 3.

Data extraction and coding was performed using NVivo software version 14 to identify patterns in the discourses and frame them into Habermas' three-dimensional rationality theory. To do this, the five-step thematic analysis approach of Braun and Clarke (Byrne 2022) was employed. It initially involved familiarization with the data through reading and re-reading, noting initial observations and patterns. Next, the initial codes were generated by identifying relevant features to the research question and naming patterns across the articles. Consequently, the codes were
organized into potential themes, considering how they fit together to form a coherent picture of the data. Then, the revision of the themes was refined to ensure they accurately represent the articles’ discourses and address the research question. Ultimately, the themes and subthemes were assigned descriptive names that capture the core concept of each one.
CHAPTER FOUR: RESULTS

Of the 122 records selected for full-text screening and after the deduplication process (resulting in 25 peer-reviewed articles and 97 grey literature records), 45 articles met the inclusion criteria and were included in the review (16 peer-reviewed articles and 29 grey literature articles). The exclusion of 77 articles was based on limited data and/or insufficient discussion on ethics and privacy concerns related to AI in Canadian healthcare. Refer to Appendix A for search queries and search results.

4.1 Descriptive Analytics

The majority of academic publications discussing the topic in the Canadian context were affiliated with academic or health institutions in Ontario (n=8), Quebec (n=4), and Alberta (n=2). Other sources included Saskatchewan (n=1) and the combined affiliations of Ontario, Quebec, and British Columbia (n=1). All academic articles addressed the three forms of rationality to varying extents and in diverse forms. Some studies predominantly mainly focused on a combination of two types of rationality, yet still provided insights into all three dimensions.

In the case of grey literature, the provinces reviewing the topic included Ontario (n=6), Alberta (n=5), Quebec (n=5), British Columbia (n=4), Manitoba (n=2), Newfoundland and Labrador (n=2), Saskatchewan (n=2), Yukon (n=2) and New Brunswick (n=1). Almost half of this literature addressed the three types of rationality (12 of 29 articles) and the rest focused mainly on a variety of combinations of two rationalities. No relevant information involving all elements of ethics, privacy, and AI applied in Canadian healthcare was found originating from the regions of Prince Edward Island, Northwestern Territories, Nunavut, and Nova Scotia. This gap indicates a possible lack of published data or early development stages of discussions about ethical debates using AI in healthcare in these regions, limiting the review to areas where information was available.

The sources from the academic literature reviewed originated from authors linked to educational and healthcare organizations. The corpus included 11 original research articles, 2 literature review articles, 2 qualitative research studies, and 1 white paper. In contrast, the grey literature came from diverse sources, including 13 instances from institutional and academic publications,
7 from journalistic entities such as trade journals and news outlets, 4 from newspapers, 3 from governmental bodies, 1 from digital magazines and 1 from organizational sources.

The academic literature covers the Canadian context broadly, while grey literature focuses on specific provincial or territorial settings. In Western Canada, discussions emphasize ethics, privacy, bias, regulatory changes, and healthcare job evolution. Central Canada concentrates on the need for better data, policies for AI in healthcare, addressing bias and racism, and re-training healthcare personnel. Atlantic Canada's literature debates around ethical AI laws and policies, professional education on AI, and AI development with ethical, legal, and political considerations. In the North, specifically Yukon, the focus is on privacy laws, reforms, bias, and AI explainability. See Appendix B for references and territorial division clarification.

4.2 Common Structure of the Three Rationalities Across Academic and Grey Literature
This section provides an analysis of the issues explored within both categories of literature, framed through the lens of Habermas's three-dimensional forms of rationality. Additionally, it elaborates on topics raised on each type of rationality, providing a detailed examination of their implications and manifestations including quotations of each subtopic.
Figure 1 illustrates how the concepts of privacy and ethics are not isolated but rather overlap among the three forms of rationality. The intersections between these rationalities indicate how discourses regarding privacy and ethical concerns are inseparable from questions related to the technological efficiency of AI, policy regulation considering social justice and the impact on patients’ experience. This diagram serves as an analytical tool to visually synthesize these complex, intertwined issues, which can be framed using the three-dimensional theory of rationality from Habermas.
Cognitive-Instrumental rationality is represented by issues about the technological efficiency and accuracy of AI in healthcare; Aesthetic-Expressive rationality is focused on the socio-emotional and personal-emotional experiences impacting patients’ privacy, autonomy, and trust; and Moral-Practical rationality can be found in legal, moral, and ethical questions, including regulations and norms to safeguard privacy rights and ensure social justice. Figure 1 also outlines the center of the frequent discussion regarding the importance of balancing technological, socio-emotional, and legal/ethical aspects while integrating AI in healthcare. Furthermore, it also highlights the issue of patient data ownership involving tech companies, patients, and regulatory bodies.

Throughout the review of the existing literature, it became evident that discussions or analyses addressing the spiritual implications of integrating AI into the Canadian healthcare system remain a significant gap. Despite the extensive exploration of ethical, practical, and technological dimensions, there appears to be an absence of discourses on how AI implementation might intersect with or impact spiritual considerations in healthcare settings.

In the following subsections, I detail the specific debates that constitute the rationalities previously explained in Figure 1. The most discussed topics by both types of literature were 1) Bias and equity, 2) Privacy and confidentiality and 3) Legal and policy framework. To enhance the clarity and depth of the discussion, direct quotations from the data are provided for each topic. These quotations serve to ensure that the insights derived are firmly anchored in the empirical evidence collected during the research process. To see references of each rationality and subtopic, consult Appendix C.

4.2.1. Cognitive-Instrumental Rationality Discussions

- **Algorithmic Reliability**: Focuses on the challenges posed by AI errors, limitations, and safety concerns. It encompasses the risk of malfunctioning, the necessity for continuous technological improvement, and the dangers of over-reliance on algorithms for critical decisions, emphasizing the importance of human intervention to ensure proper technical function and to validate the algorithm's results.
“Quality and assessment (statistical testing and validation): What is not scientifically valid cannot be ethical. Correct scientific expertise is needed to ensure that models are robust and validated accurately” (Aiken et al. 2021:2).

“AI algorithms are not infallible and can produce false or misleading information, known as AI hallucinations. These errors can arise from biases in the training data or limitations in the AI’s understanding of complex medical scenarios” (Hryciw et al. 2023:2).

“Model auditing is consistent with a focus on continuous quality improvement, and should collect and retain evidence that the ML-based decision-making tool is safe to use in the intended population” (McCradden, Joshi, Anderson et al. 2020;2026).

- **Data Security Vulnerabilities:** This theme highlights the escalating threats to the confidentiality and integrity of patient data, addressing the frequency and impact of security breaches. It calls for robust protective measures to safeguard sensitive patients’ information against unauthorized access and cyber threats. It also involves anonymization processes to ensure privacy within the databases.

  “There is a risk of breaches through exploitable vulnerabilities in the AI system. This is similar to the risk faced by regular ICT [information and communication technology] systems, but there is an additional risk related to opening an AI system up to public access” (Office of the Information and Privacy Commissioner of Alberta 2023:3)

  “Preparing for potential security breaches, including those that result in reidentification by machine learning algorithms, is a legal obligation of corporate data custodians” (Murdoch, Jandura, & Caulfield, 2022:46)
• **Professional Role Evolution**: Reflects on the changing landscape of healthcare responsibilities and roles due to AI integration. To ensure that patients receive the most precise and efficient care, this topic delves into the discussions of job displacement (especially in the case of doctors), skill adaptation, and the need for healthcare workers to receive constant training to work effectively alongside AI technologies.

“One of the questions that arose during the news conference was whether AI would eventually replace doctors. Park said he did not believe that would ever happen. "I think there's always a fear around emerging technologies that we do need to address," said Park (Gillis 2023, para. 21-22)

“One concern with algorithm-based care is that doctors will spend less time listening to patients, trying to understand the complex social determinants that factor into health, and more time looking at screens” (Glauser 2020:22).

“In parallel with the new ethical challenges presented by AI, it can be expected to change the role of individuals working in healthcare just as other technologies before it have done. How many – and which – of the litany of healthcare professional roles will be altered remains unclear” (Brenna 2021:108).

4.2.2. *Aesthetic-Expressive Rationality Discussions*

• **Acceptance and perception**: Explores patient interactions with AI in healthcare, focusing on feelings, perceptions, and whether they approach AI with trust or skepticism. This includes examining the expressive aspects of technology acceptance and the relationship between patients and AI systems. It contrasts with how physicians’ and healthcare providers' experiences are discussed, usually framing their perspective in terms of delivering services (cognitive-instrumental rationality).
“The issues of public trust and public input into the governance of ML initiatives in health care have been widely discussed as the popularity of AI has grown, with advocates suggesting that future developments of AI ought to be explicitly supporting a broader public interest” (Shaw et al. 2019:5)

“The collection, storage, and use of bulk medical data present additional challenges related to social acceptability and public perception, legislative obstacles, information technology barriers, and the risk of breaches in data security” (Jaremko et al. 2019:108).

“The role of human control (delegating to the machine) may cause concern for patients and publics” (Aiken et al., 2021:2).

- **Data sovereignty**: Under the umbrella of autonomy and emancipation, this theme addresses patients' rights to control their personal health data and how it is used. It is also immersed into the ethical considerations surrounding data sovereignty, emphasizing the importance of empowering patients to manage their health information in the context of AI integration.

  “As health institutions deploy these technologies, the amount and intrusiveness of health data collection may grow; the clinical utility (validated or merely promised) of AI may be used to justify extensive digitization of health status, healthcare management practices, treatment use, and so on, even including data collection via wearable devices or other forms of surveillance” (Racine, Boehlen, and Sample 2019: 272)

  “Patient involvement in the decision-making process is paramount for promoting responsible AI integration in healthcare. By engaging patients with AI-generated insights, physicians can ensure that these are considered alongside human expertise and experience, as well as the patient’s preferences and unique circumstances” (Hryciw et al. 2023:02).
• **Consent complexities:** This theme discusses the new questions emerging based on the integration of AI and consent procedures, like the processes and methodologies to ensure that consent is fully informed and voluntary, and it is revisited as algorithms evolve. This reflects how these processes must align with expressive values and patients’ decisions.

> “With a novel technology such as generative AI, it may be challenging to communicate the risks and benefits of the use of the technology to patients. At a minimum, patients should be asked to consent to record the clinical encounter (if applicable) and made aware of potential risks involving data integrity, bias and privacy” (College of Physicians & Surgeons of Alberta [CPSA] 2023:4).

> “How can patients be informed about possible negative outcomes of AI-based health technologies if we do not really know how the data will be used and what it might reveal? It may be impossible for human minds to foresee all uses of AI and health information. However unforeseeable changes are, institutions should respond to anything that could impact consent and consent processes” (Racine, Boehlen, and Sample 2019: 273).

• **Safety and integrity:** These issues are concerned with ensuring that the integration of AI in healthcare does not compromise patient safety or integrity. They involve assessing and mitigating risks associated with implementing AI in healthcare settings.

> “Unlike other tools, ML might realize its errors and correct its performance issues. But, at the same time, this aspect of ML creates real challenges: if an ML tool is “approved” as safe prior to entry to the market but then evolves, how can a regulator guarantee that it remains safe for use as it evolves?” (Da Silva, Flood, Goldenberg & Singh 2022:65)
“Cost overruns and system inefficiency arising from poor data design and use are a material source of harm in the healthcare system. The use of generative AI to support charting and clinical decision-making has the potential to improve efficiency and access to care while reducing costs. However, it must be carefully evaluated to avoid unintended harm to patients” (College of Physicians & Surgeons of Alberta [CPSA] 2023:5).

- Privacy and confidentiality: Examines the challenges and potential solutions for protecting patient privacy and confidentiality, as well as the eventual need to use patient data to improve research and innovation. It includes anonymization efficacy and is distinct from cognitive-instrumental privacy and data breaches as it focuses on expressive concerns about fundamental rights in healthcare systems.

“While AI presents potential benefits across many domains and in everyday life, the regulators note that there are also risks and potential harms to privacy, data protection, and other fundamental human rights if these technologies are not properly developed and regulated” (Government of Newfoundland and Labrador 2023:30)

“There are potential concerns about anonymity protection even if identifiers are removed from diagnostic images. Appropriate privacy protections are critically important from both ethical and practical (i.e., technological) perspectives. This issue must be foremost in future planning regarding this technology” (Zarzeczny et al. 2021:171).

“We argue that the time has come for meaningful public deliberations about the appropriate balance between privacy risks and the harms of not using data for health research and innovation. We know that patients and their families want to see their health data used to improve care” (Sandusky 2023, para. 6).
4.2.3 Moral-Practical Rationality Discussions

- **Bias and Equity**: This theme questions how AI systems can perpetuate or exacerbate social inequalities due to biases in data and algorithms. It seeks to ensure social justice and fairness in AI healthcare applications by minimizing systemic biases that may disproportionately affect vulnerable populations.

  “Health-related AI is designed by humans, who have explicit and implicit biases. The algorithms that humans develop could accordingly be inadvertently biased against marginalized groups/patients. AI innovators make many design choices that may increase or decrease algorithmic bias risks, such as decisions about whether or not to pursue and build upon data sets that are more representative but may be more costly or otherwise difficult to obtain” (Da Silva, Flood, Goldenberg & Singh 2022:69)

  “For example, marginalized populations — including certain ethnic groups and Indigenous people — are under-represented in national datasets because Canada does not systematically collect information on these groups in a comprehensive way, according to the research team. This is why health equity is a critical factor in AI innovation” (UM Today Staff 2020, para. 8)

- **AI Ethical Integration**: AI, as a non-human agent, inherently lacks human values and ethical reasoning. This issue explores the moral challenges in embedding ethical principles into AI systems and ensuring that these technologies align with human values, particularly in sensitive healthcare contexts where ethical decision-making is crucial.

  “Whether it is ethically appropriate to use AI depends on whether the AI will, in fact, contribute positively to individual patients and/or society” (Jaremko et al. 2019:115)

  “Ethicists play a crucial role in guiding the responsible development and implementation of these advances, allowing for the benefits of innovation
**Legal and Policy Framework**: This theme examines the existing and evolving legal frameworks and policy-making processes concerning AI in Canadian healthcare. Furthermore, it highlights the need for rigorous legal standards and policies that address the issues posed by AI in healthcare. In other words, it seeks to ensure that these technologies operate in a regulated framework and are developed and deployed responsibly.

“AI remains a fairly novel frontier in global healthcare, and one currently without a comprehensive global legal and regulatory framework” (Murdoch 2021:2).

“Public bodies and custodians of personal health information in Newfoundland and Labrador have a responsibility to ensure that their use of AI complies with existing privacy laws, however, those privacy laws will likely prove inadequate to effectively protect the public interest. Governments in Canada and around the world have begun to recognize this, and many have begun to enact specific laws and legislative amendments that are meant to address the unique challenges of AI” (Government of Newfoundland and Labrador 2023:4).

**Liability and Accountability**: Explores who is responsible when AI systems may cause harm or make errors. It reflects on the moral and legal implications of AI decision-making, emphasizing the need for clear accountability mechanisms and policies to protect patient rights and ensure justice.
“It is not yet clear who – if anyone – ought to be accountable for errors made by AI, nor are we fully able to predict all of the consequences which would follow from its displacement of human beings from jobs we now perform” (Brenna 2021:108).

“The question of liability, in particular, needs clarification. To disincentivize potentially harmful applications of AI, it is important that legal obligations do not disappear when machines are replacing humans. For now, it is not entirely obvious from whether legal responsibility for adverse outcomes lies with developers or those deploying the technology” (Ontario Chamber of Commerce n.d.:48).

**Explainability and Transparency**: Deeply connected to the accountability and legal subtopics, concerns of this type stress that we need AI to be transparent in the decision-making process, accentuating that AI technologies should be clear and comprehensible to both healthcare professionals and patients to foster informed decisions and trust.

*AI is only as powerful as the data it’s built on, and often those data reflect historic biases – that’s why many advocate for what’s called “explainability,” which requires solutions to demonstrate how and why they came to a particular conclusion, as opposed to a “black box” approach, in which AI makes determinations with little or no transparency into its decision-making process” (Lindzon 202, para. 21).*

*“Transparency is a cornerstone of the evolving physician-patient relationship in the era of AI-driven healthcare. As AI systems can sometimes be perceived as “black boxes” with their complex decision-making processes, physicians must highlight that while AI can provide useful information, it may not yet consider all relevant factors or nuances of a patient’s unique circumstances that are considered by a human physician” (Hryciw et al. 2023:02).*
4.3 Main Differences Across Grey and Academic Literature

In this section, the distinctions between both types of literature are reviewed, with a specific focus on their divergent emphases and subthemes within the designated rationalities. This comparative analysis aims to discuss the unique characteristics and contributions of each type of literature. To see references of the subtopics on Human error, Data quality assurance and Public education, see Appendix C.

Table 1: Differences Between the Three Forms of Rationality – Types of Issues Raised

<table>
<thead>
<tr>
<th>Forms of rationality</th>
<th>Type of Literature</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Academic Literature</td>
</tr>
<tr>
<td>Cognitive- Instrumental</td>
<td>• Human error</td>
</tr>
<tr>
<td>Aesthetic-Expressive</td>
<td>• More focused on patient safety</td>
</tr>
<tr>
<td>Moral-Practical</td>
<td>• More focused on accountability</td>
</tr>
</tbody>
</table>

Table 1 compares the differences between academic and grey literature with respect to the content discussed based on the three forms of rationality. On the one hand, discourses on patient safety and accountability are more detailed and reiterated more often in the academic literature. As opposed to the grey sources, academic literature frequently highlights how human error can compromise the accuracy of AI technologies in healthcare, particularly through the lack of proper use by health professionals. This issue manifests in two critical ways: underutilization and overreliance. Healthcare providers may underutilize AI technologies due to uncertainty about their workings, thereby missing chances to reduce medical errors. Conversely, overreliance on
these tools can occur if providers fail to consider potential flaws, such as unrepresentative training data, leading to inaccurate patient diagnoses and or treatment responses.

On the other hand, grey literature is more concerned about patients’ perceptions of potentially interacting with AI in health contexts, in addition to the implications for consent procedures. In contrast to the academic literature, grey publications often conduct a more thorough analysis of the importance of data quality. It notes that the effectiveness of AI in healthcare critically depends on high-quality, reliable data. Training AI with inaccurate or indiscriminately collected data not only reduces its effectiveness but also raises the risk of spreading misinformation in an era already fraught with data quality issues.

Ultimately, grey literature underscores the need for legal and health authorities to educate both healthcare workers and the public about AI's potential and limitations, making AI tools explainable and informing the public about how their data are used. This involves demystifying the benefits of using AI effectively and initiating comprehensive education efforts about data practices, thereby ensuring that AI technologies are understood and trusted by all stakeholders.
CHAPTER FIVE: DISCUSSION AND CONCLUSIONS

The aim of this research paper was to examine academic and grey literature to identify and classify the different discussions found in the selected literature regarding the integration of AI in Canadian healthcare. This was achieved by applying Habermas's model of three worlds/three rationalities to analyze the discourses.

In the discussion of integrating AI into Canadian healthcare, it becomes evident that the conceptual framework under consideration is both applicable and insightful, capturing the interconnectedness of various themes. This is particularly true when examining the issues of ethics, privacy, algorithmic reliability, bias and equity, and legal and policy frameworks. Habermas argues that modernization in society is marked by the increasing integration of the three forms of rationality within societal systems (Dickinson 2021, 2020; Johnson 2006). It is possible to distinguish the three forms of rationalities in the literature reviewed: Cognitive-Instrumental-Rationality is represented in the discussion on algorithmic reliability and the development of AI technologies. The focus on improving AI policies, not only for privacy but in the practical implementation of AI in healthcare, demonstrates an engagement with cognitive-instrumental rationality as well, as it emphasizes the need for effective and efficient AI integration in healthcare systems. The push for more reliable algorithms and comprehensive AI policies reflects a modernization effort toward optimizing healthcare delivery through technology.

Aesthetic-Expressive Rationality might be best appreciated through the dynamics and emotions of patients regarding AI, encompassing issues of consent, acceptance and data sovereignty. This approach is relevant as it points out the importance of incorporating individual perspectives and emotional responses into the broader discourse on AI applications in healthcare delivery. The interconnectedness of these themes suggests that addressing one invariably impacts the others, indicating the need for cross-collaboration between disciplines; nonetheless, ethics and privacy emerge as foundational elements to understand broader implications of AI—such as bias, equity, algorithmic design and legal ramifications.

Finally, in the case of Moral-Practical Rationality, the widespread concern for privacy and the development of policies, along with efforts to address bias and ensure equity in AI applications, illustrate a societal movement to greater moral and ethical consideration in the
integration of AI in healthcare. This aligns with Habermas's view of modernization as involving the increasing collaboration of moral-practical understandings.

In the examination of the current literature, it was clear that there is a notable lack of focus on a fourth sphere of discourse: the spiritual aspects related to incorporating AI into Canadian healthcare. The absence of discourses on how AI implementation might impact or problematize spiritual considerations in healthcare settings does not necessarily reflect a lack of potential issues. Instead, this gap may be indicative of the early stages of emerging concerns regarding AI within the healthcare setting. It suggests that as the integration of AI in Canadian healthcare continues to evolve, the complexity and scope of these discussions will likely expand, engaging with religious and spiritual debates. Recognizing these concerns at this initial stage underscores the need for an eventual proactive approach with the spiritual dimensions of healthcare to ensure a holistic view of AI integration, taking into account the spiritual aspects that influence patient care and provider perspectives. This is and will be particularly interesting in a culturally and ethnically diverse country such as Canada.

The rational discourses identified in the literature may reflect how Canadian society is integrating the different forms of rationality in response to the challenges and opportunities presented by AI in healthcare. In this case, the discourses around AI in healthcare in Canada demonstrate an ongoing negotiation between technological advancements, ethical considerations, and social values. This would represent a balance according to Habermas's framework, mainly concerned about the disproportionate focus on instrumentality over other spheres of rationality (Dickinson 2020; Johnson 2006).

Provinces and territories present consensus on the types of issues raised; however, variation in the data is primarily related to the amount of available information rather than the nature of the concerns themselves. Even though there is an uneven distribution of accessible information on this topic across Canada, in general terms, the results show collective national awareness of the need for a comprehensive AI analysis guided by ethical considerations integrated into AI technological development, legislation of policies, and adherence to fundamental rights.

The current Canadian policies on this matter primarily focus on data protection. Moreover, there is a noticeable absence of standardization across provinces regarding data protection policies. These two factors may result in insufficient measures for addressing all potential challenges posed by AI in the healthcare context. Nonetheless, the ongoing discussions and concerns raised
indicate a first step towards addressing these gaps, stating Canada, in general terms, as a country with a proactive yet cautious approach towards the integration of AI into healthcare. The evolving dialogue represents a national aim to apply the potential of AI in improving healthcare but also ensuring that such integration is ethical, socially equitable, and legally solid. This aligns with Romanow’s (2002) cautious vision for the healthcare system adaptation to current and future realities, focused on greater accountability, technological enhancement and ownership of personal health information. From Habermas’ perspective, these findings could represent how rational debates are taking place in Canada, demonstrating the ongoing process of modernization while trying to reconcile the rapid pace of technological advancement with enduring ethical and social values.

This research presents limitations on the generalizability of the findings: the focus on the Canadian healthcare system might limit its applicability to other countries with different healthcare models, technological infrastructures, and privacy laws. A second limitation is related to the rapid technological advances; as AI research and development moves at a fast pace, some examples or technologies discussed in this paper may quickly become outdated, making it challenging to stay relevant. A third limitation includes the absence of data from some provinces, which restricts the ability to present a complete overview. Consequently, the analysis should be treated as a general approximation of the current situation, primarily based on the available information. This may inadvertently lead to an overrepresentation of certain provinces over others.

Future research could aim to further integrate interdisciplinary research that combines perspectives and methodologies from ethics, law, computer science, and healthcare to address the complex challenges of AI in healthcare more comprehensively. It can also expand on the methodological approach to primary data, incorporating interviews, surveys, or case studies, to gather firsthand insights from various stakeholders, including patients who are the recipients of AI-driven care, healthcare professionals who interact with AI systems, and policymakers who govern the use of technology in healthcare. This would allow gaining deeper insights into each form of rationality delineated in this paper.

The conclusions of this paper can be stated in four key points regarding the ethical and privacy considerations of AI in Canadian healthcare, framed within Habermas's theory of three-dimensional rationality. The first one involves modernization and rationality: findings align with
Habermas's view of modernization, illustrating how social involvement with the three forms of rationality reflects the ongoing stage of modernization in Canada. The second one is the interconnectedness of themes, validating the applicability of Habermas's three-dimensional rationality (Cognitive-Instrumental, Aesthetic-Expressive and Moral-Practical) as a comprehensive framework to assess the multifaceted implications of AI in healthcare. The third conclusion is the national awareness and need for policy development in Canada. Findings show a general national consensus regarding the need for a holistic strategy to implement AI in healthcare, considering ethical implications, technological advancements, and policy formulation. Finally, the fourth point defines ethics and privacy as central components to understand broader implications of AI technologies in healthcare, including stages of development, implementation, and governance.
REFERENCES


Dickinson, H. 2021. Class Notes from SOC 303.6 Course (History of Sociological Theory): Modernization as a two-sided rationalization process: The theory of Jurgen Habermas


Faculty of Law, University of Alberta. 2020-2021. Privacy and Artificial Intelligence: Protecting Health Information in a New Era. Health Law Institute, University of Alberta.


Sandusky, K. 2023. CIFAR AI Insights Policy Brief outlines strategies for federated health data access in Canada. CIFAR.


APPENDICES

APPENDIX A: Tables of search queries and search results

ACADEMIC LITERATURE

<table>
<thead>
<tr>
<th>Database</th>
<th>Search Query</th>
<th>Filters</th>
<th>Results</th>
<th>Titles selected</th>
<th>Titles included in the review</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scopus</td>
<td>(&quot;Artificial Intelligence&quot; AND &quot;health*&quot;) AND (privacy OR ethics) AND (Canada OR Canadian)</td>
<td>Fields: Social sciences, health professions, medicine</td>
<td>40</td>
<td>10</td>
<td>5</td>
</tr>
<tr>
<td>Medline</td>
<td>See search queries broken down on next table</td>
<td></td>
<td>3</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Web of Science</td>
<td>(&quot;Artificial Intelligence&quot; AND healthcare) AND (ethics OR privacy)</td>
<td>Fields: Sociology, Social Issues, Social sciences, Healthcare services, medical ethics</td>
<td>64</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>Semantic Scholar</td>
<td>(&quot;Artificial Intelligence&quot; OR &quot;AI&quot;) AND (health*) AND (privacy OR ethics) AND (Canada OR Canadian)</td>
<td>Fields: Sociology, Law</td>
<td>27</td>
<td>7</td>
<td>6</td>
</tr>
</tbody>
</table>

TOTAL                                           16
## Medline search queries

<table>
<thead>
<tr>
<th>No.</th>
<th>Search terms/limits</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>healthcare.mp</td>
<td>320,238</td>
</tr>
<tr>
<td>2</td>
<td>limit 1 to (english language and yr=&quot;2018-2023&quot;)</td>
<td>156,724</td>
</tr>
<tr>
<td>3</td>
<td>artificial intelligence.mp. or Artificial Intelligence/</td>
<td>49,450</td>
</tr>
<tr>
<td>4</td>
<td>limit 3 to (english language and yr=&quot;2018-2023&quot;)</td>
<td>25,054</td>
</tr>
<tr>
<td>5</td>
<td>2 and 4</td>
<td>2284</td>
</tr>
<tr>
<td>6</td>
<td>privacy/ or privacy.mp</td>
<td>191,326</td>
</tr>
<tr>
<td>7</td>
<td>limit 6 to (english language and yr=&quot;2018-2023&quot;)</td>
<td>7840</td>
</tr>
<tr>
<td>8</td>
<td>ethics/ or ethics.mp</td>
<td>191,326</td>
</tr>
<tr>
<td>9</td>
<td>limit 8 to (english language and yr=&quot;2018-2023&quot;)</td>
<td>37,991</td>
</tr>
<tr>
<td>10</td>
<td>7 or 9</td>
<td>44,745</td>
</tr>
<tr>
<td>11</td>
<td>5 and 10</td>
<td>296</td>
</tr>
<tr>
<td>12</td>
<td>Canada.mp</td>
<td>161,168</td>
</tr>
<tr>
<td>13</td>
<td>limit 12 to (english language and yr=&quot;2018-2023&quot;)</td>
<td>41,319</td>
</tr>
<tr>
<td>14</td>
<td>11 and 13</td>
<td>3</td>
</tr>
<tr>
<td>Search Engine</td>
<td>Search Query</td>
<td>Results</td>
</tr>
<tr>
<td>---------------</td>
<td>--------------</td>
<td>---------</td>
</tr>
<tr>
<td>Google</td>
<td>AI in healthcare Saskatchewan privacy ethics</td>
<td>56,600</td>
</tr>
<tr>
<td></td>
<td>AI in healthcare “Saskatchewan” “privacy” OR “ethics”</td>
<td>78,800</td>
</tr>
<tr>
<td></td>
<td>AI in healthcare “Saskatchewan”</td>
<td>310,000</td>
</tr>
<tr>
<td></td>
<td>Artificial Intelligence in health “Alberta” “privacy” OR “ethics”</td>
<td>86,000</td>
</tr>
<tr>
<td></td>
<td>AI in healthcare Alberta privacy ethics</td>
<td>186,000</td>
</tr>
<tr>
<td></td>
<td>AI in healthcare “Alberta” ethics</td>
<td>170,000</td>
</tr>
<tr>
<td></td>
<td>AI in healthcare Ontario privacy ethics</td>
<td>291,000</td>
</tr>
<tr>
<td></td>
<td>AI in healthcare “Ontario”</td>
<td>9,140,000</td>
</tr>
<tr>
<td></td>
<td>Artificial Intelligence in health “Ontario” “privacy” OR “ethics”</td>
<td>413,000</td>
</tr>
<tr>
<td></td>
<td>AI in healthcare Quebec privacy ethics</td>
<td>1,590,000</td>
</tr>
<tr>
<td></td>
<td>AI in healthcare “Quebec” “Artificial Intelligence” OR “AI” in “health” “Quebec” “privacy” OR “ethics”</td>
<td>7,930,000</td>
</tr>
<tr>
<td></td>
<td>AI in healthcare British Columbia privacy ethics</td>
<td>4,830,000</td>
</tr>
<tr>
<td></td>
<td>Artificial Intelligence in health “British Columbia” “privacy” OR “ethics”</td>
<td>141,000</td>
</tr>
<tr>
<td></td>
<td>AI in healthcare Manitoba privacy ethics</td>
<td>69,000</td>
</tr>
<tr>
<td></td>
<td>Artificial Intelligence in health “Manitoba” “privacy” OR “ethics”</td>
<td>106,000</td>
</tr>
<tr>
<td></td>
<td>AI in healthcare “Manitoba”</td>
<td>355,000</td>
</tr>
<tr>
<td></td>
<td>AI in healthcare Nova Scotia privacy ethics</td>
<td>48,000</td>
</tr>
<tr>
<td></td>
<td>Artificial Intelligence in health “Nova Scotia” “privacy” OR “ethics”</td>
<td>87,500</td>
</tr>
<tr>
<td></td>
<td>AI in healthcare “Nova Scotia”</td>
<td>82,000</td>
</tr>
<tr>
<td></td>
<td>AI in healthcare New Brunswick privacy ethics</td>
<td>158,000</td>
</tr>
<tr>
<td>Region</td>
<td>Search Term</td>
<td>Results</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>----------------------------------------------------------------------------</td>
<td>---------</td>
</tr>
<tr>
<td>New Brunswick</td>
<td>Artificial Intelligence in health “New Brunswick” “privacy” OR “ethics”</td>
<td>108,000</td>
</tr>
<tr>
<td></td>
<td>AI in “healthcare” “New Brunswick”</td>
<td>107,000</td>
</tr>
<tr>
<td>Prince Edward Island</td>
<td>Artificial Intelligence in “health care” “Prince Edward Island” “privacy” OR “ethics”</td>
<td>32,500</td>
</tr>
<tr>
<td></td>
<td>AI in “healthcare” “Prince Edward Island”</td>
<td>144,000</td>
</tr>
<tr>
<td>Newfoundland and Labrador</td>
<td>Artificial intelligence in health “Newfoundland and Labrador” “privacy” OR “ethics”</td>
<td>44,300</td>
</tr>
<tr>
<td></td>
<td>AI in healthcare “Newfoundland and Labrador”</td>
<td>148,000</td>
</tr>
<tr>
<td>Northwest Territories</td>
<td>Artificial Intelligence in “health care” “Northwest Territories” “privacy” OR “ethics”</td>
<td>29,100</td>
</tr>
<tr>
<td></td>
<td>AI in healthcare “Northwest Territories”</td>
<td>83,200</td>
</tr>
<tr>
<td>Yukon</td>
<td>Artificial intelligence in “health” “Yukon” “privacy” OR “ethics”</td>
<td>42,600</td>
</tr>
<tr>
<td></td>
<td>AI in healthcare “Yukon”</td>
<td>173,000</td>
</tr>
<tr>
<td>Nunavut</td>
<td>Artificial intelligence in health “Nunavut” “privacy” OR “ethics”</td>
<td>36,200</td>
</tr>
<tr>
<td></td>
<td>AI in healthcare “Nunavut”</td>
<td>113,000</td>
</tr>
</tbody>
</table>
### APPENDIX B: References and territorial division clarification table

<table>
<thead>
<tr>
<th>Area</th>
<th>Provinces and territories</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Western Canada</td>
<td>British Columbia</td>
<td>Kim 2023; Provincial Health Services Authority 2023; Samuels 2022; Aiken et al., 2021; Alberta Innovates 2023; College of Physicians &amp; Surgeons of Alberta. 2023; Office of the Information and Privacy Commissioner of Alberta. 2023; Ireland, 2023); Faculty of Law, University of Alberta 2020-2021</td>
</tr>
<tr>
<td></td>
<td>Alberta</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Saskatchewan</td>
<td>Saskatoon Star Phoenix 2023; Longo &amp; Zarzeczny 2021.</td>
</tr>
<tr>
<td></td>
<td>Manitoba</td>
<td>UM Today Staff 2020; Prychitko &amp; Siddall 2022.</td>
</tr>
<tr>
<td>Central Canada</td>
<td>Quebec</td>
<td>Tomesco 2023; Laverdière &amp; Régis 2023; Caruso-Moro 2022; Serebrin 2022; Risling 2018.</td>
</tr>
<tr>
<td></td>
<td>Ontario</td>
<td>Lindzon 2023; Sandusky 2023; Gillis 2023; Sultan 2022; Glauser 2020; Ontario Chamber of Commerce. (n.d.).</td>
</tr>
<tr>
<td>Atlantic Canada's (excluding PEI and Nova Scotia)</td>
<td>Newfoundland</td>
<td>Government of Newfoundland and Labrador 2023; Health Accord NL 2022.</td>
</tr>
</tbody>
</table>
### APPENDIX C: References based on rationality and subtopics table

<table>
<thead>
<tr>
<th>Type of Rationality</th>
<th>Subtopic (Debates raised)</th>
<th>Number and type of articles discussing</th>
<th>Grey references</th>
<th>Academic references</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cognitive- Instrumental</td>
<td>Algorithmic Reliability</td>
<td>10 grey 12 academic</td>
<td>Tomesco, 2023; Laverdière &amp; Régis 2023; Alberta Innovates 2023; College of Physicians &amp; Surgeons of Alberta (CPSA) 2023; Office of the Information and Privacy Commissioner of Alberta, 2023; Provincial Health Services Authority 2023; Gillis 2023; Caruso-Moro 2022; Ireland 2023; Aiken et al., 2021.</td>
<td>Murdoch, Jandura, &amp; Caulfield, 2022; Da Silva, Flood, Goldenberg &amp; Singh 2022; Zarzeczny et al., 2021; Brenna 2021; Murdoch 2021; Forcier, Khoury, and Vézina 2020; McCradden, Joshi, Anderson et al. 2020; Hardcastle &amp; Ogbogu 2020; Kluge 2020; Racine, Boehlen, and Sample 2019; Shaw et al., 2019.</td>
</tr>
<tr>
<td>Data Security Vulnerabilities</td>
<td></td>
<td>8 grey 7 academic</td>
<td>Lindzon 2023; College of Physicians &amp; Surgeons of Alberta (CPSA) 2023; Office of the Information and Privacy Commissioner of Alberta, 2023; Ireland 2023; Gillis 2023; Prychitko &amp; Siddall 2022; Aiken et al., 2021; Faculty of Law, University of Alberta (2020-2021)</td>
<td>Murdoch, Jandura, &amp; Caulfield, 2022; Da Silva, Flood, Goldenberg &amp; Singh 2022; Murdoch 2021; Mörch, Gupta &amp; Mishara 2020; Racine, Boehlen, and Sample 2019; Shaw et al., 2019; Jaremko et al., 2019.</td>
</tr>
<tr>
<td>Professional Role Evolution</td>
<td></td>
<td>9 grey 3 academic</td>
<td>Kim, 2023; Laverdière &amp; Régis 2023; Saskatoon Star Phoenix 2023; Gillis 2023; Health Accord NL 2022; Longo &amp; Zarzeczny 2021; Aiken et al., 2021; Glauser 2020; Risling 2018.</td>
<td>Hryciw et al., 2023; Brenna 2021; Racine, Boehlen, and Sample 2019</td>
</tr>
<tr>
<td>Data quality assurance</td>
<td></td>
<td>6 grey</td>
<td>Saskatoon Star Phoenix 2023; Kim 2023; Ireland 2023; Sultan 2022; Aiken et al., 2021; Glauser 2020.</td>
<td></td>
</tr>
<tr>
<td>Human Error</td>
<td></td>
<td>4 academic</td>
<td>-</td>
<td>Da Silva, Flood, Goldenberg &amp; Singh 2022; Da Silva et al., 2022; Forcier, Khoury, and Vézina 2020; McCradden, Joshi, Anderson et al. 2020.</td>
</tr>
<tr>
<td>Aesthetic-Expressive</td>
<td>Acceptance and perception</td>
<td>5 grey</td>
<td>Gillis 2023; Saskatoon Star Phoenix 2023; Sultan 2022; Aiken et al., 2021; Ontario Chamber of Commerce. (n.d.).</td>
<td></td>
</tr>
<tr>
<td>---------------------</td>
<td>--------------------------</td>
<td>--------</td>
<td>-------------------------------------------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>10 academic</td>
<td>Hryciw et al., 2023; Da Silva et al., 2022; Brenna 2021; McCradden et al., 2020; Hardcastle and Ogbogu 2020; Kluge 2020; Forcier, Khoury, and Vézina 2020; Shaw et al., 2019; Racine, Boehlen, and Sample 2019; Jaremko et al., 2019.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Data sovereignty</td>
<td>6 grey</td>
<td>Alberta Innovates 2023; College of Physicians &amp; Surgeons of Alberta (CPSA) 2023; Serebrin 2022; Longo &amp; Zarzeczny 2021; Faculty of Law, University of Alberta (2020-2021); Samuels 2020.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>5 academic</td>
<td>Hryciw et al., 2023; Zarzeczny et al., 2021; Murdoch 2021; Racine, Boehlen, and Sample 2019; Jaremko et al., 2019.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Consent complexities</td>
<td>4 grey</td>
<td>College of Physicians &amp; Surgeons of Alberta (CPSA) 2023; Government of Newfoundland and Labrador, 2023; Ireland 2023; Faculty of Law, University of Alberta (2020-2021).</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2 academic</td>
<td>Racine, Boehlen, and Sample 2019; Shaw et al., 2019.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Safety and integrity</td>
<td>2 grey</td>
<td>College of Physicians &amp; Surgeons of Alberta (CPSA) 2023; Aiken et al., 2021.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>5 academic</td>
<td>Da Silva, Flood, Goldenberg &amp; Singh 2022; Zarzeczny et al.; 2021; Forcier, Khoury, and Vézina 2020; McCradden, Joshi, Anderson et al. 2020; Kluge 2020.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Privacy and Confidentiality</td>
<td>12 grey</td>
<td>Saskatoon Star Phoenix 2023; Laverdière &amp; Régis 2023; Sandusky 2023; College of Physicians &amp; Surgeons of Alberta (CPSA) 2023; Office of the Information and Privacy Commissioner of Alberta, 2023; Government of Newfoundland and Labrador 2023; Gillis 2023; Serebrin 2022; Prychitko &amp; Siddall 2022; Chalke, McEvoy &amp; McLeod-McKay 2021; Faculty of Law, University of Alberta (2020-2021); Samuels 2020.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>15 academic</td>
<td>Hyciwi et al., 2023; Murdoch, Jandura, &amp; Caulfield, 2022; Da Silva, Flood, Goldenberg &amp; Singh 2022; Da Silva et al., 2022; Zarzeczny et al., 2021; Murdoch 2021; McCradden et al., 2020; Mörch, Gupta &amp; Mishara 2020; Kluge 2020; Racine, Boehlen, and Sample 2019; Shaw et al., 2019; Jaremko et al., 2019.</td>
<td></td>
</tr>
<tr>
<td>Moral-Practical Bias and Equity</td>
<td>16 grey</td>
<td>13 academic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------------------------</td>
<td>---------</td>
<td>------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Saskatoon Star Phoenix 2023; Laverdière &amp; Régis 2023; Lindzon 2023; College of Physicians &amp; Surgeons of Alberta (CPSA) 2023; Office of the Information and Privacy Commissioner of Alberta, 2023; Ireland 2023; Kim 2023; Gillis 2023; Caruso-Moro 2022; Serebrin 2022; Chalke, McEvoy &amp; McLeod-McKay 2021; Longo &amp; Zarzeczny 2021; Aiken et al., 2021; Samuels 2020; UM Today 2020; Glauser 2020.</td>
<td>Hryciw et al., 2023; Murdoch, Jandura, &amp; Caulfield, 2022; Da Silva, Flood, Goldenberg &amp; Singh 2022; Da Silva et al., 2022; Zarzeczny et al., 2021; Brenna 2021; Murdoch 2021; McCradden, Joshi, Anderson et al. 2020; McCradden et al., 2020; Mörch, Gupta &amp; Mishara 2020; Kluge 2020; Shaw et al., 2019; Racine, Boehlen, and Sample 2019.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>AI Ethical Integration</th>
<th>9 grey</th>
<th>6 academic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Saskatoon Star Phoenix 2023; Lindzon 2023; Kim 2023; Government of Newfoundland and Labrador 2023; Canadian Healthcare Technology 2023; Alberta Innovates 2023; College of Physicians &amp; Surgeons of Alberta (CPSA) 2023; Aiken et al., 2021; Samuels 2020.</td>
<td>Brenna 2021; Mörch, Gupta &amp; Mishara 2020; McCradden, Baba, Saha, et al. 2020; Kluge 2020; Racine, Boehlen, and Sample 2019; Jaremko et al., 2019.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Explainability and Transparency</th>
<th>7 grey</th>
<th>8 academic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Topic</td>
<td>Grey</td>
<td>Academic</td>
</tr>
<tr>
<td>----------------------------------</td>
<td>------</td>
<td>----------</td>
</tr>
<tr>
<td>Legal and Policy Framework</td>
<td>14</td>
<td>12</td>
</tr>
<tr>
<td>Liability and Accountability</td>
<td>4</td>
<td>10</td>
</tr>
<tr>
<td>Public Education</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Data ownership</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Balance</td>
<td>5</td>
<td>4</td>
</tr>
</tbody>
</table>