UNIVERSITY OF SAN FRANCISCO

CHANGE THE WORLD FROM HERE

Nonprofit Corporate Governance: How Can Nonprofit Corporate Mechanisms Govern Artificial Intelligence?

by
Linda A. Logemann
llogemann@usfcs.edu

Capstone Research Report Submitted in Partial Fulfillment
of the Requirements for the
Master of Nonprofit Administration Degree
in the School of Management
directed by Dr. Marco Tavanti

San Francisco, California
Spring 2024

Abstract

Artificial intelligence (AI) has the potential to rapidly transform economic and societal norms at a pace unseen by previous technologies, resulting in cross-sector discussions about its benefits and utility that also include worrisome discussions about its risks and ethical challenges in a social context. Currently, there is a gap in the literature about the role of nonprofit corporate governance structures and mechanisms in the development of nonprofit AI applications making this study timely and relevant. This paper takes an exploratory approach by asking the question, how can nonprofit corporate mechanisms govern AI? The discussion of technology will remain at a high, abstract level. To frame the discussion, the paper will begin with defining terms and concepts and asking the key questions relevant to the papers subject matter. In the data analysis section, the data are presented in an understandable visual framework to add clarity to the implications and recommendations for this study.

Acknowledgments

This paper would not exist without Prof. Louise Carroll. Her guidance, encouragement, wisdom, and intellect kept me going and, mostly, on track. I would also like to thank my expert interviewees for their invaluable depth of knowledge, wisdom, insights: Mr. Gene Takagi, Dr. Brian P. Green, and Mr. Phil Albert. And, most importantly, with deep respect and admiration to the final USF Masters Nonprofit Administration cohort, Abdiel, Christine, Norma, and Sally, thank you!

Table of Contents

Section 1. Introduction 1
Section 2: Terms and Concepts 2
Section 3: Key Questions 3
Section 4: Literature Review 4
Overview 4
Artificial Intelligence 4
Corporate Mechanisms 6
Organizational Failure Framework 7
Summary 9
Section 5: Methodology & Approach 10
Overview 10
Literature Review 10
Expert Interviews 11
Qualitative Analysis 11
Section 6. Data Analysis 12
Overview 12
Phase One: Data Analysis 12
Relevant Nonprofit Corporate Mechanisms 13
Risk Mitigation of Misaligned AI Applications 13
Extreme Capital to Build AI Models and Applications 14
Nonprofit Mission Alignment Training Data 14
Phase Two: Corporate Mechanisms Analysis 16
Phase Three: Visual Data Analysis 16
Summary 20
Section 7: Implications and Recommendations 21
Implications 21
Recommendations 21
Section 8: Conclusions 23
Section 10. Limitations 24

References 25
Author's Bio 28

List of Tables

Table 1. Thematic Data Analysis Results p.15

Table 2. Nonprofit Corporate Mechanisms and the Swiss Cheese Model. p.18

List of Figures

Figure 1. Swiss Cheese Accident Causation Model (Reason, 1990, 2000) p. 8

Figure 2. Taxonomy for Artificial Intelligence Hazard Analysis (TAIHA) p. 9

Figure 3: Visual Data Analysis p. 20

Section 1. Introduction

Since November of 2022, when ChatGPT-3 became available for public use, many corporations have begun to incorporate AI technology into their organizations. This includes nonprofit organizations developing artificial intelligence (AI) applications to expand the organization's mission. (Hadley, et al., 2023, Google.org, 2024) Because AI has long-term implications to be both a transformational opportunity and a complex challenge, it necessitates a deep dive into the governance mechanisms that can effectively oversee its integration. (Papagiannidis, 2023). This capstone explores the multidimensional implications of the development of AI applications within nonprofits organizations by asking the research question, how can nonprofit corporate mechanisms govern AI?

As AI becomes increasingly available, leaders and board members face pivotal questions: How can AI further the mission of the organization? What are the potential impacts on the organization's operations? How should they prepare for AI implementation?

The structure of this capstone is designed to comprehensively address these inquiries through a literature review, expert interviews, and a theory development approach. The literature review identifies a research gap at the intersection of AI and nonprofit governance. This research attempts to bridge this gap and to offer a visual framework that aids in the communication of the findings in an accessible and actionable way for nonprofit leaders.

Section 2: Terms and Concepts

The field of artificial intelligence is rapidly evolving, as are definitions of terms and concepts used by technologists, organizational leaders, and the public. Therefore, having an agreed upon set of definitions to relevant terms and concepts associated with AI is essential to engaging any study or discussion about the technology.

Artificial Intelligence (AI): "AI can be understood as a collection of techniques that to some degree seek to mimic human capabilities (e.g., language recognition, movement through robotics) or develop alternatives that substitute for human decisions and/or actions. "(Niederman & Baker, 2023).

AI Applications: Examples of AI applications are ChatGPT, speech recognition, customer service, computer vision, supply chain robotics, weather forecasting (IBM, 2024)

AI Alignment: "AI alignment considers how we can encode AI systems in a way that is compatible with human values." (Hou & Green, 2023).

AI Governance: "According to Butcher and Beridze (2019), AI governance 'can be characterized as a variety of tools, solutions, and levers that influence AI development and applications.'" (Papagiannidis, 2023)

AI Misalignment: AI misalignment occurs when AI systems act in ways that do not align with human intentions, beliefs, and/or values that result in unintended and/or harmful outcomes.

AI Model: "An AI model is a program that has been trained on a set of [internet-based] data to recognize certain patterns or make certain decisions without further human intervention." (IBM, 2024)

AI Training Data: Training data is '…incorporating proprietary content into an…" AI model. (Davenport & Alavi, 2023).

Section 3: Key Questions

As nonprofit leaders and board members increasingly seek the integration of artificial intelligence into their organizational strategies, a series of questions can help frame a discussion around the technology. Nonprofit leaders are asking their own questions, such as, "Is AI too risky for our mission?", "Will AI compromise our values?", and "Can we control AI's impact?" (Winer, 2024) The questions below are an attempt to help leaders ask new questions they may have not thought about. I hope to answer a few of the questions in the sections below with some left unanswered for future research.

1. What is AI? Is there a single definition?
2. How can I prepare my organization for AI?
3. Why is AI risk mitigation important to my nonprofit?
4. What is AI governance?
5. What is an AI application?
6. What is AI alignment?
7. What is AI misalignment?

Building on these foundational questions, the study that follows provides insights to aid nonprofit leaders in incorporating AI applications into their organization.

Section 4: Literature Review

Overview

The literature review explores how nonprofit corporate mechanisms govern AI. The literature review includes three areas: (a) artificial intelligence (AI), (b) corporate mechanisms, and (c) an organizational failure framework. This review will be done to provide a functional level of understanding of the material as it relates to the research subject matter, research question, and data analysis. This literature review is based on 18 sources, including books, journal articles, peer reviewed research, forthcoming research, and various online sources.

An initial review of literature exposed gaps in the research on the study of the intersection of AI and nonprofit governance. As a result, the review required multiple thematic lenses. The primary lens of the literature review was done with the objective of answering the research question, how can nonprofit corporate mechanisms govern AI? Two sub-lenses, nonprofit corporate mechanisms and AI application development and deployment were used to gain a comprehensive understanding of each subject matter. As a theory emerged from the data, a third sub-lens, organizational failure framework, was used to explore a visual framework to communicate the research findings.

Artificial Intelligence

An education computer science and a background in professional software development gave me the ability to comprehend terms, concepts, implementations, and implications of AI models and applications. This skill set offered benefits and limitations to this research subject matter. The benefits were the ability to synthesize complex AI technologies including its implications for an

organization. The limitations were my developing a positional point of view of assumed knowledge that nonprofit leaders or board members may not possess that had the potential to result in conceptual leaps in analysis. A review of the AI literature produced by technical journals fell outside of the scope of this capstone project but did influence my perspective and insights for this research.

The concept of AI began in the 1950 when Alan Turing, "often referred to as the father of computer science, [created the] Turing Test, where a human interrogator would try to distinguish between a computer and human text response… [it is] an important part of the history of AI, as well as an ongoing concept within philosophy…" (IBM, 2024)

In November 2022, when ChatGPT-3 became available for public use, the term of generative AI become widely known, but not well defined. "Generative AI refers to deep-learning models that can take raw data—say, all of Wikipedia or the collected works of Rembrandt—and 'learn' to generate statistically probable outputs when prompted. At a high level, generative models encode a simplified representation of their training data and draw from it to create a new work that's similar, but [hopefully] not identical, to the original data." (IBM, 2024) "As to the future of AI, when it comes to generative AI, it is predicted that foundation models will dramatically accelerate AI adoption in enterprise." (IBM, 2024). Some argue that future is already here. (Arcas, 2023)

> "Many companies are experimenting with ChatGPT and other large language or image models. They have generally found them to be astounding in term of their ability to express complex ideas in articulate language. However, most users realize that these systems are primarily trained on internet-based information and can't respond to prompts

> or questions regarding proprietary content or knowledge…The most common approach to customize content…the orginal [AI] model is kept frozen, and is modified through prompts in the context window that contain domain-specific knowledge." (Davenport & Alavi, 2023)

Corporate Mechanisms

A corporate mechanism refers to the organizational structures, processes, policies, and operational strategies used to manage, govern, and sustain the corporation. (Renz, 2016). "As businesses adopt Artificial Intelligence (AI), they are faced with new value propositions..." (Papagiannidis, 2023) Nonprofit mechanisms such as governance by board of directors, oversight by committees, leadership by the executive director, and policy enforcement by individual employees are all responsible to ensure the organization's corporate compliance, and these activities must be aligned with the organization's mission. (G. Takagi, April 24, 2024, personal communication).

A study by Schuett, Reuel, and Carlier (2024) on designing AI ethics boards offers comprehensive insights into the formation and function of such boards within AI companies developing and deploying AI models. The paper highlights the necessity of having a structured approach to ethical oversight aimed at mitigating societal risks associated with AI technologies and its complexities. A critical takeaway from the study is the detailed exploration of the AI ethics board's potential functions, which include providing insights to leadership overseeing AI development and deployment and interpreting ethics principles to ensure mission alignment and guard against potential harms.

> "If the company already has an internal audit function, the additional value of an ethics board with similar responsibilities would be limited. It would merely provide an additional defense layer. However, if the company does not have an internal audit function, the added value of an [AI] ethics board could be significant. Without a deliberate attempt to identify ineffective risk management practices, some limitations will likely remain unnoticed. But the value ultimately depends on the individuals who conduct the assessment and their ability and willingness to identify ineffective risk management practices. This might be problematic because conducting these kinds of assessments requires a rare combination of specific expertise and character traits (e.g. objectivity, truth-seeking, and epistemic honesty) that not many board members will have. Assessing risk management practices is also very time-consuming, so a part-time board might not be able to do it properly." (Schuett, 2024)

This research supports the need for boards to be equipped with the appropriate authority and resources to perform their duties effectively, ensuring that AI applications are developed in line with the organization's mission and values. (Schuett, 2024)

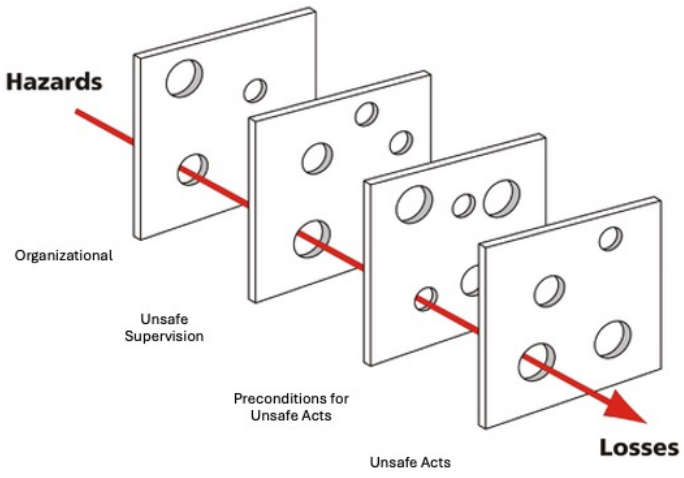
Organizational Failure Framework

The most significant shift in the iterative literature review process occurred during an expert interview with Dr. Green suggesting the Swiss Cheese Model as a framework to structure my data analysis.

Reason's Swiss Cheese Model (SCM) of accident causation is a risk analysis and management framework that illustrates how multiple layers of defense, like slices of Swiss cheese, can prevent hazards resulting in losses to the organization. (Cummings, 2023, Larouzee,

2020, Reason, 1990, 2000) Acknowledging no defense is infallible, each layer has vulnerabilities illustrated as holes. It is only when these holes align that hazards can fully penetrate through the defenses, causing a bad outcomes. “The model draws upon a general, easy-to-remember and adaptable graphical representation that makes it easy to visualize (and conceptualize) the notion of the organizational accident.” (Larouzee, 2020). The model “…focuses uniquely on human actives that lead to accidents...” (Cummings, 2023).

Figure 1: Swiss Cheese Accident Causation Model (Reason, 1990, 2000)



A study by M.L. Cummings (2023) offered strong support to the developing theory emerging from the iterative literature review and ongoing thematic data analysis. The study examined three autonomous vehicle case studies “…to determine if and how AI contributes to accidents.” The research showed that “… the Swiss Cheese model can be updated to better

understand how an AI-enabled system's ability to cope with uncertainty leads to accidents, and what can be done to mitigate such problems." (Cummings, 2023). The study proposed a new Taxonomy for AI Hazard Analysis that "…modifies the organizational level and replaces the lower three levels in the original model layer that look at the design, maintenance, and testing of AI." (Cummings, 2023) See Figure 2.

Figure 2: Taxonomy for Artificial Intelligence Hazard Analysis (TAIHA) (Cummings, 2023

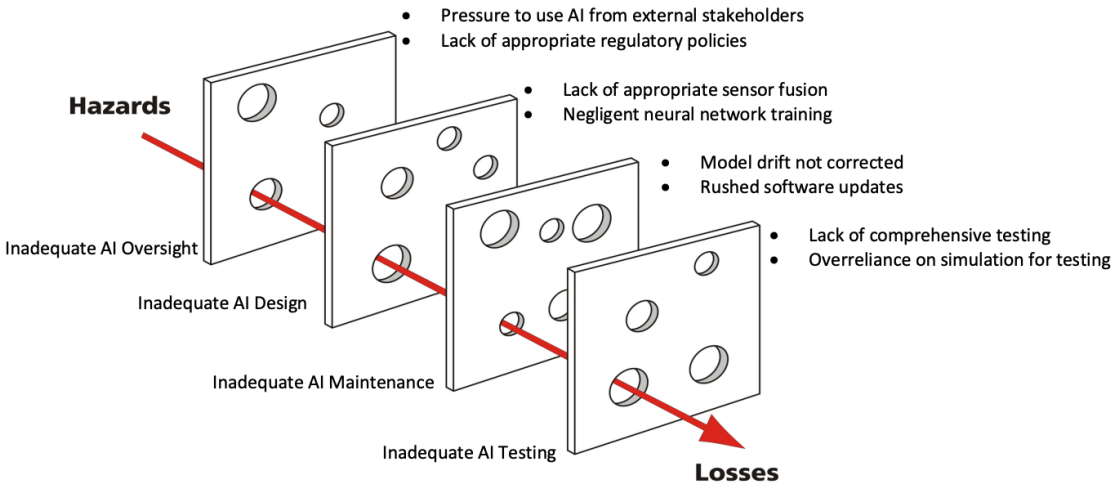


Figure 3: Taxonomy for Artificial Intelligence Hazard Analysis (TAIHA)

Summary

The literature review investigates the governance of AI by nonprofit organizations, covering three key areas: AI technology, corporate governance mechanisms, and an organizational failure framework. To address research gaps at the intersection of AI and nonprofit governance, a multi-lens analytical approach is used. The review looks at how nonprofits can govern AI by the establishment of an AI ethics board and its implications for organizational compliance and mission alignment. The adaptation of Reason's Swiss Cheese

Model to understand AI-related accidents and mitigate risks serves as a crucial tool in understanding how misaligned governance mechanisms can result in organizational failures in the context of internal AI application development and deployment. (Larouzee, 2020).

Section 5: Methodology & Approach

Overview

The research design was grounded in theory. Data was collected from the literature and expert interviews to find patterns and themes to analyze a theory that began to emerge from the data. Due to the complexity of the subject matter, the approach was iterative in data collected from the literature and interview transcripts.

Literature Review

The literature review method was used to gain a comprehensive understanding of AI technologies, nonprofit corporate mechanisms, and a visual framework to communicate a theory that emerged from the data analysis. A focused literature review was done to prepare for each expert interview. After a qualitative analysis of interview transcripts, a deeper and boarder review of the literature was done. The literature review section and the data analysis section provide more insight into these methods and outcomes.

Expert Interviews

Expert interviews were critical to gaining insights into establishing a coherent understanding of the subject matter. Three interviews were completed: Phil Albert, Partner, Intellectual Property, Haynes and Boone, LLP, Brian P. Green, Ph.D., Director of Technology Ethics, Santa Clara University, Gene Takagi, Principle, Neo Law, LLC. The semi-structured and unstructured interviews were with individuals with different areas of expertise relevant to the study. The interviews were 60 to 75 minutes long and conducted over Zoom. Questions were developed and revised iteratively as study progressed. All interview participants were knowledgeable about contemporary AI technologies.

Qualitative Analysis

Data collection was done using a qualitive thematic analysis of the interview transcripts combined with qualitative analysis of the literature. The analysis was conducted on the literature review and expert interview transcripts. Using an iterative process, themes were coded, sub-coded, refined, and analyzed to find repetitive patterns in the cross-sectional data. Themes were identified and coded based on frequency and relevance. Four themes emerged from the data: relevant nonprofit corporate mechanisms, risk mitigation to misaligned AI applications, the extreme capital requirements to build AI models and applications, and the alignment of nonprofit mission to AI application training data.

Section 6. Data Analysis

Overview

The data analysis is presented in three phases, each developed through the iterative data collection process that built upon the insights gained from the previous iterations to deepen and clarify understanding of AI governance in nonprofit organizations. Four themes from the initial data analysis emerged: Relevant Nonprofit Corporate Mechanisms, Risk Mitigation of Misaligned AI Applications, Extreme Capital to Build AI Models and Applications, Nonprofit Mission Alignment Training Data. The second phase takes a deeper dive into answering the research question, how can nonprofit corporate mechanisms govern AI? The third phase utilizes an organizational failure framework to establish a cohesive argument for how the findings contribute to understanding the governance of AI by nonprofit settings.

Phase One: Data Analysis

Four themes emerged from the data supporting the findings from the literature review. (See Table 1) First, nonprofit boards must ensure AI applications align with their mission and comply with legal and ethical standards to maintain organizational trust and tax-exempt status. (G. Takagi, April 24, 2024, personal communication) Second, risk mitigation strategies focus on setting clear boundaries and employing evaluations to prevent misaligned AI applications, allowing for safe innovation. (Papagiannidis, 2023, P. Albert, April 24, private communication) Third, the integration of the nonprofit's mission into AI training data is crucial, necessitating detailed mission statements to guide AI's development and deployment effectively, ensuring

alignment with core organizational values. (Davenport & Alavi, 2023, P. Albert, April 24, 2024, private communication)

Relevant Nonprofit Corporate Mechanisms

In nonprofit corporations, the board of directors holds the ultimate authority, (G. Takagi, April 24, 2024, private communication) shaping the organization's strategic direction and ensuring compliance with both legal and ethical standards, specifically in the development of AI applications.(P. Albert, April 24, private communication) The value of an AI ethics board lies in its members' willingness and ability to actively pursue and uphold the organization's mission, ensuring that the development and deployment of AI applications adhere to both ethical guidelines, and legal requirements. (Schuett, 2024, B.P. Green, April 12, private communication, P. Albert, April 24, private communication) This alignment is vital for maintaining trust and integrity within the nonprofit's operations and external relationships. When integrating AI technology, it is imperative that these entities consider the ethical implications alongside the legal requirements to safeguard the organization's integrity and public trust. This setup ensures that the organization’s AI initiatives align with its core values while meeting regulatory standards.

Risk Mitigation of Misaligned AI Applications

The most dominant theme in the data is risk mitigation of misaligned AI applications. In defining risk mitigation, rules are often framed negatively. Negative framing provides the freedom to explore and innovate by clearly specifying prohibited actions, "This is the area where you can't go. You can do everything else." (B.P. Green, April 12, 2024, personal communication).

Evaluation also stood out as a pivotal tool ensuring that applications are not misaligned but adhere to ethical standards and function as intended without unforeseen consequences. "...evaluation has emerged as a critical mechanism for safe and responsible development, release, and governance of generative AI systems." (Weidinger & et al., 2024) The combination of evaluation and clearly delineated rules provides guardrails for AI applications setting boundaries on potential harms while allowing for the development of AI applications. "Most rules are framed negatively because it gives you more freedom… this is the area where you can't go. You can do everything else." (B. P. Green, April 12, 2024, personal communication)

Extreme Capital to Build AI Models and Applications

The development of AI models and applications demands extreme capital investment. and promise of substantial returns. As these applications evolve, the revenue generated has the potential to come from activities that are not mission related. The "... the limitations of the nonprofit governance model are that it is quickly co-opted by the money." (G. Takagi, April 24, 2024, personal communication)

Nonprofit Mission Alignment Training Data

To ensure that AI applications align with nonprofit missions, it is crucial to infuse the organization's mission into the training data of the AI application. "The mission needs to be infused in the training of the AI [application]." (P. Albert, April 24, 2024, personal communication) As AI models and applications inherently develop distinct behaviors and perspectives based on their training data, (Davenport & Alavi, 2023) aligning these behaviors

and perspectives with the nonprofit's mission is crucial. "... your AI [model] will have a particular point of view." (P. Albert, April 24, 2024, personal communication)

Therefore, to maintain alignment and legal clarity, the mission statement of a nonprofit may need to be detailed and potentially complex, serving as a definitive guide for the AI's development and deployment strategies. "Values of an organization, if not codified, don't really mean anything legally, right? So, the values are going to have to be codified within the mission. So, you might have a very unruly mission statement for legal purposes." (G. Takagi, April 24, 2024, personal communication).

Table 1: Thematic Data Analysis Result

Relevant Nonprofit Corporate Mechanisms	Risk Mitigation Of Misaligned AI Models	Extreme Capital Requirements To Build AI Models	Nonprofit Mission Alignment With Custom Training Data
Ethics board "value depends on it members and their willingness and ability to pursue it mission." (Schuett, 2023)	Compliance with government AI policies and regulatory rules with help mitigate risk.	AI technologies have the potential for significant return on investment.	"The mission needs to be infused in the training of the AI [model]." (Albert, personal communication)
"The board of directors has ultimate power in a nonprofit corporation." (Takagi, personal communication)	"...evaluation has emerged as a critical mechanism for safe and responsible development, release, and governance of generative AI systems." (Weidinger, 2024)	Most earned revenue will come from activities that are not mission-related.	"... your AI [model] will have a particular point of view." (Albert, personal communication)
"...legality and ethics need to be taken into account when you build AI [models]." (Albert, personal communication)	"Most rules are framed negatively because it gives you more freedom… this is the area where you can't go. You can do everything else." (Green, personal communication)	"... the limitations of the nonprofit governance model is that it is quickly co-opted by the money." (Takagi, personal communication)	"Values of an organization, if not codified don't really mean anything legally right? So, the values are going to have to be codified within the mission. So, you might have a very unruly mission statement for legal purposes." (Takagi, personal communication)

Phase Two: Corporate Mechanisms Analysis

Following the thematic analysis on collected data in the development of AI applications in nonprofit organizations, this section takes a deeper dive into answering the research question, how can nonprofit corporate mechanisms govern AI?

The thematic analysis showed the pivotal role of nonprofit boards of directors in their responsibility for setting the strategic direction and ensuring that the deployment of AI technologies adheres to stringent ethical and legal standards. (G. Takagi, April 24, 2024, personal communication, P. Albert, April 24, 2024, personal communication) This is particularly critical in the deployment of AI applications, where ethical dilemmas and regulatory compliance are at the foreground. The need to establish an AI ethics board (Schuett, 2024) within these organizations reenforces the importance of ensuring that AI technologies are developed and implemented in ways that align with both the organization's core values.

Furthermore, the thematic analysis reveals that risk mitigation strategies for AI applications are equally significant to address potential misalignments that could jeopardize ethical standards, or potentially cause harm to beneficiaries of nonprofit services. (P. Albert, April 24, 2024, personal communication) The data suggest that effective risk management involves setting strict boundaries defined through negative framing or what can't be done. Papagiannidis, 2023) "Most rules are framed negatively because it gives you more freedom… this is the area where you can't go. You can do everything else." (B.P. Green, April 24, 2024, personal communication)

Phase Three: Visual Data Analysis

To establish a coherent linkage between the data analysis and the Swiss Cheese Model (SCM) Table 2 offers an analysis of the various governance mechanisms through which nonprofit

organizations can oversee the implementation and integration of AI application into their operations. The table draws from both literature review and expert interviews to illustrate how a layer in the Swiss Cheese Model adds protection and reveals vulnerabilities. The "Board of Directors," "Legal," "AI Ethics Board," and "Employee Inputs" are identified as the critical corporate mechanisms.

> “So, you have legal is the number one thing that's used right now, where somebody says that's against the law. You can't do that. You also have the judgment of individual engineers. So, an individual engineer says, oh. we don't want to make a product that breaks. So, we need to have another safety system here. And then, of course, there's the leadership, the corporate leadership, the tone that's set from the top. Is this an organization that takes risks?” (B.P. Green, April 12, 2024, personal communication).

The application of the SCM makes clear these components are not only important in terms of their individual responsibility, but also in how they interact with and support one another in the broader context of the organization. Quotes from experts and the relevant literature emphasize their roles and the expectations placed upon them in the safe and ethical deployment of AI applications. For instance, the "Board of Directors" is emphasized as having the ultimate decision-making power within a nonprofit (G. Takagi, April 25, 2024, private communication), a sentiment echoed by insights from the literature which describes their role as central to organizational success. (Renz, 2016). Similarly, the establishment of an AI Ethics Board is noted as crucial for maintaining ethical oversight, emphasizing the need for board members to be

actively engaged and committed to the nonprofit's mission. (Renz, 2016, P. Albert, April 24, 2024, personal communication).

Table 2: Nonprofit Corporate Mechanisms and the Swiss Cheese Model

Relevant Nonprofit Corporate Mechanisms	**Literature Review & Expert Interview**	**Swiss Cheese Model Risk Mitigation Slice**	**Relationship**
Board of Directors	"The board of directors has ultimate power in a nonprofit corporation." (Takagi, April 24, 2024, personal communication)	Organizational processes that state goal, assess, organize, and manage (Larouzee, 2023)	"Nonprofit "governance is a central and essential component of the leadership of nonprofit organizations, and the board of directors that engage in the work of governance are central to the success of the organizations they serve," (Renz, p. 163)
Legal	"...legality and ethics need to be taken into account when you build AI [applications]." (Albert, April 24, 2024, personal communication)	Unsafe supervision that create conditions for violations. (Larouzee, 2023)	
AI Ethics Board	Ethics board "value depends on its members and their willingness and ability to pursue it mission." (Schuett, 2023)	Preconditions for unsafe acts is the failure to evaluate	"… evaluation has emerged as a critical mechanism for safe and responsible development, relea
Employee Inputs		Unsafe acts are intended and unintended actions	

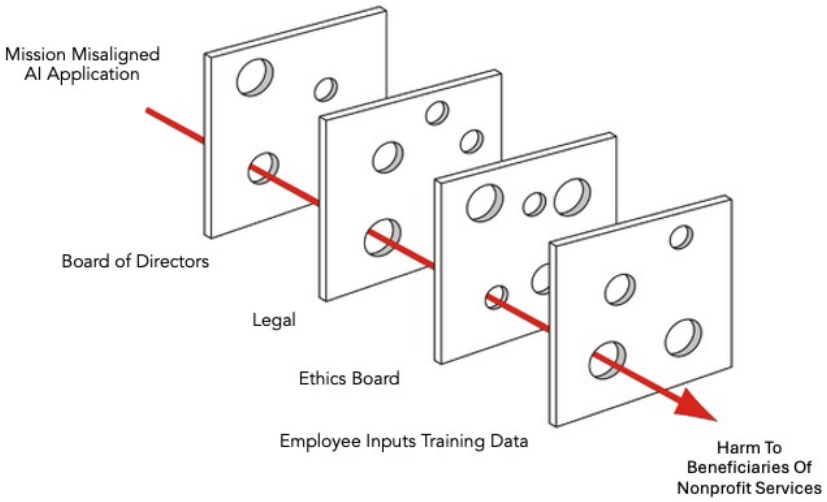
As seen in Figure 3, The Board of Directors, Legal Department, AI Ethics Board, and individual employees are all depicted as layers of defense against the risks associated with AI application deployment. However, as shown in the graphic, each layer, like slices of Swiss cheese, has its holes or weaknesses, which can align to create a pathway for risks to manifest as actual harm to the organization's mission or its beneficiaries. This alignment of holes across multiple governance layers signifies the need for a more robust, interlinked approach where the

gaps in one layer are covered by strengths in another. When the holes are in alignment an AI application misaligned with the nonprofit's mission can lead to potential harm to beneficiaries of organization's services. (Larouzee, 2020).

From this analysis, it is evident that while each mechanism individually provides certain controls, their effectiveness is contingent upon the organization's ability to recognize and mitigate gaps proactively. For instance, the Board of Directors might set the strategic direction and oversee AI integration, but without proper legal oversight and an ethical framework, as emphasized by the literature and expert opinions, there might be unintended negative outcomes. Similarly, the input from individual employees can be a vital layer of defense, but only if they have the resources and an awareness of the broader ethical standards within the organization. (Griffin, 2024) Thus, the Swiss Cheese Model not only emphasizes the risks of misaligned AI applications but also stresses the importance of a comprehensive, multi-layered approach to governance in mitigating these risks within nonprofit organizations.

> "Within the bounds of the law to say, you know what type of biases are permitted and what type of biases are unpermitted, even if there are ethics issues right? So, the board could say, you know, these are our values and ethics. And therefore, you know, we're going to safeguard it in making our decisions about how to tell our staff about what to do in terms of feeding the AI and what product comes out. So, it could be that first layer of defense, if you will, in the Swiss cheese model." (G. Takagi, April 24, 2024, private communication)

Figure 3: Visual Data Analysis



Summary

The findings from the data analysis are presented in three phases to make a cohesive argument for the research question, how can nonprofit corporate mechanisms govern AI? The third phase presents the findings with a visual framework. The Swiss Cheese Model is utilized to visualize how the various layers of oversight—board of directors, legal, AI ethics board, and employee inputs—interact to prevent risks and align AI applications with the nonprofit's mission and govern its development and deployment. Moreover, the findings bring to light the importance of layer independence and the necessity for organizational mechanisms to work cohesively to safeguard against potential harms.

Section 7: Implications and Recommendations

Implications

The implication of the integration of AI applications into nonprofit organizations poses a unique set of challenges and opportunities, showing the need for a comprehensive long-term strategy for how existing nonprofit corporate mechanisms can effectively govern AI. It is critical to recognize the potential limitations within existing nonprofit corporate mechanisms and their capacities to adequately govern the complexities of AI technologies. As this study has shown, the development and deployment of AI applications come with inherent risks, demanding a governance framework to mitigate potential harms to beneficiaries of the communities they serve.

More specifically, addressing the governance of AI within a nonprofit organization requires strategic adjustments and proactive leadership. For nonprofit boards, active involvement is essential to navigate the complex landscape of AI effectively. Leaders within these organizations must cultivate a long-term risk mitigation mindset to anticipate and manage the potential hazards associated with AI technologies. Given the evolving nature of AI and its impacts, further research is necessary to continuously refine and adapt governance mechanisms to meet future challenges and opportunities AI offers in the nonprofit sector.

Recommendations

The following recommendations are aimed at ensuring that the development and deployment of AI applications in a nonprofit corporation serve to advance organizational mission while maintaining ethical and legal integrity. First, leveraging the use of a governance framework, like

the application of the Swiss Cheese Model in the study, can provide a structured approach to navigate these challenges. Second, AI policies can establish protocols to mitigate risks associated with AI, ensure alignment with the nonprofit's values, and prevent ethical and legal issues. Third, an AI ethics board is seen as essential for protecting the organization’s integrity and public trust. Lastly, the recommendation that training data be mission aligned brings to light the impact training data has on AI behavior.

1. Develop an AI Governance Framework

The findings bring to light the complex challenges and risks associated with integrating AI into a nonprofit organization. The use of a governance framework is a useful tool to address these challenges. The Swiss Cheese Model used in this study illustrated how corporate mechanisms, such as the integration of an AI ethics board, provide a structure to navigate AI complexities and capitalize on the opportunities for an organization to further its mission.

2. Develop Comprehensive AI Policies

The research findings support the development of comprehensive AI policies. These policies would help in establishing clear protocols for risk assessment, ensuring AI alignment with organizational mission and values, while preventing potential ethical and legal violations. Setting specific boundaries and expectations are crucial in addressing the risks of AI misalignment, as identified through the study's application of the Swiss Cheese Model, where each layer of defense prevents potential hazards from penetrating and causing organizational failures.

3. Implement an AI Ethics Board

The recommendation to implement an AI Ethics Board within nonprofit organizations is supported by the findings of the literature and expert interviews. An AI ethics board could play a critical role in ensuring that AI deployments are not only legally compliant but also uphold organizational values, preserving the organization's integrity, reputation, and public trust.

4. Align AI Training Data to the Mission

As the research makes clear, AI training data has a significant impact on the development of AI applications to be aligned or misaligned with the organization's mission. This study showed that without careful alignment, AI applications can inadvertently undermine the nonprofit's goals, leading to unintended consequences that may contradict the organization's purpose. Mission alignment in AI training helps ensure that AI outputs are congruent with the nonprofit's values and objectives, thereby enhancing the effectiveness and trustworthiness of AI applications in supporting the organization's goals.

Section 8: Conclusions

In conclusion, the complexities of the development and deployment of AI applications requires nonprofit organizations to adapt and evolve their governance structures proactively, thinking long-term, and with a risk mitigation mindset. By embracing a comprehensive approach to AI governance, nonprofits can use the benefits of AI to advance their mission while managing the risks

associated with the technology. This requires the establishment of governance mechanisms, such as AI ethics boards and detailed policies, to oversee AI activities and ensure they do not deviate from intended ethical and operational parameters.

Furthermore, this study emphasizes the critical role of training data in shaping AI behavior and backs for the alignment of this data with the nonprofit's mission. This ensures that AI applications support rather than undermine organizational goals, enhancing the trust and integrity of nonprofits in the eyes of their stakeholders.

Section 10. Limitations

A combination of the subject matter's complexity in data analysis and developing theory, the rapid development and deployment of AI technologies, and the publication of relevant studies near the conclusion of this capstone's research resulted in a study and paper with limitations. As a result, study's depth, breadth, and theory refinement would have benefited by more time and resources to do a broader review of the literature and an expansion of expert interviews.

References

Arcas, B. A. y. (2023, October 10,). *Artificial general intelligence is already here.* NOEMA. https://www.noemamag.com/artificial-general-intelligence-is-already-here

Cummings, M. (2023). Identifying AI hazards and responsibility gaps [forthcoming]. In M. E. Noorman, & M. Verdicchio (Eds.), *Computer ethics across disciplines: Applying Deborah Johnson's philosophy to algorithmic accountability and AI.* Springer Nature.

Davenport, T., & Alavi, M. (2023, July 06). *How to train generative AI using your company's data.* Harvard Business Review. https://hbr.org/2023/07/how-to-train-generative-ai-using-your-companys-data

Google.org. (2024) *Nonprofits & generative AI.* www.google.org. https://services.google.com/fh/files/blogs/nonprofits_and_generative_ai.pdf

Griffin, T. A., Green, B. P., & Welie, J. V. M. (2024). The ethical wisdom of AI developers | AI and ethics. *AI Ethics,* https://doi.org/10.1007/s43681-024-00458-x

Hadley, E., Blatecky, A., & Comfort, M. (2024). Investigating Algorithm Review Boards for Organizational Responsible Artificial Intelligence Governance. *arXiv preprint arXiv:2402.01691*.

Hou, B. L., & Green, B. P. (2023). A multi-level framework for the AI alignment problem. Paper presented at *The Thirty-Sixth Annual Conference on Neural Information*

Processing Systems, New Orleans Convention Center. https://search.ebscohost.com/login.aspx?direct=true&AuthType=shib&db=edsarx&AN=edsarx.2301.03740&authtype=sso&custid=s3818721&site=eds-live&scope=site&custid=s3818721

IBM Research. (2023, Nov. 16). *What is AI governance? [video].* YouTube. https://www.youtube.com/watch?v=f6dx3Yh-Tww

IBM. (2024, Mar. 14). *What is an AI model?* ibm.com. https://www.ibm.com/topics/ai-model

Larouzee, J., & Le Coze, J. (2020). Good and bad reasons: The swiss cheese model and its critics. *Safety Science, 126*, 104660. 10.1016/j.ssci.2020.104660

Niederman, F., & Baker, E. W. (2023). Ethics and AI issues: Old container with new wine? *Information Systems Frontiers, 25*(1), 9-28. 10.1007/s10796-022-10305-1

Papagiannidis, E., Enholm, I. M., Dremel, C., Mikalef, P., & Krogstie, J. (2023). Toward AI governance: Identifying best practices and potential barriers and outcomes. *Information Systems Frontiers, 25*(1), 123-141. 10.1007/s10796-022-10251-y

Renz, D. O. (2016). *The jossey-bass handbook of nonprofit leadership and management* (4th ed.). Wiley.

Reason, J. (1990). *Human error*. Cambridge University Press.

Reason, J. (2000). Human error: Models and management. *British Medical Journal,* (320), 768-770.

Schuett, J., Reuel, A., & Carlier, A. (2024). How to design an AI ethics board | AI and ethics. *AI Ethics,* https://doi.org/10.1007/s43681-023-00409-y

Weidinger, L., & et.al. (2024). *Holistic safety and responsibility evaluations of advanced AI models* . [Manuscript submitted for publication]. https://doi.org/10.48550/arXiv.2404.14068

Winer, C. (2024, Mar. 18). *AI for nonprofits: Time to ask new questions.* linkedin.com. https://www.linkedin.com/pulse/ai-nonprofits-time-ask-new-questions-camille-winer-u2ege

Author's Bio

Linda A. Logemann has a Bachelor of Science degree in computer science from Rochester Institute of Technology. She spent many years coding and designing large software systems. In May of 2024, she will complete her Masters in Nonprofit Administration at University of San Francisco.