UNIVERSIDADE DO VALE DO RIO DOS SINOS – UNISINOS
UNIDADE ACADÊMICA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO
NÍVEL DOUTORADO EM ADMINISTRAÇÃO

VIRGINIA WESTPHALEN MOREIRA

IMPROVING DATA COLLABORATIONS FOR THE COMMON GOOD:
DEVELOPING A GOVERNANCE FRAMEWORK

PORTO ALEGRE

2024

VIRGINIA WESTPHALEN MOREIRA

IMPROVING DATA COLLABORATIONS FOR THE COMMON GOOD:
DEVELOPING A GOVERNANCE FRAMEWORK

Tese apresentada como requisito parcial para obtenção do título de Doutora em Administração, pelo Programa de Pós-Graduação em Administração da Universidade do Vale do Rio dos Sinos (UNISINOS).

Orientador: Prof. Dr. Jorge Verschoore

PORTO ALEGRE
2024

FICHA CATALOGRÁFICA

M838i Moreira, Virginia Westphalen

Improving data collaborations for the common good: developing a governance framework / por Virginia Westphalen Moreira. – 2024.
82 f. : il.; 30 cm.

Tese (doutorado) — Universidade do Vale do Rio dos Sinos, Programa de Pós-Graduação em Administração, 2024.
Orientação: Prof. Dr. Jorge Verschoore.

1. Micro governance functions. 2. Wicked problems.
3. Data collaboratives. 4. Gestão de dados. I. Título.

CDU 658:004.62

Catalogação na Fonte:
Bibliotecário Maicon Juliano Schmidt - CRB 10/2791

VIRGINIA WESTPHALEN MOREIRA

IMPROVING DATA COLLABORATIONS FOR THE COMMON GOOD:
DEVELOPING A GOVERNANCE FRAMEWORK

Tese apresentada como requisito parcial para obtenção do título de Doutora em Administração, pelo Programa de Pós-Graduação em Administração da Universidade do Vale do Rio dos Sinos (UNISINOS).

Aprovado em 30 de setembro de 2024

BANCA EXAMINADORA

Professor Dr. Jorge Renato Verschoore Filho

Presidente da Banca Examinadora - Unisinos

Profa. Dra. Amarolinda Iara da Costa Zanela Klein

Componente da Banca Examinadora - Unisinos

Prof. Dr. Bruno Anicet Bittencourt

Componente da Banca Examinadora - Unisinos

Profa. Dra. Rosiléia Milagres

Componente Externo da Banca Examinadora - Fundação Dom Cabral

Prof. Dr. Eduardo Henrique Diniz

Componente Externo da Banca Examinadora - Fundação Getulio Vargas

AGRADECIMENTO À CAPES

Esse estudo foi financiado em parte pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código Financeiro 001.

AGRADECIMENTOS

Uma pandemia, um negócio encerrado, um casamento, uma gravidez, um adeus muito difícil ao meu padrasto, o nascimento da minha filha, um puerpério complicado, aprender a ser mãe, o desafio das noites acordadas com o bebê, o dilema entre ser mãe e trabalhar, uma grande enchente. Viver todas essas situações, das mais fantásticas às mais difíceis, e continuar o doutorado foi um grande desafio. Em vários momentos, achei que não conseguiria chegar ao fim. De forma alguma eu teria conseguido sozinha. Foi lembrando de todas as pessoas que me deram força que encontrei o "empurrão" necessário para finalizar esse ciclo.

Um profundo agradecimento ao Prof. Dr. Jorge Verschoore, meu orientador desde o mestrado. Muito obrigada pelo tempo e paciência, pelos questionamentos sempre cirúrgicos e por todos os empurrões. Sem sombra de dúvidas, tua orientação foi chave no meu desenvolvimento. Aos fantásticos professores e a toda a equipe do PPGAdm, meu agradecimento pela generosidade em sala de aula e fora dela. Aos colegas e amizades que fizemos pelo caminho, que nossos caminhos continuem se cruzando e que as parcerias se fortaleçam.

Meus amigos, muito obrigada por entenderem as ausências e por compartilharem comigo desde as angústias até as vitórias. Uma amiga ganha destaque nesses agradecimentos, pois vem me apoiando na vida acadêmica desde o princípio: Juliana Pedroso Rohde, do fundo do coração, obrigada! Pelas infinitas conversas e encorajamento sempre que tive dúvidas.

Não poderia deixar de agradecer à minha família, pelo apoio constante e também pelos exemplos. Foi em casa que aprendi a baixar a cabeça e trabalhar. Ao meu pai, um agradecimento especial, certamente um dos meus maiores exemplos de dedicação ao trabalho e de vontade de aprender cada vez mais. À minha mãe, é difícil colocar em palavras o quanto teu apoio e incentivo foram importantes. Mesmo nos momentos mais difíceis das nossas vidas, tu estavas ali, como sempre, me ajudando a manter a calma, me ensinando a cuidar da Catita e me estimulando diariamente a me dedicar ao doutorado, lembrando-me o quanto isso era importante. Muito, muito obrigada.

Ao meu marido, Fabricio. Tu me apoiaste todos os dias, suportaste meus momentos de estresse, me deste força e seguraste todas as pontas. Com a chegada da Catarina, fizemos juntos uma série de esforços para que eu conseguisse conciliar a maternidade com a carreira acadêmica. E tu não mediste esforços para que eu conseguisse finalizar a tese. Não tenho dúvidas de que eu não teria chegado aqui sem teu apoio, meu amor. Meu mais profundo agradecimento.

Para Catarina.

Que nem imagina a força que me dá.

RESUMO

É no contexto do reconhecimento crescente do valor do compartilhamento de dados para diversos propósitos, incluindo pesquisa, desenvolvimento de políticas públicas e inovação, que as iniciativas de compartilhamento de dados para o bem comum têm surgido. Mesmo com o número crescente de casos e evidências que mostram os benefícios e o valor das colaborações em torno de dados, até hoje, a maioria dessas iniciativas tem sido experiências pontuais e limitadas, embora substanciais em impacto. Apesar de estudos destacarem os desafios enfrentados, existe uma lacuna na pesquisa voltada para soluções dos obstáculos enfrentados por colaborações de dados. Esta pesquisa defende que é necessário expandir o conhecimento sobre governança em colaborações de dados, apoiando a tese de que apenas uma visão ampla de governança não é suficiente para a sustentabilidade de longo prazo dessas iniciativas. Propõe-se, então, a seguinte questão de pesquisa: Qual é o papel da governança na sustentação de colaborações de dados para o bem comum? Essa tese propõe que os desafios enfrentados pelas iniciativas de colaboração de dados podem ser abordados por meio de uma perspectiva baseada nos problemas a jusante. Além disso, as funções propostas pela microgovernança podem apoiar o desenvolvimento de um ambiente colaborativo. Foi escolhida uma abordagem qualitativa de pesquisa com um desenho exploratório, utilizando especificamente o método de estudo de caso. Optou-se por um estudo de caso único devido ao acesso e à complexidade do fenômeno, especialmente no Brasil. O caso selecionado foi o "Minha Saúde Digital" (MSD), ativo no Rio Grande do Sul, Brasil, desde 2020. As evidências coletadas durante a pesquisa enfatizam a importância das funções de microgovernança na promoção de um ambiente colaborativo dentro das colaborações de dados. Quando desempenhadas de forma eficaz, essas funções contribuem substancialmente para a coesão e o sucesso do esforço colaborativo. No entanto, o debate gira em torno da identificação dos papéis e responsabilidades apropriados dentro da colaboração. Assim, são apresentadas três novas proposições relacionadas aos papéis do convener e do orchestrator no contexto da microgovernança.

Palavras-chave: Colaborativos de Dados. Governança Colaborativa. Problemas Complexos. Funções da Microgovernança.

ABSTRACT:

It is in the context of increasing recognition of the value of data sharing for various purposes, including research, policy development, and innovation, that the data-sharing initiatives for common good have emerged. Even with the growing number of cases and increasing evidence showing the benefits and value of data collaboration, to this day, most of the initiatives have been one-off, limited experiences, yet substantial in impact. Although studies are emphasizing the straits, there is a gap in research geared toward the solution to the challenges faced by data collaborations. This research uphold that it is necessary to expand knowledge about governance in data collaborations, supporting the thesis that just a broad view of governance is not enough for the long-term sustainability of the initiative. The following research question is proposed: What is the role of governance in sustaining data collaborations for the common good? This thesis proposes that the challenges faced by data collaboration initiatives can be addressed by viewing them through the lens of downstream problems. Additionally, the functions proposed by micro-governance will support the development of a collaborative environment. A qualitative research approach with an exploratory design was chosen, specifically employing the case study method. A single case study was selected due to access and the complexity of the phenomenon, particularly in Brazil. The chosen case was 'Minha Saúde Digital' (MSD), active in Rio Grande do Sul, Brazil, since 2020. The evidence collected during the research emphasizes the importance of micro-governance functions in promoting a collaborative environment within data collaborations. When effectively performed, these functions substantially contribute to the cohesion and success of the collaborative effort. However, the debate revolves around identifying the appropriate roles and responsibilities within the collaboration. Therefore, three new propositions are put forward regarding the roles of the convener and the orchestrator in the context of micro-governance.

Keywords: . Collaborative Governance. Wicked Problems. Data Collaboratives.

FIGURES AND CHART LIST

TABLE OF CONTENTS

1 INTRODUCTION

As technology and data collection methods advance, the volume of data collected daily is growing exponentially. Individuals and organizations are generating daily vast amounts of data in diverse formats (Pappas et al., 2018). These data sets typically contain an extensive amount of information, originating from a wide range of sources. Some of the primary sources include: satellites, mobile networks, social media content on Facebook, Instagram, LinkedIn, and others, IoT devices like smart thermostats, wearable fitness trackers, GPS devices, online retail, and more (Blazquez, 2018). This data forms what we call big data. Big data encompasses large and complex data sets that are difficult to manage, process, and analyze using traditional data processing tools and methods (Beyer and Laney, 2012; Bello-Orgaz, 2016).

This abundance of data has made it more accessible to organizations, enabling them to leverage it for insights and decision-making (Pappas et al., 2018). Organizations have realized the immense value that data-driven insights can bring to their operations (Mikalef et al., 2018; McAfee et al., 2012). Data-driven companies tend to outperform their competitors, as they can make more informed decisions, identify patterns, and predict trends. This perception of data as a competitive and rivalry asset has led to a cultural shift where data is viewed as a strategic tool (Mikalef et al., 2018). In fact, the importance of data in the economy as a whole has become increasingly evident. Organizations that effectively collect, analyze, and utilize data are better positioned to succeed and adapt to the rapidly evolving business landscape (McAfee et al., 2012).

This data-based decision-making movement is happening in more than just the private sector. Given the potential benefits of sharing and analyzing data, governments worldwide have been promoting open data initiatives, making non-sensitive datasets publicly available (Rasche et al., 2021; Hogan et al., 2017). The aim is to promote transparency, accountability, and public participation while fostering innovation and economic growth (Zuiderwijk et al., 2014; Borzacchiello and Craglia, 2012). By making public data more accessible, governments can empower individuals and organizations to leverage data for societal, economic, and innovative purposes (Hogan et al., 2017; Susha, 2020).

By increasingly promoting interdisciplinary collaboration, governments also allow for a broader understanding and the development of more comprehensive solutions to wicked problems (Susha et al., 2023, Klievink et al., 2018). The concept of wicked problems, introduced by Rittel and Webber (1973), describes complex, multifaceted, and persistent challenges that are difficult to define, let alone solve. These problems are characterized by their

intricate and interconnected nature, and often lack a clear and well-defined problem statement, making them resistant to straightforward or traditional solutions (Rittel and Webber, 1973). Examples of wicked problems include: Climate change, poverty, global health crises, education reform, urbanization, and housing (McCrea, 2020). Since these problems have become increasingly complex for individual organizations to solve, addressing wicked problems typically requires a collaborative approach. It involves engaging diverse stakeholders, from different backgrounds, and interdisciplinary collaboration to tackle the intricate nature of these challenges (Head, 2022; Susha et al., 2023).

Big data can play a significant role in understanding, addressing, and tackling wicked problems. Both topics are related in several ways, as they often intersect in addressing complex, multifaceted challenges (Susha, 2020; Verhulst, 2021). Big data analytics can provide the tools and insights needed to navigate the complexity and uncertainty associated with these challenges, supporting more informed decision-making and adaptive strategies. Data-driven analysis can reveal patterns, correlations, and trends that may not be apparent through traditional methods (Pappas et al., 2018; Rasche et al., 2021). As well as be used for continuous monitoring and tracking of changes related to the problem (Bartalucci, 2023). It is important to highlight that much of the most relevant data, the technology, the intelligence to understand this data, and the ability to generate insights are privately owned (Susha et al., 2020; Susha and Gil-Garcia, 2019). To unleash the full potential, this resource needs to be shared with actors who can implement high-impact programs.

Therefore, it is in this context, of open data initiatives, the open government movement, and the increasing recognition of the value of data sharing for various purposes, including research, policy development, and innovation, that the data-sharing initiatives for common good have emerged (Susha et al., 2023; Klievink et al., 2018; Bartalucci, 2023). These initiatives have a common goal of harnessing the potential of data to address wicked problems, however, they can go by various names depending on their focus, goals, and the organizations or sectors involved. Data collaboratives are partnerships between different organizations, often across sectors, that collaborate to share data and expertise for the common good (Verhulst and Sangokoya, 2015). Data Philanthropy involves organizations and individuals donating data to support social causes and address various challenges (George et al., 2020). A Data Commons is a shared data repository or platform where data is pooled and made accessible to researchers, policymakers, and the public for collaborative problem-solving (Grossman et al., 2016). In this research, considering that this topic is still incipient, both in practice and in theory, we chose not to focus on a single definition. Taking the aforementioned definitions as a guide, this

research uses, interchangeably, the terms data collaboration or data collaboratives, to designate formalized partnerships between actors from different sectors that collaborate to share data and expertise to overcome public challenges.

Even with the growing number of cases (The Gov Lab, 2023), which grew from 23 in 2016 to 200 in 2021[1], and increasing evidence showing the benefits and value of data collaboration, to this day, most of the initiatives have been one-off, limited experiences, yet substantial in impact (Susha and Gil-Garcia, 2019; Susha et al., 2023). There are inherent challenges that must be carefully addressed to change this scenario. The obstacles range from legal barriers, data privacy and security concerns, interoperability, and technical challenges, to trust and changing stakeholder interests (Klievink et al., 2018; Ruijer, 2021; Susha et al., 2023). Although studies are emphasizing the straits (Klievink et al., 2018; Susha, 2020), there is a gap in research geared toward the solution to those challenges.

Recent academic research on data collaboration has seen significant advancements across various domains. However, few studies have focused on governance, governance models, or governance's role in sustaining data collaborations (Klievnik et al., 2018; Susha et al., 2018; Ruijer, 2021). The studies that focus on governance, look at governance models more broadly. For the most part, they seek to adapt existing collaborative governance models, disregarding gaps related to the integration of collaborative governance and data science concepts as well as other specific challenges (Susha and Gil-Garcia, 2019; Klievink et al., 2018). In this regard, we argue that new governance structures, processes, and practices are needed to ensure the proper functioning of the collaboration in a long-term capacity.

Klievink et al. (2018) and Ruijer (2021) sustain that given that data collaborations can be seen as a public governance tool, with a focus on the active involvement of non-governmental actors, these initiatives can then be considered in the domain of collaborative governance. Collaborative Governance can be defined as an approach to decision-making and problem-solving that involves multiple stakeholders working together in a collective formal decision-making process (Ansell and Gash, 2008). It is characterized by a collaborative and constructive process in which government agencies, non-governmental organizations, and other relevant stakeholders collaborate to develop and implement policies, structures, programs, and solutions to achieve a public purpose (Emerson; Nabatchi and Balogh, 2012).

[1] Repository of cases around the world compiled by The Gov Lab at New York University at: https://datacollaboratives.org/explorer

Hence, in this research, data collaborative governance will be interpreted as a governance approach that comprises norms and rules that will coordinate, monitor, and regulate behaviors, influencing the decision-making processes and actions of the various parties involved, therefore allowing the conservation of partnerships and institutions, whether they are public or private (Bryson; Crosby and Stone, 2015; Ansell and Gash, 2008; Emerson; Nabatchi and Balogh, 2012). It is also important to highlight that governance is not only creating the conditions for ordered rules and collective action, but also enforcing the rules (Wegner and Verschoore, 2022).

Thus, we argue further that there is still a lack of clarity on the activities that must be carried out to overcome the challenges encountered daily. In the attempt to address this gap, some studies (Provan and Kenis, 2008; Ansell and Gash, 2008) suggest insights; however, there is still space for refinement and further understanding. Sørensen and Torfing (2021) believe that actors involved in collaborative governance should continuously monitor and assess the impact of their policies and be willing to adapt and modify their strategies when downstream problems are identified. In their work, Sørensen and Torfing (2021) use the term downstream problems to refer to the challenges and issues that can arise during the implementation of collaborative and networked governance approaches. They generally refer to resistance to change, such as opposition to new policies, non-cooperation, or conflicts among stakeholders, or involve difficulties in establishing an effective feedback loop, or yet inadequate or flawed evaluation processes that can lead to a lack of awareness about problems until they become significant.

Seeking to further fill the gap in the daily operationalization of collaboration, developing from existing literature, Wegner and Verschoore (2022) delve into the concept of network governance and its practical application in fostering collaborative environments. The authors propose a framework with the functions that must be performed by network leaders in order to stimulate cooperation. They present six functions that leaders must perform: Aligning, Mobilizing, Organizing, Integrating, Arbitrating, and Monitoring. Therefore, we maintain that a day-to-day vision of the functions that need to be carried out daily is needed, so that data collaborations can face the challenges of collaborating for data while generating a competitive advantage for all parties involved.

Since 2020, at the onset of the COVID-19 pandemic, the 'Minha Saúde Digital' project has been actively engaged in Rio Grande do Sul, Brazil. Originally conceived as an applied academic research endeavor, led by the private university Unisinos, this initiative has facilitated collaboration among a diverse array of public and private entities, including hospitals, health

plan operators, and non-profit organizations. Its initial objective was to create an intelligent communication framework using blockchain technology to integrate standardized clinical data, fostering seamless connectivity among healthcare providers and patients. Over time, 'Minha Saúde Digital' has evolved into a sustained collaborative effort, demonstrating remarkable longevity despite the challenges associated with data-sharing initiatives. This longevity, combined with its aim to address public challenges through data collaboration, makes it a compelling case for investigation in this thesis.

While the 'Minha Saúde Digital' project showcases the potential of data collaborations, it also highlights the governance challenges inherent in such initiatives. Sustaining these partnerships over time requires more than broad governance principles; it demands specific, actionable strategies to address daily operational challenges. These include aligning diverse stakeholder interests, ensuring data interoperability, maintaining trust, and adapting to changing circumstances. Despite progress in understanding collaborative governance, existing models often fail to integrate the unique demands of data collaborations or offer detailed guidance on day-to-day operationalization. In light of the foregoing, this thesis uphold that it is necessary to expand knowledge about governance in data collaborations, supporting the thesis that just a broad view of governance is not enough for the long-term sustainability of the initiative. Further on this thesis, we defend some propositions on the themes described above. Furthermore, given the backdrop, the following research question (RQ) is provided: **What role does governance play in sustaining data collaborations for the common good?**

1.1 Objectives of the research

This thesis aims to identify and detail which governance functions play a critical role in the long-term sustainability of data collaborations for the common good.

Aiming to achieve the general objective, these are the specific objectives:

i. Identify which governance functions play a critical role
ii. Identify the governance outcomes
iii. Propose a specific framework for the governance of Data Collaborations

1.2 Relevance of the research

Due to the growing interest and emergence of data collaborations for the common good around the world, in the most diverse sectors (The Gov Lab, 2023), the field began to gain attention in academic literature, with the first publications in 2017/2018 and the most recent as of 2023. A fairly new and expanding body of empirical papers (Klievink et al., 2018; Ruijer, 2021; Verhulst, 2021; Susha et al., 2023) and a few review studies (Susha et al., 2019; Ruijer, 2021) have been analyzing the main characteristics and challenges of these partnerships. To date, these publications have been mainly dedicated to understanding the conceptual and illustrative side of data collaborations. Even though there are advances in studies on the challenges faced by this type of partnership, an understanding of the importance of governance is still lacking. Few studies have focused on governance, governance models, or governance role in sustaining data collaborations (Klievnik et al., 2018; Susha et al., 2018; Ruijer, 2021).

In the study of Klievink et al., (2018), the scholars expanded upon Ansell and Gash (2018) framework to delineate the operational factors underpinning the efficacy of data collaborations. Their findings underscore the pivotal role of contextual elements in shaping the objectives of collaborative endeavors involving data. The authors also assert that technology antecedents and prior experience with open data are important for data collaborative governance. Nonetheless, they acknowledge the limitation of their study, rooted in its singular case study nature, emphasizing the need for future investigations employing more rigorous methodologies to corroborate and extend these initial insights.

Susha and Gil-Garcia (2019), used the Data for Climate Action case study as an illustrative lens to demonstrate the applicability of the Collaborative Governance Framework proposed by Emerson, Nabatchi and Balogh (2012) in elucidating the intricate governance dynamics within data collaborations. Their study underscores that while the framework offers insights into these dynamics, it presents limitations in comprehensively addressing the intricate complexities inherent in collaboration for data. The particularities that the authors found are mainly related to data sharing and the multi-sector nature of the partnership. It is important to point out that the study of Susha and Gil-Garcia (2019) was also carried out considering only one case and it is mainly conceptual.

Ruijer (2021) builds upon the work of Bryson, Crosby and Stone (2015). The author chose this model, as Bryson, Crosby and Stone (2015) put governance and technology at the heart of their framework while also integrating elements from various collaborative governance frameworks authored by others. Drawing insights from a living lab situated within a local

government context, Ruijer (2021) concludes that the distinctive dynamics inherent in data collaboratives necessitate tailored specifications within a collaborative governance framework. The challenges mentioned by the author are related to the definition of the public problem and the management of data sets. The results of this research also showed trust challenges underlying data collaboratives. It's noteworthy that while the article centers on data collaboratives, the case study involved solely public sector entities, lacking private organizational participation. Notwithstanding, the articles add to the literature, albeit still leaving gaps regarding the best governance model for data collaboration. As well as gaps related to how governance is expressed daily, considering activities that need to be performed to suppress the challenges found and foster long-term sustainability.

Therefore, the value of this research lies both in theory and in practice, as it involves expanding the theoretical understanding of the approaches and overcoming gaps in knowledge on governance adapted to the context of data collaborations. At the same time, it also seeks to look at the day-to-day practices and functions that need to be carried out for the long-term sustainability of the collaboration. The research developed in this thesis focuses on the evolution of literature and practice in the field of business administration and governance, reconciling the intersecting fields of knowledge, at the same time it draws on notions from many neighboring fields, especially technology.

Furthermore, data collaboratives can provide valuable insights for policymakers. Thus studying data collaborative governance is essential for public policy as it helps policymakers navigate the complex challenges associated with data-driven decision-making, while ensuring that data is used in a responsible, ethical, and effective manner to address societal challenges and improve public welfare.

1.3 Thesis Structure

The remainder of this thesis is organized as follows: The next section discusses the theoretical foundations and empirical backdrop of data collaborations for the common good and collaborative governance. This section is followed by a presentation of the methodology used to carry out this research, to collect and analyze the research. Next, the case select for the study is presented. This section begins with a general overview dedicated to explaining and detailing the context of the analysed case. Next, in fifth section, the results are presented. Here, the collected evidence is elucidated and analyzed through both theoretical and empirical lenses,

aiming to further explore the theories and expand the filed´s knowledged. Finalythe concluding remarks constitute the last chapter of this study.

2 THEORETICAL FOUNDATIONS

This section comprises the theoretical backdrop of this thesis. First, Data Collaboration for the Common Good is defined. Following data collaboration challenges are adressed. Then, using the Collaborative Governance approach (Ansell and Gash, 2008; Emerson; Nabatchi and Balogh, 2012; Bryson; Crosby and Stone, 2015), the work of Sørensen and Torfing (2021) on downstream problems and the work of Wegner and Verschoore (2022) on micro-governance, the governance challenges faced by data collaborations are explored. Due to the socio-technological nature of data collaborations (Susha et al., 2018), it was also important to bring references from the area of technology and data science to this study. Finally, a theoretical model that served as the basis for conducting this research is presented.

2.1 Data Collaboration for Common Good

Data Collaboration for Common Good initiatives exist to leverage data, expertise, and resources across sectors to address complex societal challenges in a collaborative, ethical, and impactful manner, ultimately aiming to create positive social change and improve the well-being of communities and society at large (Verhulst and Sangokoya, 2015; Susha et al., 2018; Klievink et al., 2018; George et al., 2020). In the realm of this type of initiative, it is possible to encounter various terms and concepts that are closely related, varying on their focus, and the organizations or sectors involved. These terms might be used interchangeably or with slight variations across literature and practice, however, they all revolve around the same core concept. The most common alternative names or related terms, based on publications[2], include Open Data Collaboration, Data Commons, Data for Good, Data Philanthropy, and Data Collaboratives.

Open Data initiatives refer to programs or efforts by governments, organizations, or communities to make data freely available to everyone, without restrictions on access or usage. The goal is to promote transparency, collaboration, innovation, and efficiency by providing accessible data that can be used, shared, and analyzed by individuals, businesses, researchers, and governments (Hall et al., 2012). These initiatives often involve releasing data sets collected by policymakers or organizations, allowing for analysis, interpretation, insight generation, and

[2] Number of articles with keyword in tittle - search on Scopus and WoS on december 30, 2023 - using quotation marks: Data Commons (227), Data for Good (59), Data Collaboratives (21), Data Philanthropy (16), Open Data Collaboration (8).

development of applications or solutions that leverage this information (Shadbolt et al., 2022). Open data can span various fields, including government spending, public transportation schedules, weather patterns, healthcare statistics, and more. The idea is to empower people to make informed decisions, drive innovation, and foster economic and social development by utilizing this wealth of available information (Janssen et al., 2012; Susha and Gil-Garcia, 2019).

Data Commons involves the creation of shared and accessible pools of data resources for a community, researchers, or organizations. It is a concept that promotes the open sharing of data, enabling diverse users to access, contribute to, and utilize datasets for various purposes (Dulong de Rosnay and Stalder, 2020). It encourages collaboration among individuals, organizations, or research communities to contribute data, insights, and analyses, fostering a collaborative environment for shared learning and innovation. Data Commons can be established by governments, research institutions, non-profit organizations, or collaborations among multiple stakeholders (Dulong de Rosnay and Stalder, 2020; Bollier, 2012).

In essence, Open Data focuses on making datasets available to the public, and Data Commons provides a collaborative platform for shared data resources, encouraging accessibility and interdisciplinary collaboration. While they share common goals of transparency and accessibility in the realm of data, they differ in their specific focuses and approaches toward achieving those objectives. The concept of Data for Good emphasizes the application of data-driven insights, tools, and methodologies. These initiatives utilize data analysis, predictive modeling, machine learning, and other analytical tools to derive insights and identify patterns (Mayer-Schönberger and Cukier, 2013).

Data philanthropy refers to the practice of private companies, organizations, or individuals donating their private datasets and analysis resources to contribute to the public good, social causes, research, or humanitarian efforts (Stempeck, 2014; Susha et al., 2019) It involves sharing proprietary data for purposes that benefit society, without necessarily seeking financial gain. Data philanthropy is related to Corporate Philanthropy as they are donating a company resource in a way that produces a significant impact and preserves or increases shareholder value (George et al., 2020; Stempeck, 2014). Data philanthropy can take different forms: (1) Direct Sharing, when companies or organizations share their datasets with nonprofits, research institutions, or policymakers working on projects aligned with their philanthropic goals; (2) Collaborative Projects between private entities, government agencies, and NGOs, where data is shared, pooled, or analyzed collectively to address specific social issues or challenges; (3) Data Access Programs where organizations create platforms or programs that

grant access to their data for specific purposes, such as academic research or humanitarian efforts (Stempeck, 2014; Susha et al., 2019; George et al., 2020).

Data Collaboratives, a relatively new term coined mainly by Susha et al. (2018), Klievink et. al (2018) and Verhulst and Sangokoya (2015), refer to a new form of intersectoral collaboration, in addition to the more traditional public-private partnership model, in which participants from different sectors – including private companies, research institutions, and policymakers – can share their datasets to address complex social problems. Thus, these agencies and organizations can "take advantage of the availability of diverse and complementary public and private data to better understand a specific problem and propose a solution" (Susha and Gil-Garcia, 2019, p. 2893). The definition of data collaboratives "emphasizes the process of collaboration between the parties and, therefore, suggests a more comprehensive vision that goes beyond just sharing data" (Susha et al., 2018, p. 2691).

In this research, considering that this topic is still incipient, both in practice and in theory, we chose not to focus on a single definition. Taking the aforementioned definitions as a guide, this research uses, interchangeably, the terms data collaboration or data collaboratives, to designate formalized partnerships between actors from different sectors that collaborate to share data and expertise to overcome wicked problens.

The field of data sharing for the common good encompasses a wide range of concepts, as seen above, and each of these terms represents distinct perspectives and applications. By not limiting the study to a single definition, it is possible to capture a broader spectrum of initiatives and strategies, thereby providing a more comprehensive understanding of how different organizations approach data sharing for societal benefits. Focusing on multiple definitions allows the study to be more inclusive of different models and frameworks. This inclusivity can highlight the strengths and weaknesses of various approaches, offering valuable insights into what works best in different contexts. It also encourages the exploration of hybrid models that combine elements from various definitions, potentially leading to more effective and sustainable data-sharing practices.

2.1.1 Actors Involved

Data collaborations have been conceptualized as cross-sector collaborations, thus, regardless of how they are named, involve a range of actors from different sectors (Susha et al., 2019). These actors usually are Government Agencies, Non-profit Organizations and NGOs, Private Sector Companies, and Academic and Research Institutions (Klievink et al., 2018;

Susha et al., 2019). It is possible to find different combinations among these actors for each initiative, and all four possible combinations among sectors: public-private partnerships, public-nonprofit partnerships, for-profit-nonprofit partnerships, and tripartite partnerships involving actors from all of these sectors (Susha et al., 2019).

The primary incentive for **government agencies** is related to the idea of the "common good". Participating in a data collaboration might be the means to achieve an objective that they can not on their own (Klievink et al., 2018). Their motivation might came from the potential to improve public services, enhance decision-making, and foster innovation through shared knowledge and resources (Janssen et al., 2012). According to Klievink et al. (2018), the government agencies may act as a facilitator for the initiative, being the part responsible for providing a safe environment, and infrastructure or mediating conflicts between private parties. Else, in most cases, they can stimulate and facilitate collaborative action of the other actors (Zuiderwijk and Janssen, 2014).

Given the centrality of policy intervention as an outcome of data collaborations, as highlighted by Susha et al. (2018), it becomes clear that the role of these government actors extends beyond mere administration or support. Their involvement is deeply intertwined with shaping and influencing policy (Janssen et al., 2012), which aligns with the overarching goals of data collaboratives to achieve high-impact societal outcomes. Therefore, from this point forward, the term **policymakers** will be used to more accurately reflect their role in steering these collaborations toward data-driven policy making and societal change (Susha et al., 2018).

Non-profit organizations includes organizations with different legal natures and operational structures, including charities, philanthropic foundations, voluntary and community organizations, community groups, social enterprises, and cooperatives. They are motivated to participate in data collaboratives due to the potential for enhanced impact, resource efficiency, and the ability to address complex social issues through data-driven insights (Carman and Fredericks, 2010). By pooling data with other entities, non-profits can uncover trends and patterns that inform their strategies and improve service delivery (Poel et al., 2018). The primary role of the non-profit sector is to leverage its expertise in the social sector to generate impact through the use of data. Non-profits possess deep knowledge of their beneficiaries, this knowledge enables them to align data provision with actual demand needs effectively (Susha et al., 2023). However, these also face challenges and risks. Non-profits often face constraints related to limited technical expertise and resources necessary for effective data management and analysis (Grubb and Easterling, 2012).

Private companies have various motivations to engage in data collaborations. Some of the main motivations include: gain access to diverse data sets that can drive innovation, enhanced market intelligence, and the opportunity to demonstrate corporate social responsibility (Porter and Kramer, 2011; Susha, 2020). They can contribute to the initiatives with capabilities to process and generate intelligence from data. Usually private actors plays a primary role in data sharing and sharing of technological capacity for data use. Many of the data that can be used to address social challenges are owned and handled by private entities (Susha, 2020). Engaging in data collaborations also comes with barriers and risks. Sharing data, especially sensitive information, can raise concerns. Companies must navigate issues related to data privacy and security, ensuring compliance with regulations such as GDPR to protect consumer information. Also companies might be hesitant to share proprietary data that gives them a competitive edge, fearing loss of market advantage or intellectual property (Zuiderwijk and Janssen, 2014; Susha and Gil-Garcia, 2019).

Universities and researchers are motivated to participate in data collaboratives due to the potential for advancing research, fostering innovation, and enhancing their societal impact. Similar to the scenario of private companies, collaborating with external entities allows universities to access diverse data sets, which can lead to new insights and breakthroughs across various fields of study (Mayer-Schönberger and Cukier, 2013). The academic sector may play crucial roles in data collaboratives by functioning as impartial intermediaries, effectively connecting data providers with users. Their engagement guarantees the application of rigorous scientific methodologies in data analysis and interpretation, thereby bolstering the credibility and reliability of results (Cai and Zhu, 2015). However, the interest in generating new knowledge rather than the results of its application, can result in challenges (Perkmann and Schildt, 2015). Anyhow, these institutions provide platforms for the secure and ethical sharing of data, which aids in converting raw data into actionable insights, ultimately facilitating informed decision-making and policy formulation for the public good (Eaton et al., 2017).

In addition to the actors mentioned above, it is important to highlight the **role of conveners.** According to Susha et al. (2023), conveners are the initiators of data collaboratives. Also conveners will play "various facilitating roles in developing a more sustainable data ecosystem for the data collaboratives" (Susha et al., 2023, p.2). Conveners are entities or organizations that take the lead in bringing together various stakeholders to share and collaborate to achieve a specific goal related to innovation (Harrison; Pardo and Cook, 2012). They are essential for initiating dialogue, building trust, and ensuring that all parties have a shared understanding of goals and expectations (Bryson; Crosby and Stone, 2015). Typically

have the responsibility of creating a framework for collaboration, establishing guidelines, and ensuring that the interests and concerns of all participating parties are addressed. They can help to create a neutral platform where participants can openly share data, ideas, and resources, ultimately driving collective action towards common objectives (Vangen and Huxham, 2003). They might play a central role in organizing, facilitating, and often overseeing the collaborative effort (Gupta; Panagiotopoulos and Bowen, 2020; Susha et al., 2023).

However, as stated by Susha et. al (2023), what is understood about the role of conveners to date is based on the literature on other topics, such as innovation ecosystems. Regarding the role of the convener in data collaborations this role may have some particular challenges, and still needs a clearer definition, given the challenges inherent to this type of initiative (Susha et al., 2023).

> …for conveners it is not always clear which role they should assume at which point in time to stimulate the development of the data collaborative in the most effective way. Data collaboratives as a new phenomenon might require a particular approach to convening. (Susha et al., 2023, p.2).

In their research Susha et al. (2023) find that successful data collaboratives require strong leadership from conveners to navigate the complexities of multi-stakeholder environments. The results highlight that conveners contribute significantly to building trust and legitimacy among stakeholders, which are essential for effective collaboration. Additionally, the study reveals that the presence of well-defined governance structures and processes, facilitated by conveners, is crucial for managing the dynamics of collaboration. The findings suggest that data collaboratives that benefit from proactive and capable conveners are more likely to achieve their intended societal impact, demonstrating the importance of the convener role in the overall success of these initiatives. Regarding the limitations of the study, the authors emphasize that it is a generalist study focusing on early stages of development.

Going further it is possible to draw a parallel with the role of the **orchestrator**, frequently mentioned in the literature on collaborative networks (Provan and Kenis, 2008; Vangen and Huxham, 2003; Dhanaraj and Parkhe, 2006). Typically, conveners focus on initiating and setting up the collaboration (Harrison; Pardo and Cook, 2012), while orchestrators are more involved in the ongoing management and coordination of activities (Provan and Kenis, 2008; Vangen and Huxham, 2003). However, as stated by Strasser et al. (2022) few studies examine the complexities of orchestrating collaborative networks for social innovation. The authors emphasize that network leadership is a distributed practice involving various

individuals and organizations that support transformative capacity development. The orchestrator's role involves enabling and supporting the network's collective abilities to achieve transformative impact. Wegner, Hölsgens and Bitencourt (2023, p.9) study "highlights that the orchestration roles performed by leaders in social innovation initiatives may not belong to only one individual, but to the group that is an integral part of the collaborative network". The authors argue though, the approaches studied in relation to the role of the orchestrator in this context of social innovation initiatives say very little about the daily activities required for the collaboration to be successful.

2.2 Data Collaboration Challenges

Data collaboration can vary greatly in terms of organizational form, sharing, and use of data. Nonetheless, Verhulst and Sangokoya (2015) identify three main benefits of data collaboration, besides the solution to the wicked problem: (1) Data from a wide variety of sources (private and public) allows both public policy and decision-making processes in the private sector to be data-driven; (2) Add value by creating important platforms for exchanging and coordinating information between data providers and users; (3) Can increase synergy within the data community (data collectors, data integrators, data policy experts, and data scientists), facilitating the emergence of much-needed standards and frameworks to make data interoperable and useful between organizations and sectors. However, not all data inherently possesses value; rather, extracting meaningful insights requires specific skills encompassing data processing, visualization, data science, and advanced analytics (Hoffman et al., 2019; Susha et al., 2018). Thus, as stated by Susha et. al (2023) "organizations find it increasingly difficult to balance the benefits and risks that such data sharing poses".

2.2.1 Data Capabilities

The prerequisite for delivering accurate insights lies in the legitimate collection and reliable processing of data. An accurate insight is an unbias insight, which might be a challenge in itself to achieve. Observations by Hoffman et al. (2019) underscore that the delivery of unbiased insights is intricately linked to both the quality of data and the contextual background in which insights are generated. The solution begins with developing the questions that must be answered by the available data. Defining the problem and aligning the demand for data with its supply is recognized as a pivotal task and one of the biggest challenges within a collaborative

framework, involving an iterative process until an optimal scenario is identified (Susha and Gil-Garcia, 2019). In collaborative endeavors, companies can define the main goal of the collaboration in connection with the available data and cultivate the requisite capabilities to achieve this goal (Susha and Gil-Garcia, 2019). However, it is important to note that supply is generally the bottleneck; inaccessibility of data is cited as one of the challenges inhibiting the emergence of data sharing initiatives. The private sector benefits from greater data availability, but the motivations for companies to disclose their data are often considered inadequate (Zuiderwijk and Janssen, 2014; Susha et al., 2023).

Before the data is available for analysis, stemming from the diverse array of data providers involved, whether multiple entities within the same industry or several companies spanning different industries, privacy and interoperability concerns arise. Ensuring data privacy and security is paramount, as policymakers must protect sensitive information from breaches and misuse, which could lead to public distrust or adverse policy outcomes (Zuiderwijk and Janssen, 2014). There must be a prior identification of risks in data sharing, especially when it comes to personal data - data about individuals -, and the elaboration of a strategy to mitigate these risks (Susha and Gil-Garcia, 2019). Regarding Interoperability, the greater the number of data sets combined in a data collaborative, the greater the chance of adding value. On the other hand, it also increases the efforts needed to understand this information, stored in different formats and heterogeneous systems (Susha et al., 2018; Ruijer, 2021).

As Gupta and George (2016) write, companies need a combination of certain tangible (e.g. data and technology), intangible (managerial and technical skills, for example), and human resources capabilities, to build a data analysis capability. It is possible to define the data analysis capability as "the ability of a firm to provide insights using data management, infrastructure, and talent to transform business into a competitive force" (Mikalef et al., 2018, p. 557). Thus, is understood that the collaborative needs to develop big data analysis capabilities, which will be further defined here as the ability of an organization to capture, integrate, and analyze large volumes of data that can be complex and diverse, and use insights obtained from this data to make informed decisions that generate real business value (Cao and Duan, 2015; Kwon et al., 2014). It is important to note that the way companies use their big data analysis capabilities can vary greatly, directly influencing their strategic capabilities and added value through data (Mikalef et al., 2018). Reinforcing the need to understand the initial context of the actors involved and their impact on the results of collaborative data.

2.2.2 Social Impact Measurement

Even though measuring social impact is recognized as a critical element that can accelerate the diffusion of data collaborations and ensure long-term sustainability, there appears to be no shared method in the literature for defining and measuring the impact. It is still a generally understudied topic in this context. Evaluating social impact is crucial for demonstrating the value and effectiveness of data-driven initiatives (Susha et al., 2019). This evaluation helps build trust among stakeholders, attract ongoing funding, and foster a culture of accountability and transparency. Without concrete metrics to assess outcomes, Data Collaboratives risk losing stakeholder engagement and failing to scale their efforts effectively (Verhulst & Young, 2017).

Systematic impact assessment can guide strategic adjustments and optimize resource allocation, thereby enhancing the overall efficacy and sustainability of these collaborative efforts (Davenport, 2019). To effectively measure social performance, Ebrahim (2014) emphasizes the importance of clarity in defining the mission and objectives, ensuring that the metrics align with these goals. Organizations should employ a mix of quantitative and qualitative methods, combining numerical data with narrative evidence to capture the full picture of their social impact. Furthermore, continuous feedback loops and adaptive learning processes are crucial, allowing organizations to refine their strategies and interventions based on real-world outcomes and stakeholder feedback (Ebrahim, 2014). This comprehensive approach ensures that social performance measurement is not only robust and rigorous but also relevant to the specific context and objectives of the organization.

Building on this, Moore and Khagram (2006) suggest that measuring social impact should be viewed through the lens of public value creation, where the focus is not only on the immediate outputs but also on the broader societal outcomes and legitimacy of the initiative. They argue that organizations, particularly in the public and non-profit sectors, must adopt a holistic approach to performance measurement that incorporates multiple dimensions of value, including social, economic, and environmental impacts. This approach requires flexibility and adaptability, as the goals and circumstances of public value creation can evolve over time. Moore and Khagram (2006) also highlight the importance of legitimacy and stakeholder engagement in measuring social impact.

> …but it turns out to be both expensive and difficult to measure the outcomes of government action; partly because the desired results often occur some years after the government has acted, in a place that has become far removed from the government agency's current operations. Further, precisely because so much happens between the actions of the government agency on one hand, and the social results on the other, even if we can see a change, we cannot be sure it ought to be attributed to the government's action rather than some other cause. (Moore and Khagram, 2006, p.7).

This means that impact assessment should not only focus on quantitative metrics but also consider qualitative evidence that reflects the perceptions and experiences of those affected by the initiative. By integrating these insights, organizations can ensure that their measurement strategies are comprehensive, contextually relevant, and capable of capturing the full scope of the public value they aim to create.

Chart 01 - Challenges of Data Collaboration

CHALLENGE	BRIEF DESCRIPTION	MAIN AUTHORS
Data Capabilities	It is related to the capability of interoperability, storage and security, as well as analyzing and generating unbiased insights, which respond to a common goal.	Hoffman et al. (2019) Susha and Gil-Garcia (2019) Zuiderwijk and Janssen (2014) Susha et al. (2023) Gupta and George (2016) Cao and Duan (2015) Kwon et al. (2014) Mikalef et al. (2017)
Social Impact Measurement	It is related to the fact that there is no shared method in the literature for defining and measuring the impact. However evaluating social impact is crucial for demonstrating the value and effectiveness of data-driven initiatives	Susha et al. (2019) Verhulst & Young (2017) Davenport (2020) Ebrahim (2014)

Source: prepared by the author (2024)

Beyond the specific points discussed in this section, the documented studies in the field, highlight broader issues related to governance as a whole, extending beyond the realm of data collaborations. To provide a more comprehensive analysis, the next section will delve into these issues through the lens of Collaborative Governance, offering greater depth and clarity.

2.3 Collaborative Governance Challenges

Data collaboration initiatives can vary widely regarding organizational form, data sharing, or data usage. Often this type of collaboration does not have traditional hierarchical structures and the decision-making flow is not top-down either. Even if it does not have formal structures, coordination mechanisms and the roles of stakeholders and leadership are important components for the collaboration to thrive (Susha et. al., 2019). Overall, governance structures and mechanisms have the potential to significantly affect the success of the initiative (Ruijer, 2021). Taking Ansell and Gash (2008) definition into consideration, Klievink et al. (2018) proclaim that given that data collaboration initiatives can be seen as a public governance tool, with a focus on the active involvement of non-governmental actors, these initiatives can then be considered in the area of collaborative governance. Ruijer (2021), based on the concept proposed by Emerson, Nabatchi and Balogh (2012) that collaborative governance is the processes and structures of public policy management that involve different actors in a constructive way to achieve a public purpose, also understands that the definition of governance for data collaboration must involve collaborative governance.

Thus, in this research, data collaboratives governance will be interpreted as follows: a governance approach that comprises norms and rules that will coordinate, monitor, and regulate behaviors, influencing the decision-making processes and actions of the various parties involved, therefore allowing the conservation of partnerships and institutions, whether they are public or private (Bryson; Crosby and Stone, 2015; Ansell and Gash, 2008; Emerson; Nabatchi and Balogh, 2012) to ensure that efforts are directed toward achieving collective goals and generating relational gains (Dyer; Singh and Hesterly, 2018). It is also important to highlight that governance is not only creating the conditions for ordered rules and collective action, but also enforcing the rules (Wegner and Verschoore, 2022).

2.3.1 Conflicting Goals

Even though the main goal of the collaboration is the solution to a wicked problem, each of the different actors involved might have additional divergent objectives, priorities, or purposes for participating in a collaborative effort (Klievink et al., 2018). For instance, academic and research institutions often aim to contribute to scientific knowledge and innovation. Their goals may differ significantly from those of commercial entities, which are driven by market demands, shareholder value, and competitive advantage. In collaborative

settings, partners may pursue their own organizational objectives while also working towards common goals. This dual focus can create tensions and challenges in maintaining the collaboration, as partners must constantly negotiate and balance their individual and collective aims. (Hoffman et al., 2019; Susha and Gil-Garcia, 2019; Klievink et al., 2018).

On the other hand, the ability of the initiative to add value can increase with the number of participants, once the data and overall capabilities increase. Klievink et al. (2018), based on the studies of Vangen and Huxham (2003) referred to this challenge as the goal paradox, since "the fact that the involvement of diverse actors is a source of both value and conflict" (p.381). The goal paradox highlights the importance of continuous negotiation, mutual understanding, and the development of trust among partners to ensure that the collaboration remains effective and sustainable over time.

Without a shared vision or strategic alignment among collaborators, defining and maintaining common objectives and developing a cohesive plan becomes challenging. Conflicting interests may manifest in disagreements over how resources, including funding, personnel, and technology, should be allocated within the collaborative effort. This lack of cohesion can lead to confusion, inefficiencies, and suboptimal outcomes.Therefore, achieving stakeholder alignment and engaging in discovery of interests, concerns, and values at the outset of a partnership, is critical (Hoffman et al., 2019; Susha and Gil-Garcia, 2019; Klievink et al., 2018).

2.3.2 Asymmetries of Power

Bianchi, Nasi and Rivenbark (2021) state that collaborative governance is a leadership task, suggesting that some forms of hierarchy must exist within the network to create a vision, motivate, and lead to success. Nonetheless, as stated by Ansell and Gash (2008) power imbalances among stakeholders can significantly impact the effectiveness of collaborative processes. Power asymmetry can affect who participates in the collaborative process and how much influence they have, can hinder the development of trust among participants or disproportionately favor the more powerful participants, thus undermining the legitimacy and effectiveness of the collaborative effort (Ansell and Gash, 2008).

In the context of data collaboration, there is an inherent asymmetry of power, since in most cases it is a public-private collaboration of some kind. The policymakers might be seen as a key facilitator of collaboration or a source of tension taking the fact that it is also a part of the partnership (Ruijer, 2021; Hoffman et al., 2019; Klievink et al., 2018). However, much of the

most relevant data, the technology, the financial resources, the intelligence to understand that data and the ability to generate insights are privately owned (Susha et al., 2018; Susha and Gil-Garcia, 2019). Power asymmetry often manifests in terms of who controls the data. Entities that own or generate large volumes of data can set the terms of access and usage, potentially marginalizing smaller participants who may have less leverage to negotiate favorable terms. Furthermore, the institutional environment has an important impact in this context, since collaboration can be voluntary or mandatory, and be influenced by political dynamics or legal and regulatory issues (Ruijer, 2021; Verhulst and Young, 2017; Bryson; Crosby and Stone, 2015; Emerson; Nabatchi and Balogh, 2012).

Susha et al. (2019) argue that the coordination of data-driven partnerships might be costly or difficult, nonetheless, the lack of coordination of roles, resources, and activities is a challenge that needs to be dealt with given that it will affect the dynamics of collaboration. Verhulst and Young (2017) argue that effective governance structures are crucial to managing power asymmetries. That inclusive governance mechanisms should be established to ensure that all stakeholders have a voice in the decision-making processes. This includes creating transparent processes, equitable participation rules, and mechanisms for conflict resolution.

2.3.3 Trust

Bryson, Crosby and Stone (2015) argue that trust and collaboration are widely recognized as mutually reinforcing. They create a positive feedback loop, particularly when there is an existing relationship between the actors. According to Ansell and Gash (2008), when stakeholders have a history of low cooperation and high conflict, the collaborative network may struggle with diminished trust and commitment among its members. This lack of trust can hinder the collaborative process, leading to challenges in achieving consensus and shared goals. Conversely, a successful history of collaboration builds social capital and fosters trust, reinforcing a virtuous cycle of cooperation. Ansell and Gash (2008) also emphasize the importance of trust-building as an iterative process, where the gradual development of trust and social capital can help overcome initial conflicts and create a more resilient and effective collaborative network.

Edelenbos and Klijn (2007) explore the importance of trust in complex decision-making networks, particularly in the context of public-private partnerships. They argue that trust is crucial for effective cooperation and the successful management of such networks. Their research highlights that trust can facilitate smoother interactions, reduce transaction costs, and

enhance the overall efficiency of collaborative efforts. However, initiating this process can be challenging. Ansell and Gash (2008) recognise that building trust is a process that requires time and emphasise the role of collaborative leadership in doing so.

Regarding the relationship between trust and contracts, Edelenbos and Klijn (2007) point out that while contracts are necessary to formalize agreements and provide a legal framework, they cannot replace the role of trust. Instead, trust and contracts are not mutually exclusive but rather complementary mechanisms that can enhance relationship stability and effectiveness. Contracts play a crucial role in the initial stages by setting clear terms of engagement, while trust gradually builds through ongoing interactions. Trust reduces transaction costs by minimizing the need for extensive monitoring and enforcement associated with formal contracts (Edelenbos and Klijn, 2007; Woolthuis; Hillebrand and Nooteboom, 2005).

Woolthuis, Hillebrand and Nooteboom (2005) further add that both trust and contracts are essential in mitigating opportunistic behavior. Trust helps mitigate opportunism by fostering a sense of mutual obligation and ethical behavior. At the same time, contracts provide legal recourse and clarity that can deter opportunistic actions. Ansell et al. (2020) add that agents can participate in the network to monitor what is happening or to protect their specific interests. This balanced approach is especially crucial in data collaborations involving the sharing of business intelligence, where the potential for opportunistic behavior is heightened. Susha and Gil-Garcia (2019) reinforce "convincing companies to trust their commercial data to an outsider can be challenging" (p. 2896). Trust is about risk, and in the case of data collaboratives, risk is about sharing sensitive information. In such contexts, ensuring robust contractual frameworks while simultaneously cultivating trust can significantly mitigate risks and enhance collaborative outcomes. Klievink et al. (2018) highlight that the antecedents of collaboration are crucial for the success of data collaborations.

> Data are a key organizational asset, and opening up data to new users, potentially by others, means giving up some control over that asset and relinquishing some autonomy to the collaborative. Fear over what others might do with the data could be a disincentive to collaboration (Klievink et al., 2018, p.382).

Furthermore, the literature shows the influence of trust on the quality of information shared (Klievink et al., 2018). In the case of data collaboratives, the concept of trust is strictly linked to the core data activities, especially when sensitive information is involved, where trusted data structures and processes are necessary (Stalla-Bourdillon et al.,2021).

Trust is also reflected in the governance model. Provan and Kenis (2008) argue that the density of trust between the parties influences the formation of the governance structure. Specifically, shared governance is more likely when trust is distributed, while intermediary organisation arises when trust is less or more concentrated in the hands of a single member. Edelenbos and Klijn (2007) argue that trust reduces the need for hierarchical control, thereby enabling more flexible and cooperative relationships. They also highlight that the presence of trust can mitigate the challenges posed by power imbalances and hierarchical structures, leading to better governance outcomes.

Chart 02 - Governance Challenges of Data Collaboration

CHALLENGE	BRIEF DESCRIPTION	MAIN AUTHORS
Conflicting Goals	It is related to the fact that each actor involved might have additional divergent objectives, priorities, or purposes for participating in a collaborative effort. This dual focus can create tensions and challenges in maintaining collaboration, as partners must constantly negotiate and balance their individual and collective aims	Klievink et al. (2018) Hoffman et al. (2019) Susha and Gil-Garcia (2019) Vangen and Huxham (2011)
Asymmetries of Power	It relates to the fact that there is an inherent asymmetry of power in data collaborations, since in most cases it is a public-private collaboration of some kind. Power imbalances among stakeholders can significantly impact the effectiveness of collaborative processes.	Ansell and Gash (2008) Ruijer (2021) Hoffman et al. (2019) Klievink et al. (2018) Susha et al. (2018) Susha and Gil-Garcia (2019) Verhulst and Young (2017)
Trust	It relates to the fact that trust is about risk, and in the case of data collaboratives, risk is about sharing sensitive information. However, Trust reduces complexity, reduces transaction costs and facilitates effective cooperation and problem-solving among diverse stakeholders.	Bryson et al. (2006) Ansell et al. (2008) Klievink et al. (2019) Susha and Gil-Garcia (2019) Provan and Kenis. (2008) Woolthuis et al. (2005) Edelenbos and Klijn, 2007

Source: prepared by the author (2024)

2.4 Downstream Problems of Collaborative Governance

Several studies have already been conducted to develop specific collaborative governance frameworks (Emerson; Nabatchi and Balogh, 2012; Ansell and Gash, 2008; Bryson; Crosby and Stone, 2015). The frameworks presented by Bryson, Crosby and Stone, 2015, Ansell and Gash (2008), and Emerson, Nabatchi and Balogh (2012), along with studies

by Provan and Kenis (2008) and other authors in the field, primarily address the backgrounds of the actors involved in the network, the dynamics of collaboration regarding macrostructures or modes of governance, and the expected outcomes. In other words, these studies focus on so-called upstream problems (Sørensen and Torfing, 2021). However, as stated by Sørensen and Torfing (2021), there is still a gap on how collaborative governance will happen in the day-to-day life of a collaborative network, solving downstream challenges. The authors remark that *"the main problems with collaborative governance are possibly to be found after a joint decision is reached and thus relate to the implementation, evaluation, and accountability*" (p.2).

Sørensen and Torfing (2021) suggest that despite the inherent chaos of the collaborative governance process, it can be analytically divided into distinct stages. These stages include problem identification and agenda setting, where stakeholders recognize issues and set objectives; initiation and design, involving the formation of the governance structure and engagement of stakeholders; deliberation and decision-making, where stakeholders negotiate and formulate action plans; implementation, focusing on executing the strategies and coordinating resources; and evaluation and learning, which assesses outcomes, facilitates learning, and informs future initiatives. This structured framework aids in understanding and managing the complexities of collaborative governance.

Previous studies on data collaborations and governance have predominantly focused on a conventional understanding of collaborative governance, addressing only upstream problems. In the study of Klievink et al. (2018), the scholars expanded upon Ansell and Gash (2018) framework to delineate the operational factors underpinning the efficacy of data collaborations. Their findings underscore the pivotal role of contextual elements in shaping the objectives of collaborative endeavors involving data. The authors also assert that technology antecedents and prior experience with open data are important for data collaborative governance. Susha and Gil-Garcia (2019), demonstrate the applicability of the Collaborative Governance Framework proposed by Emerson, Nabatchi and Balogh (2012) in elucidating the intricate governance dynamics within data collaborations. The particularities that the authors found are mainly related to data sharing and the multi-sector nature of the partnership. Ruijer (2021) builds upon the work of Bryson, Crosby and Stone (2015). The challenges mentioned by the author are related to the definition of the public problem and the management of data sets.

In this thesis, we argue that the main challenges encountered by data collaborations extend beyond the initial stages of data governance. These challenges arise in the initiation and design stage, where overcoming the challenge of trust (Susha and Gil-Garcia, 2019; Klievink et al., 2018; Hoffman et al., 2019) is necessary for stakeholders to engage in discussions,

negotiations, and deliberations to develop a shared understanding and consensus on the issues at hand. This involves joint problem-solving, idea generation, and the formulation of strategies and action plans. Challenges also emerge during the implementation phase, where it is crucial to apply tangible, intangible, and human resource capabilities (Gupta and George, 2016; Mikalef et al., 2017) to coordinate and mobilize resources, assign responsibilities, and execute collaborative initiatives. Additionally, challenges are present in the evaluation and accountability, or social impact measurement (Susha et al., 2019; Verhulst & Young, 2017) phases, where measuring the results of the initiative becomes essential. Therefore, it is understood that all the challenges faced by data collaboration initiatives can be overcome if viewed through the lens of downstream problems. By redirecting the analytical focus towards downstream problems, a greater understanding of how governance can efficiently address these challenges may be achieved.

Proposition 01: *Governance of downstream problems will establish an enabling environment for the long-term sustainability of Data Collaborations*

2.4.1 Micro-Governance of Collaborative Networks

In an attempt to fill the gap in the daily operationalization of collaboration, developing from existing literature, Wegner and Verschoore (2022) propose a framework with the functions and practices that must be performed by network leaders in order to stimulate cooperation between related parties. The authors argue that governance functions will create an environment that supports collaboration, thereby achieving intermediate governance outcomes. The governance of collaborative networks does not directly produce the final outcomes for participants, but it plays a critical role in establishing an environment conducive to effective collaboration. According to Bianchi, Nasi, and Rivenbark (2021), effective governance structures are essential for facilitating coordination, trust, and shared goals among network members. By providing clear frameworks and support mechanisms, governance ensures that participants can contribute their best efforts to the collective objectives, ultimately enhancing the network's overall performance and success.

Schilke and Lumineau (2018) contend that governance functions are pivotal in structuring routine interactions, helping to reduce uncertainty and foster effective collaboration. They emphasize that governance mechanisms play crucial roles in coordination, control, and adaptation, ensuring that all parties align their actions and adhere to agreed-upon terms while

remaining flexible enough to adapt to unforeseen changes. In this way, the micro-governance functions serve as a protective mechanism, safeguarding the alliance's collaborative objectives from potential disruptions.

The work of Wegner and Verschoore (2022) holds significance for the current research as the authors constructed their concepts by synthesizing and integrating pivotal elements from preceding network governance literature. The authors' research is based on studies by Zhong et al. (2017), Klijn et al. (2010), Persson et al. (2011), Smith (2020), Huxham and Beech (2008), among others. The micro-governance of collaborative networks framework aligns with the objectives of this study, offering a strategic response to the challenges elucidated in the previous section. Furthermore, its application is aptly suited to address research gaps in the identification of optimal governance models for data collaboration.

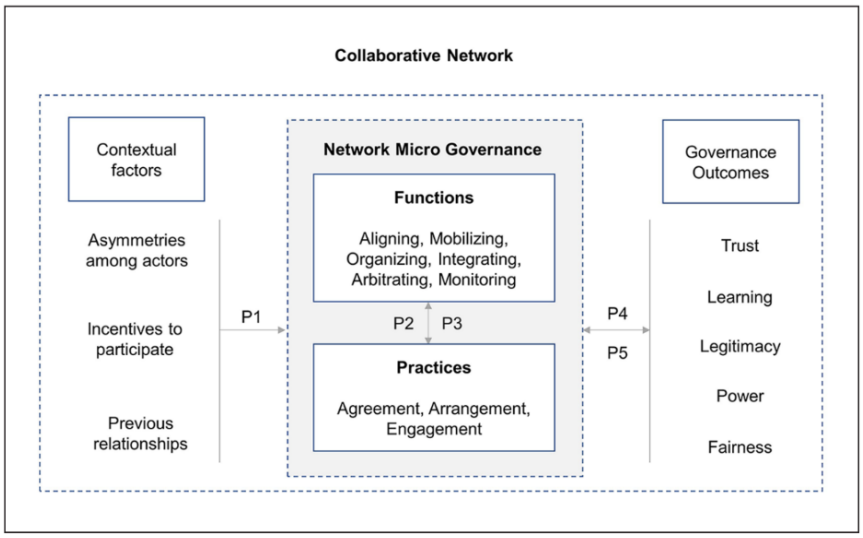
Firstly, ate the center of the model, the authors suggest six functions leaders should perform, which are: (1) the responsibility of leaders to maintain the **alignment** of goals and objectives of the participants, using this leveling between the parties as a basis for making decisions concerning the direction that the network will take; (2) **mobilize** represents the role of governance in stimulating an environment of exchange and joint action between members towards achieving the common objective; (3) **organizing** is the function related to the development of an environment conducive to the practical productivity of members, promoting the organization of human, financial, technological, and legal resources, as well as the organization of routines and processes that contribute to the network's goals; (4) **integrate** refers to the work necessary to unite the different parties involved in the network, identifying their specific resources and capabilities and ways to share them; (5) In a non-hierarchical and cooperative context, the function of **arbitration** becomes important to resolve conflicts inherent to the process; (6) **monitoring** is the function that looks closely at actions and results, promoting course corrections when necessary (Wegner and Verschoore, 2022).

Wegner and Verschoore (2022) further group the governance practices necessary to perform the aforementioned functions into three groups: (1) **agreement** groups functions related to selection and integration of partners, aligning objectives, and organizing contracts, roles, and responsibilities; (2) **arrangement** refers to functions that seek to facilitate the coordination of activities between network members; (3) **engagement** seeks connection between participants, through functions related to facilitating communication and integration activities (Wegner and Verschoore, 2022).

In their interpretation, these functions and practices of micro-governance will foster the success of the collaborative environment among network members. “Micro-governance does

not guarantee the achievement of the instrumental goals of a collaborative network, but it helps network members in developing intermediate outcomes" (Wegner and Verschoore, 2022, p. 9).

Figure 01 - Wegner and Verschoore Model



Source: Wegner and Verschoore (2022)

The outcomes proposed by the authors are (1) **trust,** a condition that promotes integration among members, stimulating communication and connection, supporting a positive environment; (2) **learning** is a mechanism that shapes the future functions and practices of governance, (3) **legitimacy**, is the network validation for its members or external stakeholders (4) **power**, is related to the capacity of a members to influence other's actions and behaviors; and (5) **fairness**, refers to the perception of network members regarding governance rules. These results stem from a carefully cultivated environment, particularly enhanced by the network's micro-governance, which fosters cooperation to achieve the collaborative's goals.

Wegner, Hölsgens and Bitencourt (2023) research, in addition to confirming the functions defined in the article by Wegner and Verschoore (2022), add three new functions to collaborative networks for social innovation. These are "**designing**", "**bridging**" and "**legitimizing**". The "designing" function refers to the activity of identifying potential members for the collaboration network. "Bridging" is about developing relevant connections with stakeholders outside the organization and at the same time relevant to the development of the initiative. "Legitimizing" is about through communication, seeking understanding from stakeholders who may be opposed to the social innovation.

> Identifying and inviting stakeholders requires a clear picture of the resources needed to achieve the expected outcome, and who may provide them. Network orchestrators who are able to map out those two threads will successfully design the collaborative network. (Wegner et al., 2023, p.8).

Rosa, Wegner and Del Ben (2023) advancing in the studies of Wegner and Verschoore (2022) propose a new contextual factor, time of collaboration has emerged as a new and relevant factor. The authors observed "that time provides opportunities to develop stronger ties, facilitates communication, and reduces the need for alignment" (Rosa, Wegner and Del Ben, 2023, p. 14). The authors also propose a new outcome of micro-governance, which is the connection between stakeholders, which goes beyond the financial gains of cooperation, thus being a relational gain (Dyer; Singh and Hesterly, 2018).

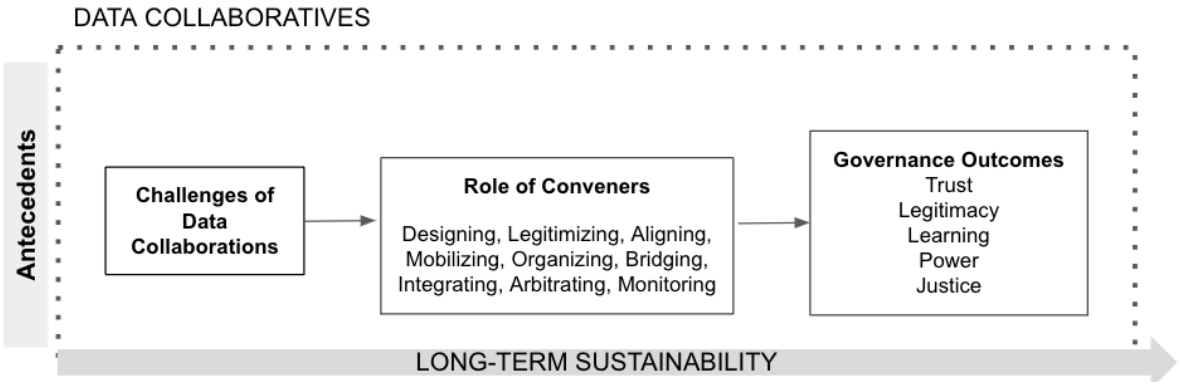
Schilke and Lumineau (2018) argue that governance functions are essential for routine interactions, which can reduce the chances of misalignment, misunderstanding, and disputes among members. Klijn et al. (2020) suggest that how a network is governed impacts the results it archives directly. Cabral and Krane (2018) emphasize the importance of guidance within networks to enhance collaborative governance. They postulate that effective guidance and facilitating in a network helps align the diverse goals of participating organizations, fostering a cohesive strategy and shared vision. Although networks tend to be a group of interdependent organizations with horizontal relations, a certain level of hierarchy remains (Klijn et al., 2020). Therefore, it is feasible to associate micro-governance functions with the long-term sustainability of a data collaboration initiative, as they provide a foundation for directing collaborative activities and addressing encountered challenges. It is also understood here that the functions of micro-governance require someone in a position of respect from others, with the role of guidance and facilitating, to execute the functions upon other members. Thus, evolving in the vision of Susha et. al (2023) on the role of conveners, it is argued here that the functions of micro-governance may embody the role of conveners.

Proposition 02: *The role of conveners should encompass establishing micro-governance functions to achieving a collaborative environment.*

Considering what was presented in the literature on the topics of data collaborations, the challenges faced, collaborative governance and micro-governance, figure 2 represents the governance process of a data collaboration. The proposed model is built based on the Wegner and Verschoore (2022) model presented above, with the additions proposed by Wegner,

Hölsgens and Bitencourt (2023). Antecents need to be considered as they will influence the entire context of collaboration formation and may increase or decrease the challenges encountered. The model illustrates the challenges faced and the role of the convener in performing micro-governance functions. These functions will lead to the results of the collaboration, identified by Wegner and Verschoore (2022) as five particular types. This process will help overcome challenges, making the convener's role crucial not only during initiation but also in ensuring these functions are carried out daily. This creates a virtuous circle.

Figure 02 - Model for Long-Term Sustainability



Source: based on Wegner and Verschoore (2022) and Wegner, Hölsgens and Bitencourt (2023)

Thus, considering the theoretical foundation and the propositions presented, the subsequent chapters present the method used, the research results, the discussion on the results, and the final considerations.

3 RESEARCH METHODOLOGY

This chapter aims to elucidate the methodology employed to address the research question and attain the objectives outlined in this study. It is structured as follows: i) research design, emphasizing the methodological approach; ii) an outline of the fundamental structure of the Case Study method, and iii) procedural steps undertaken in the research process. The presentation of the selected case is in chapter four.

3.1 Research Design

Considering the research problem and its respective objectives a **qualitative research with an exploratory approach** was chosen. In particular, a **case study was conducted**. Given the novelty of the subject matter, and the scarcity of cases available, it is understood that a case study is appropriate for this research. Following Stake (1995) definition, a case study is *"the study of the particularity and complexity of a single case, coming to understand its activity within important circumstances*' (p. 11). In other words, case study research involves an in-depth examination of a specific instance or phenomenon to gain insights into its unique characteristics and the broader context in which it occurs. The author emphasizes the importance of understanding the intricacies of individual cases and their interactions with the surrounding environment, what defines a case study is its singularity. One can also add here some of the benefits pointed out by Simons (2014) who states that case studies are valuable for exploring rare or extreme cases, offering insights that may not be attainable through other research methods, also they are well-suited for investigating complex and multifaceted issues, providing an opportunity for in-depth analysis of the interrelations between various factors. This is to say that the choice of method is appropriate for this research that seeks to investigate a type of initiative that is still very rare and complex.

According to Sampieri et al. (2013), the qualitative approach diverges from positivism, a paradigm commonly associated with quantitative research, which seeks absolute truths, dismissing any deviation as incorrect. Hence, this study adopts an interpretive paradigm, avoiding the imposition of preconceived notions by the researcher onto the researched context. Instead, the researcher immerses into the field to discern the most pertinent aspects among the participants, as emphasized by Saccol (2009).

Considering the novelty of the addressed topics within the academic sphere, the nature of the research is perceived to be exploratory. Exploratory research is apt for investigations

aimed at enhancing ideas or uncovering insights, allows researchers to gain a deeper understanding and clarify concepts that are not yet well defined. Creswell and Poth (2016) emphasizes the significance of exploratory research in situations where a comprehensive understanding of a phenomenon is not yet established. According to Creswell and Poth (2016), exploratory research is instrumental in generating hypotheses, gaining preliminary insights, and identifying key variables and relationships that warrant further investigation. In this context, case study methodology emerges as a robust option for conducting exploratory research. The author argues that case studies allow for an in-depth, contextual analysis of specific instances or events, providing a detailed and nuanced understanding of complex issues within real-life settings. This approach is particularly valuable in exploring new or ambiguous phenomena, as it enables researchers to capture the intricacies and dynamics of the subject matter, thereby laying the groundwork for more definitive research (Creswell and Poth, 2016).

Eisenhardt's (1989) approach to case study research is well-regarded for its systematic and rigorous methodology. The methodology emphasizes a balance between flexibility and rigor, allowing researchers to develop well-grounded theories while being open to new insights that emerge from the data.The author outlines a detailed process for conducting case study research that can be summarized in the following main steps:

Chart 03 - Case Study Steps

STEP		DESCRIPTION
1	Getting Started	Clearly specify the research questions. Use existing theories as a foundation.
2	Selecting Cases	Selecting cases purposefully.
3	Crafting Instruments and Protocols	Use multiple data collection methods.
4	Entering the Field	Collect data systematically. Maintain flexibility to allow for adjustments.
5	Analyzing Data	Compare and contrast findings.
6	Shaping Hypotheses	Constant comparison and contrast to ensure the theory is closely tied to the data.
7	Enfolding Literature	Compare the emergent theory with existing literature.
8	Reaching Closure	Conclude when theoretical saturation is reached

Source: adapted from Eisenhardt (1989)

In this research, step 01, presented in the introduction, was defined based on the theoretical reviews presented in chapter 02. Steps 02, 03 and 04 are presented in this chapter. The research instrument was constructed based on the final categories and previously tested. The remaining steps are presented in the following chapters.

3.2 Case Selection

The criteria for choosing cases are essential for the quality of the results and must be defined in advance and with extreme care. The researcher must decide which and how many cases are necessary to achieve the desired depth and breadth of the study (Eisenhardt, 1989). According to Creswell and Poth (2016), the selected case must be one that is both accessible and capable of providing rich, in-depth data. The case should be intrinsically interesting, unique, or particularly revealing, thereby allowing the researcher to explore the complexities and nuances of the phenomenon in question.

A single case study was chosen due to the access and complexity of the phenomenon, especially in Brazil. Creswell and Poth (2016) notes that while single-case studies do not provide the breadth of information that multiple-case studies offer, they compensate with depth. Single-case research can be particularly powerful and insightful when the case itself is unique, critical, or exemplary. Stake (1995) advocates for selecting a case that offers the potential for rich, in-depth understanding and insight into the research question or problem. Even though it is a single case study, data was analyzed at multiple levels and phases of the project, from the initial elaboration to an evolution with branching into new projects.

The case selected was, 'Minha Saúde Digital' (MSD), a project that has been active in Rio Grande do Sul, Brazil, since 2020. The project initiated as a research project led by Unisinos, with financial support from Capes. As the initiative develops, influenced by a series of factors analyzed in greater depth in the next chapter, a division occurs, with the initial initiative remaining focused on academic research and a second project emerging, with a more applied focus. The initial overarching objective was to create an intelligent information and communication framework utilizing blockchain architecture to integrate standardized clinical data, fostering seamless connectivity among healthcare providers and patients. The longitudinal nature of this case, spanning from 2020 to 2024, provides an opportunity to analyze its evolution over time, including how governance practices and stakeholder dynamics have adapted to sustain the initiative. The next section of this thesis details the case.

The selection criteria for the case necessitated its involvement in a collaborative project between public and private entities, aimed at addressing a common public issue through data sharing. The MSD case satisfies these criteria. Moreover, this case is substantiated as an empirical subject due to its achievement of surpassing initial objectives, yielding favorable outcomes, and sustaining activity to date, indicative of a degree of long-term viability. Additionally, access and convenience served as another criterion for case selection. Regional proximity and involvement of pertinent stakeholders facilitated attendance at initial meetings and enabled access to diverse internal and public project documents, including published articles.

3.3 Data Collection

A combination of three data collection procedures were employed for this research: 1) semi-structure interviews with MSD representatives and governance specialists; 2) archival analysis of documents from websites, reports, and previous meetings; 3) observation of meetings - MSD meetings took place via Teams from 2020-2023.

Using the "snowball" technique, interviews were conducted with professionals involved in the MSD project, representing various actors and roles within the initiative. Appendix A contains the question guide used during the interviews. A total of 14 interviews were conducted, each lasting between 30 and 90 minutes. These interviews were carried out online via the Google Meets platform, which allowed for a more flexible schedule and facilitated transcription as participants permitted recording. Fot he purpose of this thesis, all interviews transcriptions but one were translated from portugues to english, trying to be the most loyal to the original content. Only one interview, with an international specialist in data collaboratives governance was conducted in English. Chart 04 lists the interviewees. In order to maintain anonymity each interviewee received a code by which they will be identified for the remainder of the research. It is important to highlight that all interviewees involved in the case, have a high level of knowledge about the case, with a broad view of all the situations that occurred. Furthermore, the interviewees had experience with governance, thus contributing not only to the narration of the case, but also with an understanding of management practices.

Chart 04 - Interviewees

Code	Role	Type of Organization	Date
I01	Technical Manager	Private Hospital	Fev/2024
I02	Governance Specialist	Investor	March/2024
I03	Orchestrator - Pilot Project	Associated Entity	March/2024
I04	Entrepreneur	Tech Company	March/2024
I05	Orchestrator - Research	University	May/2024
I06	Technical Manager	Tech Company	May/2024
I07	Researcher	University	June/2024
I08	Researcher	University	June/2024
I09	Researcher	University	June/2024
I10	Governance Specialist	University	June/2024
I11	Governance Specialist	Private Hospital	June/2024
I12	NGO Founder/Manager	NGO	June/2024
I13	Entrepreneur	Private Company	June/2024
I14	Academic Professor/Researcher	University	June/2024

Source: Elaborated by the author (2024)

Regarding document analysis, we studied documents shared by interviewees during interviews, files shared during meetings, reports sent to Capes, news published on institutional and news websites, among others. From these materials, notes were taken to facilitate data analysis and identify patterns and behaviors. Chart 05 presents the main documents accessed.

Chart 05 - Documents

Type	Title	Link
Document - Submitted Project Capes	MinhaSaúdeDigital: Modelo Inteligente de Blockchain para Informações de Saúde e Interação com Pacientes no âmbito da COVID-19	https://drive.google.com/file/d/1Q3tG-YQ7YzP2L805DPU9UsOS-wgL4kjx/view
Document - Capes Execution Report	Minha Saúde Digital Execution Report - Ana 3	NA
Recordings of Minha Saúde Digital meetings	Recordings of 17 MSD meetings between dates 03/22 and 04/24	Via Unisinos Teams
Unisinos Website	Através do Projeto Minha Saúde Digital, estudantes da Escola Politécnica realizaram parte do Doutorado no exterior	https://www.unisinos.br/noticias/atraves-do-projeto-minha-saude-digital-estudantes-da-escola-politecnica-realizaram-parte-do-doutorado-no-exterior/
Doucment - Icolab Case presentation	Construindo Futuros Possíveis, Prováveis e Desejáveis na Saúde Público e Privada	NA
Icolab Website	Projeto MinhaSaúdeDigital é aprovado na Capes	https://icolab.org.br/noticia/projeto-minhasaudedigital-e-aprovado-na-capes/
Icolab Website	Educação em Saúde Digital com iCoLab, Feevale e Inova RS!	https://icolab.org.br/noticia/o-icolab-feevale-e-o-programa-inova-rs-juntas/
Icolab Website	Minha Saúde Digital transformando a saúde pública e privada com inovação...	https://icolab.org.br/noticia/minha-saude-digital-transformando-a-saude-publica-e-privada-com-inovacao-e-tech/
News	Iniciativa de Open Health do Hospital Ernesto Dornelles ganha vida com conexões via BioHub do Tecnopuc	https://tecnopuc.pucrs.br/iniciativa-de-open-health-do-hospital-ernesto-dornelles-ganha-vida-com-conexoes-via-biohub-do-tecnopuc/

Source: Elaborated by the author (2024)

3.4 Data Analysis

Content Analysis was chosen as the analysis procedure for this research. Content analysis aims to use a set of methodological instruments that can be applied to diverse discourses, to systematically and objectively analyze the content of textual data (Bardin, 2011). The primary objective of content analysis is to interpret the qualitative aspects of communication by categorizing and quantifying the presence of specific words, themes, or concepts within the text. Bardin (2015) delineates this process into several key phases: (i) the pre-analysis phase, where the material is selected and the analytical framework is established;

(ii) the coding phase, involving the systematic breakdown of the text into manageable categories based on predefined criteria; and (iii) interpretation phase, where the categorized data are analyzed to identify patterns, draw conclusions, and infer meanings that address the research questions. This structured approach ensures a comprehensive and replicable analysis, facilitating the extraction of meaningful insights from complex textual data.

The pre-analysis phase involved organizing the research material to make it operational and to systematize the initial ideas. During this stage, transcripts of all interviews were created, providing a comprehensive overview of the information collected. This process ensured that all data were systematically prepared for subsequent analysis.

During the exploration phase, the material was coded, and categories were defined a priori. These categories, known as final categories, were formed from initial and intermediate categories. The initial categories emerged from the direct coding of the transcript analysis and comprised words and expressions that referred to and synthesized processes, capacities, or skills. These initial categories were then grouped to form intermediate categories, which were subsequently combined to create the final categories. The final categories follow the Challenges of Data Collaboration summarized in Chart 01 and the Governance Challenges of Data Collaboration summarized in Chart 02. In the interpretation phase, the raw results were analyzed and interpreted to ensure they were meaningful and valid. Inferences were made with the support of additional sources of evidence and were related to the current literature on the topic, as presented in Chapter 2.

To enhance the reliability and validity of the findings, a triangulation procedure was performed. Triangulation involves using multiple sources or methods of data collection to cross-verify findings, reducing potential biases inherent in individual data sources and providing greater robustness to the results (Stake, 1995; Flick, 2013). In this study, triangulation was particularly crucial given the complexity of the MSD initiative and its governance dynamics. By combining interviews, archival data, and meeting observations, it was possible to confirm key findings, identify discrepancies, and integrate new evidence that complemented the original data. This methodological approach not only strengthened the validity of the results but also provided a more nuanced understanding of the MSD case.

Triangulation is particularly valuable in **e**x-post-facto explanatory research—such as this study—which seeks to retrospectively analyze cases to uncover causal relationships. In this instance, triangulation allowed for cross-sectional evaluation of the MSD initiative, comparing insights from different sources to construct a more accurate and holistic account of the governance processes and challenges involved. By corroborating evidence across interviews,

documents, and observations, this study mitigated collection biases, ensuring that conclusions were supported by consistent, multi-faceted data.

4 RESULTS

This section presents the results of the research. It begins with the characterization of the case study, followed by the evidence collected, considering the challenges listed in the bibliographic review and the Model for Long-Term Sustainability presented in Figure 02. By meticulously gathering and analyzing all the evidence during the research process, following the data analysis method outlined in Section 3, we aimed to identify the presence or absence of micro-governance functions in the case studied. The analysis focused on pinpointing specific micro-governance functions. By systematically evaluating the evidence, we sought to uncover patterns and discrepancies that could indicate the operational dynamics of micro-governance within the studied context. This rigorous process not only facilitated a detailed mapping of the governance landscape but also allowed for a nuanced interpretation of the factors influencing the effectiveness and efficiency of micro-governance in the case under consideration.

4.1 Case Presentation

The "Minha Saúde Digital" (MSD) project was conceived in the second semester of 2020, with its official approval occurring in August 2020, during the ongoing COVID-19 pandemic, with a completion date initially scheduled for the second half of 2024, being postponed to the first half of 2025. Originally envisioned as an applied academic research initiative, the project was spearheaded by private university Unisinos, receiving financial support from CAPES (Coordination for the Improvement of Higher Education Personnel), which an institution of the Ministry of Education (MEC).

> During the pandemic, some individuals were discussing research projects. There was a significant opportunity with Capes. Thus, we created the Minha Saúde Digital initiative. This project was coordinated with hospitals and other collaborators. (Interviewee 04).

The context in which the project emerged is highly relevant as it aims to address a pre-existing problem within the Brazilian healthcare system, which became significantly more apparent during the COVID-19 pandemic. Health systems consist of multiple agents and services that often struggle to share patient clinical data adequately and at the necessary speed for a more effective response to a pandemic. Health data have historically functioned in isolation, without integration among the various agents of the healthcare system. In Brazil, there

are still no initiatives that effectively benefit patients by making their information available among healthcare institutions. With each new service demanded by a patient, a new record is created without retrieving existing information. This situation not only complicates and slows down patient care but also increases public health and health insurance costs due to the repetition of unnecessary tests.

> We have a lot of waste in healthcare. We have requests for duplicate tests…the patient has just had a blood test in one hospital, ends up going to another for some reason and redoes the test…so there is a waste of time and resources. With this logic of centrality in the patient, that the data is theirs and they will carry it with them, the question of interoperability arises. (Interviewee 03)

In this context, the effective, integrated, and secure recording, management, and monitoring of clinical data at various levels of a healthcare system are crucial for better confronting the challenges posed by COVID-19 (Heymann & Shindo, 2020). It is understood that, especially when humanity faces the enormous challenge of combating COVID-19, it becomes essential to apply technologies that enable the reliable and secure exchange of data. This allows not only for better-informed decision-making by healthcare managers but also ensures that patients receive the most prompt and accurate treatment based on the completeness of their information.

Thus, the MSD project envisioned a transformative impact on society by revolutionizing the healthcare system. This initiative promised to reshape the way we approach and manage public health crises. An essential premise of the MSD project is rooted in the concept of individual data sovereignty, positing that individuals inherently possess the authority to determine the availability of their own data among various institutions.

> …this research movement comes due to a change in the healthcare scenario, which is the patient-centered care model. This centrality has stimulated a lot of research and development, including investment. (Interviewee 02).

The inniciative aimed to address this pressing need for improved management of health information and communication between healthcare providers and patients, particularly amidst the challenges posed by the COVID-19 pandemic. Therefore, the central research question of the MSD project was: "How can different healthcare service providers be connected to establish an intelligent model of information and distributed, standardized communication, facilitating interaction and remote assistance to citizens during pandemics like COVID-19?" (MSD CAPES Project, 2020).

The overarching objective of the MSD project was to develop an intelligent information and communication framework using blockchain architecture to integrate standardized clinical data, fostering seamless connectivity among healthcare providers and patients. Specific objectives included enhancing clinical efficiency, reducing duplicate exams and redundant care, and improving service times. Moreover, the project aimed to ensure effective continuity of healthcare, particularly for chronic patients with care distributed throughout the healthcare system. The application of blockchain to establish a transversal interoperability structure between public and private entities, hospitals, and operators represents a significant milestone in Brazilian health technology (MSD CAPES Project, 2020).

> Health professionals who work in the network's hospitals can benefit from the data as can the patient. an example is the patient who may arrive unconscious in an emergency, unaccompanied. and perhaps they have relevant information in other hospitals that can benefit at that moment, in this treatment. (Interviewee 03)

MSD project embarked on an ambitious journey with clear deliverable goals aimed at transforming healthcare. To pursue these objectives, the first group of stakeholders was formed. It included private hospitals: Ernesto Dornelles Hospital and Hospital Moinhos de Vento; public hospitals: Conceição Hospital Group, Hospital de Clínicas de Porto Alegre, and Santa Maria University Hospital; philanthropic hospital: Hospital Santa Casa de Misericórdia, alongside health plan operator Unimed Central de Serviços RS, and Icolab (Collaborative Blockchain Institute) a non-profit organization. In addition to these actors, the Unisinos team comprised nine scholarship recipients, distributed among 3 post-doctoral scholars, 4 doctoral aplicants, and 2 master's aplicants, and financial resources for additional funding for the execution of the project.

This group of actors was brought together by the initiative's proponents, responsible for making the project official with CAPES. It is important to be clear that even though the project has managerial implications, it was originally a project focused on academic research. This influences, among other things, the speed at which objectives could be delivered, not being limited to a rigid and short timeframe, but having from its beginning the expectation of being completed by the second half of 2024. Furthermore, it is important to highlight that starting as a research project, there was no commercial objective. Regarding contracts and rules, there was a collaboration agreement for participation in the research project, signed by all. However, there was no other document detailing the responsibilities, duties and rights of any of the parties.

> There was an attempt. The agreement did not have an adequate format. There were several attempts, several revised documents, some parts were in agreement. But a cooperation agreement was not signed. (Interviewee 02)

> There was an agreement on the research project with Capes, there was an agreement on the pilot project, but no collaboration agreement was reached. But I don't think the agreement would have made a difference. (Interviewee 04)

The consortium, comprising hospitals, Unimed, and Icolab, intended to have biweekly meetings to discuss project progress, research findings, and collaborative initiatives; however, not all actors were available for every meeting. These meetings took place mainly online. A significant challenge encountered during the project's initial eighteen months was the stringent limitations imposed by the COVID-19 pandemic, which hindered the project team's access to partner hospitals. Unisinos researchers, including faculty, undergraduates, master's students, doctoral aplicants, and postdoctoral fellows, held weekly meetings to review project advancements, accomplishments, and future directions. These meetings also took place online or in person, depending on the availability of participants and pandemic-related restrictions.

Regarding the roles of those involved, Unisinos professors and scholarship holders were the primary actors in the operationalization of the project, taking on leadership responsibilities and coordinating various aspects of the initiative. The professor who proposed the project was responsible for setting it up and initiating it, later taking on the role of orchestrator. The scholarship holders were tasked with contacting the hospitals and leading the interoperability of the data. The hospitals played a crucial role by providing essential data and serving as research fields, although their collaboration, while essential, also presented significant challenges. It is particularly important to note the participation of private hospitals, which primarily engaged in a monitoring capacity. Throughout the process, it became evident that the objectives of the initiative could potentially conflict with those of the institutions. Icolab entered the initiative with a belief in its commercial potential, but its role was not clearly defined from the outset, leading to conflicts and subsequent derived initiatives.

> We have good experiences of companies and universities working together, with products and innovation for the company. A project developed with many hands. Perhaps what was different here is that ideas were being created and people thought: this will make money, let's find a partner... but the ideas weren't solid. (Interviewee 1) ...some actors were more involved from the beginning. Private hospitals were more like listeners at meetings than participating in a practical way. (Interviewee 04).

> I understand that the market and academia operate at different speeds, partly because we have human resource constraints. If you develop a project already designed for the company, you assemble a project team, reinforce the staff, and work together. (Interviewee 1)

Some technological difficulties were reported during the interviews conducted as part of the MSD project. Each hospital involved in the project utilized a distinct data storage format, leading to interoperability issues where the data could not effectively communicate or integrate with one another. Furthermore, the absence of a strong data culture within each institution contributed to inconsistencies in data entry and maintenance. As a result, part of what was agreed upon at the beginning of the project could not be delivered by the institutions or was rendered unusable. In fact, some hospitals initially involved ended up withdrawing from the project as they realized they lacked the technological maturity to participate. These hospitals acknowledged the need to do their "homework" and evolve in digital culture before being able to participate in such a project. This significantly influenced the initial expectations of the deliverables. However, the researchers adapted their projects to work with the information that was available, showcasing their flexibility and problem-solving skills in the face of these technical challenges.

> Data is important information. Why would an actor share their information? It's hard to believe that projects will be successful without some kind of obligation. Even if they have a beautiful social objective. (Interviewee 10)

> In the context of data, the government creates basic legislation and defines standards...so the government is an important actor to help reinforce this interoperability and the need for standards. (Interviewee 8)

As the project progressed, new goals emerged among the participants, leading to disagreements and subsequent branching out. The main reasons for this division, identified during interviews, were differing opinions on the project's commercialization and the speed at which it should occur. This even led to concerns about opportunism among some actors. While the original group continued its academic efforts eventually achieving all the objectives, a new group was formed, comprising some of the initial stakeholders along with new participants. Consequently, two new initiatives were launched: the Minha Saúde Digital Pilot Project, which operated from 2021 to 2022, and Minha Saúde Education, which took place in 2023.

> ...for some players, it became a path driven by personal commercial interests, diverging significantly from the original research-focused objectives we initially sought. (Interviewee 02).

> We needed partners who could deliver the project within a certain timeframe. (Interviewee 03).

The project Minha Saúde Digital Education, even though it's a branch of the original Minha Saúde Digital project, cannot be classified as a data collaboration project. It is focused

on the professional training of healthcare professionals. ICoLab, in partnership with Feevale University and the Inova RS program of the Secretary of Innovation, Science, and Technology (SICT), delivered three free courses for researchers, enthusiasts, entrepreneurs, managers, and healthcare professionals in Rio Grande do Sul state. The courses covered relevant and current topics in healthcare, such as telemedicine, digital transformation, and health innovation centered on the person's care pathway. The connection with the MSD project was mainly because the courses delved deeply into the interoperability of sensitive health data, their medical record context, standardization, and data unification, aiming at improving care delivery and service by healthcare managers and professionals (iColab document, 2023).

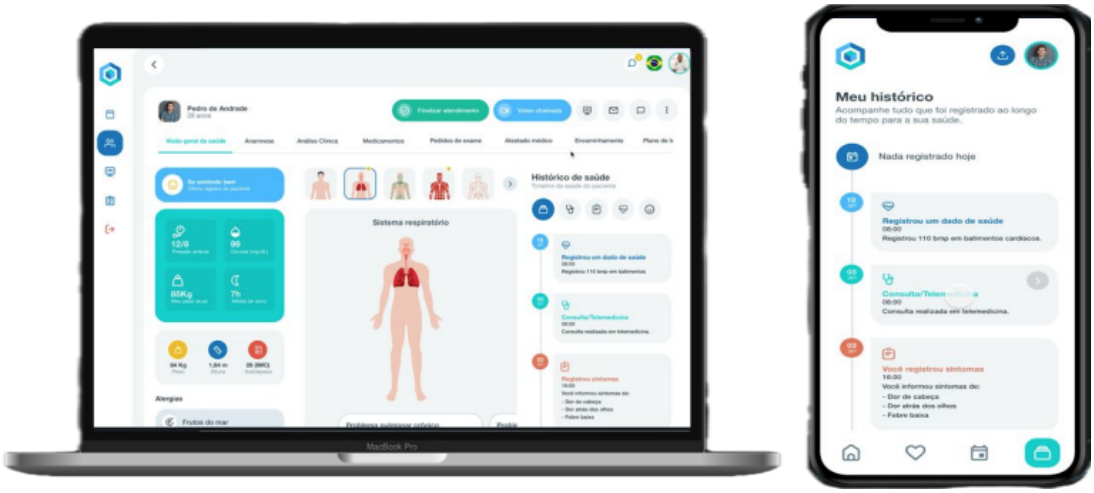
While maintaining the same purpose as the original initiative, the derivative Minha Saúde Digital Pilot Project, also sought a faster delivery speed, with a more commercial nature. ICoLab acted here as orchestrator, along with the entry of new actor IASIS Health as the technology provider, and some of the original actors, being then: Conceição Hospital Group, Ernesto Dornelles Hospital, Santa Casa de Misericórdia Hospital and health plan operator Unimed Central de Serviços RS. The MSD pilot project integrated information from the line of care Covid and Long Covid from the players involved, into an application for the patient and healthcare professional, in a simple, auditable, traceable, immutable way and with the consent of the citizen. The interoperability built connected all health players involved in the project, recovering and standardizing health data, such as: name of the healthcare establishment, date and time of entry/exit of the establishment, CNES and care modality (urgency and emergency, hospitalization, outpatient unit, etc.), origin code (initial ICD-10), clinical outcome (clinical discharge, hospitalization), other ICDs applied to the patient, procedures performed, test results (laboratory and imaging), immunization data, medication prescription and dispensing and discharge note.

> The project emerged from two main factors: external pressure to prioritize patient-centered care and the proliferation of suppliers who recognized this shift. Consequently, the platform logic was developed. (Interviewee 02).

The project was divided into four phases. The first phase was the longest, taking almost 12 months, and included the development of the cooperation agreement between all parties involved. This stage involved a series of meetings, negotiations on the role of each party involved, rights and duties, as well as analyzes by lawyers from each party. The second phase was dedicated to structuring data, profiles and information security; Follow by the third phase, integration and interoperability of clinical documents; and finally structuring the visual

evolution of the platform. Each phase had an average investment of R$ 50.000,00 considering the six signatory institutions of the Technical Cooperation Agreement signed between them. There were also two working groups. A quality group, involving doctors from the institutions. This group was responsible for analyzing the quality of shared data. The second group was the technical group, specialized in blockchain. All those involved in the first phase were invited to participate in the Pilot Project presentation meetings.

Figure 02 - MSD Pilot Project Plataform Interface



Source: Icolab Documents (2023)

Regarding the governance of the original project, it is evident that, as a research initiative, coordination falls under the purview of the educational institution. This includes signing the agreement with Capes and being accountable for delivering the promised results. The MSD Pilot Project, as confirmed through interviews, operated under a model of collaborative governance, with representatives from each institution jointly making decisions. The orchestration of the project was managed by the institution Icolab.

> ...we created governance groups. and we have a certain collaborative leadership. each participant has their representative. there is no single representative of MSD. (Interviewee 02)

> We had weekly meetings. We were building collaboratively. And we reached a point where we proposed a cooperation agreement. It was even necessary to be able to share the data. (Interviewee 05)

The MSD pilot project, despite being a shorter-term initiative with clear objectives from the beginning, also experienced moments of disagreement among the participants. The

interviewees' reports reveal that at certain points, not all actors involved were satisfied with every partner. However, these disagreements were mediated, and all the original actors remained involved until the end of the project.

Minha Saúde Digital Pilot Project garnered attention in 2022 when it was evaluated by a technical jury, and the winners were announced on the stage of the HIS – Healthcare Innovation Show, amidst over 220 entries from across Brazil. Subsequently, it earned recognition as a healthcare reference in the Administrative-Financial Management category, according to an annual study conducted by the Saúde Business portal (Icolab, 2022). Some of the resulting benefits identified in this pilot project were: reduced costs with duplicate exams; less bureaucracy, more efficiency, assertiveness, agility in care and the patient's journey; greater transparency, security and data privacy; healthcare ecosystem - reliable, traceable, more inclusive, fairer, smarter and sustainable. However, it is important to note that this project was treated as an MVP among the participants. And the results found were in relation to this MVP. It was not tested on a large scale.

In relation to the original project, registered with Capes, considering that the project is heading towards its final months, most of the goals have already been achieved. The MSD team mapped the primary limitations and challenges faced by partner institutions in managing and monitoring COVID-19 cases. This foundational work led to the creation of a secure, distributed blockchain model for data communication. Throughout the project, the MSD team modeled a knowledge base for EHR semantic interoperability using the OpenEHR standard. They defined a telemedicine model tailored to pandemic conditions, enabling remote interaction and healthcare assistance for patients. Leveraging Deep Learning, MSD proposed an intelligent model to analyze patient prognosis and generate decision-making indicators. They developed prototypes for clinical data interoperability using blockchain technology to seamlessly connect health institutions battling COVID-19. Additionally, MSD implemented a mobile application facilitating remote patient interaction and assistance. The project's impact was evaluated through numerous experiments with partner institutions, and the findings were disseminated through internationally impactful scientific publications, showcasing the project's groundbreaking contributions to healthcare.

The results delivered thus far also include the publication of 25 scientific articles in journals of international impact and 22 articles in national and internacional conferences. In addition, several actions were carried out to disseminate the knowledge acquired in the project, including a minicourse held at the XXII Brazilian Symposium on Computing Applied to Health (SBCAS2022), a lecture at the 42nd Congress of the Brazilian Computing Society (CSBC

2022), a lecture at the National Meeting of Production Engineering (ENEGEP 2022), and participation in a round table at the XXIII Brazilian Symposium on Computing Applied to Health (SBCAS 2023). Several lectures on the topic were given at universities, training courses, conferences, and other institutions, as well as actions with partner hospitals (Relatório de Execução Capes, 2023).

Moreover, alongside the initiatives discussed herein, certain stakeholders engaged in both the primary and pilot phases of the MSD project embarked on additional similar endeavors concurrently, but independently. Currently, there is an ongoing project named Minha Saúde Digital II, which shares similar interoperability objectives with its predecessor. This new initiative focuses specifically on diabetes care. Notably, two hospitals that participated in both the original Minha Saúde Digital project and the subsequent pilot project are also involved in Minha Saúde Digital II. These endeavors involved distinct institutions and shared akin objectives and experimental approaches.

> ..so in this next project, which has some similar objectives to what we had at MSD, we are going faster, because we learned a lot from MSD. And now we have the participation of two important public bodies together. This is very interesting. (Interviewee 02)

> In the evolution, in pilot two, the health department of Porto Alegre is involved. We talk about pilot two of the MSD because of the name it became known as, but we are now talking about a municipal health data network. (Interviewee 05)

The selected case for study is particularly appropriate due to the remarkable success and ongoing impact of the original Minha Saúde Digital (MSD) project. This project has not only demonstrated excellent results but is also still active, with additional deliverables planned, such as the publication of academic articles. The derivative MSD Pilot Project further underscores the suitability of this case by successfully delivering its proposed outcomes and garnering recognition for its achievements. Furthermore, the MSD initiative has spurred participation from various actors in similar initiatives, significantly enhancing the impact of data interoperability within the Brazilian healthcare system. Despite facing challenges, these efforts have consistently delivered substantial results and influenced other movements aimed at improving healthcare data collaboration. This comprehensive context makes the MSD case an exemplary model of data collaboration for the common good and a valuable subject for this research. The challenges encountered, although often seen as obstacles to governance, required careful navigation in turbulent waters to reach decisions and reduce disagreements. However,

they also represented an opportunity to break with biases and more traditional ways of viewing management and look at the downstream problems of governance, creating new solutions.

4.2 Evidence of Micro Governance Functions

As in Wegner and Verschoore's (2022) propositions, and the functions later added by Wegner, Hölsgens, and Bitencourt (2023), micro-governance presents functions that are used to guide the network towards the best path to achieve its goals. This section presents evidence of the governance functions observed in the case study, highlighting the varying extent of their presence and effectiveness. While some functions were easily identifiable due to their prominent role in the case studied, others were scarcely mentioned or could have been executed with greater precision, such as the designing and arbitrating functions. Notably, bridging function was entirely absent, and legitimizing was identified in the Minha Saúde Digital Pilot Project. This analysis will explore these findings in detail, shedding light on the governance dynamics at play in the case.

Throughout the research, the "designing" function emerged as a critical micro-governance role, consistent with the assertions of Wegner, Hölsgens and Bitencourt (2023). This function, which involves strategically identifying and integrating potential members into the collaboration network, was underscored by multiple interviewees. For instance, one interviewee highlighted the importance of engaging companies aware of their societal role, noting that "some companies already have this notion, this awareness, that they are part of society and need to contribute to improving population health, bringing them into this type of project is very interesting" (Interviewee 06). However, the case also revealed challenges associated with this function, as some actors, despite initial interest, exhibited limited involvement, suggesting gaps in the initial design process. One interviewee observed that certain participants "were more interested in staying in the loop than actually collaborating" (Interviewee 04), reflecting Provan and Kenis's (2008) observation that actors sometimes participate in networks primarily to maintain their status or reputation rather than to actively contribute to the network's goals. While another noted that some institutions "did not participate or left halfway, because they needed to do some homework" (Interviewee 05). Furthermore, while the evidence acknowledges the significant challenges that initiatives often face, it also highlights a clear connection between the motivations of the actors to participate in collaboratives and their willingness to work together to overcome these challenges. Thus, the mix of actors involved in data collaborations emerged throughout the research as a crucial and

complex point. This finding reinforces the proposition by Wegner, Hölsgens and Bitencourt (2023) regarding the designing micro-governance function.

These challenges imply that while the designing function might be present, its effectiveness could have been enhanced. A more rigorous application might have mitigated these difficulties by ensuring that all participants were both capable and committed to the collaborative effort from the outset. Thus, the MSD case illustrates the importance of the designing function not just as a foundational step, but as an ongoing process essential for maintaining collaboration momentum and reducing the likelihood of later friction.

The micro-governance function of "aligning" was identified as having been performed, though not without its challenges. According to Wegner and Verschoore (2022), this function, grounded in the work of Damgaard and Torfing (2010) and Acar et al. (2008), involves aligning the interests and objectives of all participants within a collaborative network. The aligning function is carried on during the daily routine as a process to in order to keep all the actors on the same page, following the same protocols, towards the same goals. The effectiveness of the collaboration is heavily dependent on this alignment. However, in the MSD initiative, the network leaders were not fully successful in identifying and addressing the diverse objectives of all the involved parties from the outset. This oversight led to challenges as differing goals among participants surfaced during the collaboration. Such issues might have been avoided with a more thorough alignment process, underscoring the critical importance of this function in ensuring the smooth operation and success of collaborative efforts.

> ...during the pandemic, some individuals were discussing research projects. There was a significant opportunity with Capes. Thus, we created the Minha Saúde Digital initiative. This project was coordinated with hospitals and other collaborators (Interviewee 04).
>
> Some parties had the idea that a company would be opened quickly, and it didn't work out. Doing it quickly wasn't good, and there were different interests. The original idea was research…a project that takes longer. (Interviewee 01)
>
> It was a path with personal commercial interest from some partners. A completely different line was developed from what we (founding group) wanted. (Interviewee 07)

One of the micro-governance functions that was particularly well-developed throughout the MSD project was "mobilizing", which represents the role of governance in fostering an environment of exchange and joint action, as proposed by Wegner and Verschoore (2022). Despite the challenges posed by the COVID-19 pandemic, including significant restrictions that affected how the group could interact and collaborate, the project leader, acting as orchestrator,

effectively maintained the momentum and engagement of the participants. The leader's ability to keep the group motivated, even in the face of such difficulties, was a testament to the strength of the mobilizing function. This function played a critical role in ensuring that the participants remained committed to the project's goals, demonstrating the leader's skill in rallying the team and sustaining their collective energy throughout a period of unprecedented challenges.

In addition to navigating the pandemic, the project also faced other critical moments where conflicts arose due to differing goals among the participants. During these times, some members considered stepping away from the project. However, it was the project leader who successfully kept everyone involved, reinforcing the group's cohesion and commitment. Notably, the project has been ongoing for four years, with many actors continuing their involvement even after the initial sense of urgency brought on by the pandemic had passed. This sustained engagement underscores the effectiveness of the mobilizing function in not only overcoming immediate crises but also in fostering long-term dedication to the project’s objectives.

> We didn't get all the information we had planned. We had several technical problems with the hospital data. But we adapted our own research projects and achieved a good result with what we had available. (Interviewee 08)

The evidence from interviews and observation of documents showed that the functions of organizing and integrating were pivotal throughout the MSD project. Organizing, as defined by Wegner and Verschoore (2022), involves coordinating the human, financial, technological, and legal factors to instigate organizational development and establish processes and routines. Integrating refers to engaging the members who have already joined the network and identifying their resources and capabilities. These functions are closely linked to the general objectives of a data collaborative, which aims to share data and knowledge to address wicked problems (Susha et al., 2018; Klievink et al., 2018; Verhulst and Sangokoya, 2015).

In the MSD case, these functions were manifested in several ways, including the development of contracts, continuous alignment of processes, and efforts to bring new actors into the collaborative. At this point, some of the challenges typically encountered in data collaboratives also became clear. As Gupta and George (2016) write, companies need a combination of certain tangible (e.g., data and technology), intangible (e.g., managerial and technical skills), and human resources capabilities to build a data analysis capability. In the case of MSD, it was no different; there were challenges to be overcome in relation to data

interoperability. As mentioned in the case description, some actors chose not to participate as they understood that they needed to evolve technically.

Most interestingly, the interviews revealed that these functions were distributed within the collaborative, with different leaders assuming responsibility for each stage or division of the tasks. It was not solely the role or responsibility of the orchestrator or convener to carry out these tasks, highlighting a more distributed approach to leadership and task execution within the network (Strasser et al., 2022).

> In this collaborative leadership, there must be moments when someone gives in so that another party can take the lead on a specific point, be it technological, political, or any other. Otherwise, it won't move forward. (Interviewee 08)

> It has been a difficult experience. We need to take a series of precautions, especially in communication, to ensure that we do not have situations that generate "sensitiveness" on some sides. (Interviewee 02)

As a network matures, functions like alignment and organizing become less prominent since members develop a history of cooperation and mutual understanding (Rosa; Wegner, and Del Ben, 2023). This observation held partially true in the MSD case, where it was noted that among those surveyed, including the leaders of the project and the team of the derived project, individuals with a history of collaboration found these functions to be less evident. However, in situations like the attempt to introduce a new actor to the collaborative, the alignment function, along with the arbitrating function, proved essential in keeping the work on track. The "arbitrating" function appeared at different times, both in the original MSD project as well as in the pilot project. It proved to be a very important function to maintain actors engaged with the projects and aligned with objectives. According to Susha et al. (2023), tensions occur depending on how different objectives and levels of control over data arise, bringing an additional alignment challenge to the convener. The research indicates that micro-governance functions are vital in navigating these shifts, ensuring that the collaboration continues to function effectively even as individual motivations and goals change.

> I didn't like her stance. She wanted to use the work done for a different purpose, which had not been agreed upon before hand. I said that if she continued like that I wouldn't be a part of it. He (the orchestrator) reassured me that it would be ok. (interviewee 9)

> When there were conflicts between the actors, we had to get together, talk and go back to what was agreed, understanding how everyone was important to deliver the expected results. We didn't want the siloed view of healthcare...the one who pays for this is the patient, the system, the manager without the information to treat the patient. (Interview 5)

According to Wegner and Verschoore (2022) monitoring is the function that closely examines actions and results, ensuring that activities were carried out and goals were achieved. Rosa, Wegner and Del Ben (2023) emphasize the critical importance of monitoring within collaborative networks, advocating for robust mechanisms to evaluate the effectiveness and outcomes of these collaborations. They argue that such monitoring is not just a procedural necessity but a fundamental tool for ensuring that network objectives are met and for identifying areas needing improvement. Continuous assessment, as they suggest, helps maintain goal alignment among participants and allows for the adaptation of strategies to meet emerging challenges, which is crucial in the dynamic and evolving nature of collaborative networks. In the case of the MSD initiative, rigorous monitoring was indeed a key aspect of the project. The initiative, rooted in research, naturally prioritized the delivery of clear and tangible outputs. As the pilot project, it successfully produced a concrete product, demonstrating the practical application of its research findings. However, while these deliverables were achieved and monitored effectively, they did not provide a direct means of assessing whether the overarching wicked problem, the interoperability issue, was resolved. The monitoring focused on the progress and success of the proposed solutions rather than on measuring the resolution of the problem itself.

Interestingly, this apparent disconnect between the deliverables and the ultimate goal did not seem to be problematic for the participants. Interviews conducted as part of this research revealed that the lack of measurable outcomes related to the wicked problem was not perceived as a significant challenge. This suggests an understanding among participants that wicked problems, by their nature, resist straightforward solutions and require iterative, long-term approaches that may not yield immediate, quantifiable results. This perspective aligns with the insights of Moore and Khagram (2023), who argue that in the context of complex public challenges, traditional measures of success may not adequately capture the public value being created.

This perspective is consistent with broader insights into collaborative networks addressing complex social issues. Such problems are intricate and multifaceted, often necessitating gradual, sustained efforts over time. While the MSD initiative confirms the necessity of robust monitoring mechanisms to ensure the sustainability of data collaborations, it also highlights the limitations of such monitoring when it comes to directly measuring the resolution of a wicked problem. The findings suggest that while monitoring is crucial for tracking progress and refining strategies, its relevance may lie more in assessing intermediate

objectives rather than in determining whether the ultimate, complex problem has been fully resolved, and this acceptance of ambiguity is both recognized and accepted by those involved.

According to Wegner, Hölsgens, and Bitencourt (2023, p.8), "Legitimizing through communication raises awareness of the initiative with potential users, or possibly even with detractors who disagree with the idea." In the original MSD project, which focused primarily on academic production, this function was not a visible concern. The emphasis was on scholarly output rather than market engagement, and as a result, there was little to no effort to communicate the project's value to the broader market. In contrast, the MSD pilot project showed a marked shift in focus, with the legitimizing function becoming much more prominent. An analysis of the available documents reveals that the project leaders actively communicated with the market through various digital channels, news outlets, and even by participating in an award. These efforts indicate a clear commitment to establishing the project's legitimacy and raising awareness among potential stakeholders, including the market.

This MSD Pilote Project focus aligns with Provan and Kenis's (2008) argument that legitimacy is a critical factor in the success of collaboration networks. They emphasize that for a network to be effective, it must be perceived as legitimate by both internal and external stakeholders. In the context of the MSD pilot project, the active communication with the market and broader efforts to engage stakeholders can be seen as a strategic move to enhance the network's input and output legitimacy. During interviews, it became evident that for the interoperability solution to move beyond just an experiment, patient engagement will be crucial. Engaging patients in understanding the benefits of the solution is necessary to overcome the barrier posed by some hospitals' lack of interest. This highlights the need for the legitimizing function to extend beyond communication with the market and to include efforts to engage end-users directly, ensuring that the initiative gains the necessary traction to succeed in the long term.

> The ideal would be to reach a point where the patient arrives at the health unit and requests that their data be there. That way, everyone will be in a position of obligation to make this happen. (Interviewee 10)

> Our project was recognized and awarded as a health reference in 2022 in the Administrative-Financial Management category, evaluated by a technical jury. (Interviewee 05)

In conclusion, the evidence demonstrates that the functions of micro-governance were indeed present in the analyzed case. While there were opportunities for some functions to be performed with greater focus, their execution still made a meaningful difference in the final

outcome. In the same way, the challenges commonly faced by data collaboratives, as outlined in the studied bibliography, were also experienced in this case. The micro-governance functions acted as responses to these challenges, supporting their resolution and reinforcing the collaborative process. This underscores the significant relevance of micro-governance functions and their impact on fostering a more collaborative environment, as advocated by Wegner and Verschoore (2022). Chart 06 shows the micro-governance functions, their description within MSD case and examples of evidences found for each function.

Chart 06 - Micro-Governance Functions: MSD Descriptions and Evidence

Function	MSD: Description of Function	Evidence
Designing	The partners were defined before the MSD project began. However, as the project progressed, it became clear that some had different goals, which created challenges.	Some of the parties involved at the beginning, who showed interest and said they wanted to participate, in practice ended up getting very little involvement. They participated in meetings, followed the project, but didn't get involved much beyond that...they were more interested in staying in the loop than actually collaborating (Interviewee 04).
Aligning	Carried out by the orchestrator and the researchers, this was an essential activity, particularly in terms of understanding how to achieve the objective.	It was a path with personal commercial interest from some partners. A completely different line was developed from what we (founding group) wanted. (Interviewee 07)
Mobilizing	Mobilizing was an important function, as the initiative extended over more than four years, part of which occurred during a pandemic, and the results take time to materialize.	We didn't get all the information we had planned. We had several technical problems with the hospital data. But we adapted our own research projects and achieved a good result with what we had available. (Interviewee 08)
Organizing	Organizing was a collaboratively performed function, as several workstreams were in action at the same time	In this collaborative leadership, there must be moments when someone gives in so that another party can take the lead on a specific point, be it technological, political, or any other. Otherwise, it won't move forward. (Interviewee 08)
Integrating	This role was very important at the beginning and throughout the project. However, it was carried out by more than one person.	I didn't like her stance. She wanted to use the work done for a different purpose, which had not been agreed upon before hand. I said that if she continued like that I wouldn't be a part of it. He (the orchestrator) reassured me that it would be ok. (interviewee 9)
Arbitrating	It was a necessary function when conflicts arose, and it was through Arbitrating that the derived initiative emerged as a solution to these conflicts	Some parties had the idea that a company would be opened quickly, and it didn't work out. Doing it quickly wasn't good, and there were different interests. The original idea was research, a project that takes longer.
Monitoring	Monitoring happened naturally due to the format and objective of the initiative	We had a series of publications. We have to show the execution report to Capes. But the results are very good.
Bridging	Not identified	NA
Legitimizing	Strongly identified in the pilot project, this innovation requires ongoing communication	Our project was recognized and awarded as a health reference in 2022 in the Administrative-Financial Management category, evaluated by a technical jury

Source: Elaborated by the author (2024)

5 THE ACTORS AND ROLES OF MICRO-GOVERNANCE

As seen in the previous section, the evidence confirms the challenges faced by data collaboratives described in the literature review. The research results empirically demonstrate that micro-governance functions are present in the data collaboration initiative studied, playing a crucial role in creating a collaborative environment that enables the initiative to achieve its objectives. This section will deepen these findings and respond to the propositions brought up in the literature review, showcasing the empirical validation of the theoretical claims. The first proposition posits that governance of downstream problems, as discussed by Sørensen and Torfing (2021), establishes an enabling environment for the long-term sustainability of data collaborations. The second proposition asserts that the role of conveners should include establishing micro-governance functions to foster a collaborative environment, as highlighted by Wegner and Verschoore (2022).

The voluntary nature of participation in data collaboratives introduces additional complexity to the dynamics within these initiatives (Susha et al., 2018). Despite a general willingness to engage, the lack of maturity in both the market and the participating companies, particularly concerning big data technology, presents significant challenges. Numerous practical issues arise, ranging from technical problems related to interoperability to obtaining authorization to share data, often limiting the volume of information available for collaboration (Klievink et al., 2018; Ruijer, 2021; Susha et al., 2023). Many organizations hesitate to share data, perceiving it as highly valuable and fearing the loss of their competitive edge. Additionally, these organizations often lack a comprehensive understanding of data intelligence and its full potential, exacerbating their reluctance to share (Susha et al., 2023). Furthermore, some organizations join these initiatives merely to ensure their involvement and feel secure, yet fail to genuinely engage in the collaborative process. This lack of true commitment hinders the overall effectiveness of the data collaborative, as these participants do not fully contribute to the collective effort.

> Some of the parties involved at the beginning, who showed interest and said they wanted to participate, in practice ended up getting very little involvement. They participated in meetings, followed the project, but didn't get involved much beyond that...they were more interested in staying in the loop than actually collaborating. (Interviewee 04).

For a data collaborative to function effectively and achieve long-term sustainability, the parties involved must be open to evolving their data culture. Participants need to embrace a

data-driven decision-making approach and foster a shared commitment to leveraging data (Susha et al., 2023). This requires not only constant strategic alignment of interests and recognition of the mutual benefits of collaboration but also a proactive approach to addressing the daily challenges that arise from the varying levels of maturity among the parties involved. Without these elements, the potential for conflict and misalignment increases, which can jeopardize the sustainability of the collaborative effort (Wegner; Hölsgens and Bitencourt, 2023). In a recent study, Wegner, Hölsgens and Bitencourt (2023, p.3) argue that: "orchestrating distinct partners requires an alignment of interests, the integration and organization of resources, the mobilization and monitoring of partners, and the arbitration of potencial conflicts between them". The evidence further illustrates that the key challenge was not merely the obstacles themselves but the actors' willingness and commitment to overcoming them, which ultimately determined the success of the collaboration.

> ...actors are not willing to interoperate. Data is a very valuable asset. And it is very difficult for hospitals to be interested. At some levels, they are. But the one who is really involved is the government or the citizen. (Interviewee 04).
>
> Some companies already have this notion, this awareness, that they are part of society and need to contribute to improving population health. Others are only concerned with profit. (Interviewee 06).

This perspective aligns with the idea that effective governance of downstream problems involves not only addressing logistical and operational issues but also fostering a culture of commitment and adaptability among participants (Sørensen and Torfing, 2021). It is not just that the selection of actors should consider the willingness of participants to engage constructively, collaborate, and work towards shared goals, as advocated by Emerson, Nabatchi and Balogh (2012). Sørensen and Torfing (2016) argue that introducing new forms of social accountability can help address the motivation of actors within collaborative governance. This approach suggests that accountability mechanisms can enhance participant commitment, ensuring that the actors are not only willing but also incentivized to actively contribute to the collaborative process. Moreover, while Sørensen and Torfing (2016) emphasize the importance of solving upstream governance problems, they also caution against allowing process-related issues to overshadow the need for securing desirable outcomes from collaborative governance. Therefore, addressing the daily challenges that arise from varying levels of maturity among the parties involved is necessary to prevent conflict and misalignment, which can jeopardize the sustainability of the collaborative effort. Thus, the evidence from this research supports the proposition that governance of downstream problems will establish an enabling environment.

Proposition 01: Governance of downstream problems will establish an enabling environment for the long-term sustainability of Data Collaborations.

Regarding the second proposition made in the theoretical section of this thesis, the research evidence partially confirms it. It is indeed confirmed that the micro-governance functions should be performed, as these functions are vital in fostering a collaborative environment, building trust, reducing power imbalances, and establishing clear processes and routines within data collaborations (Wegner and Verschoore, 2022). By performing these functions, collaborative participants can address the various challenges that arise in collaborative settings and promote a culture of cooperation and mutual respect among stakeholders. The importance of these functions is well-supported by this research, as seen in the previous section, and aligns with existing literature. The evidence demonstrates that the governance of downstream problems, combined with the micro-governance functions, can significantly enhance the collaborative environment and ensure the long-term sustainability of data collaborations. By addressing the theoretical and empirical challenges identified, stakeholders can create a more effective and resilient framework for collaboration, ultimately leading to greater success and impact in their initiatives.

The research evidence also supports the critical roles of the convener and the orchestrator in collaborative networks. As established by Provan and Kenis (2008), Huxham and Vangen (2005), and Dhanaraj and Parkhe (2006), there must be a clear distinction and understanding the roles of these two actors. Conveners are essential in the initial stages, focusing on the formation and facilitation of the collaborative environment (Harrison; Pardo and Cook, 2012). They bring together diverse stakeholders, setting the foundation for successful collaboration. Furthermore, the literature on data collaboratives, particularly the work of Susha et al. (2023), underscores the critical role of conveners in initiating and establishing collaborative efforts. Conveners contribute to building trust, legitimacy, and effective governance structures, which are essential for navigating multi-stakeholder environments (Susha et al., 2023) and exactly what is expected as intermediate outcomes of micro-governance functions (Wegner and Verschoore, 2022). Thus, evolving in the vision of Susha et al. (2023) on the role of conveners in the theoretical framework section of this thesis, the following proposition was made:

Proposition 02: The role of conveners should encompass establishing micro-governance functions to achieving a collaborative environment.

The research confirms the significant role of the convener in the performance of micro-governance functions, yet it also advances our understanding by raising a crucial question: Who should bear the responsibility for these functions? The debate is not about the necessity of these functions, as stated at the beginning of this section, the evidence collected during the research underscores the significance of micro-governance functions in fostering a collaborative environment within data collaborations. When effectively performed, these functions contribute substantially to the cohesion and success of the collaborative effort (Wegner and Verschoore, 2022). However, the debate is rather about identifying the appropriate roles and responsibilities within the collaborative framework. Clarifying these roles is vital to ensuring the effective execution of micro-governance functions and the success of the collaboration in achieving its goals (Susha et al., 2023).

Initially, it was proposed that conveners should be responsible for establishing these micro-governance functions to achieve a collaborative environment. However, the research reveals that this role should not be exclusive to the convener or orchestrator, aligning with the vision of Susha et al. (2023). Instead, the evidence suggests that the functions of micro-governance within a collaborative initiative are most effective when distributed among multiple individuals rather than concentrated in a single actor. This distributed approach aligns with Susha et al. (2023), who emphasize the importance of conveners in setting the stage for collaboration, while also recognizing that the complexities of managing a data collaborative require contributions from multiple actors across the ecosystem. This viewpoint is further supported by Strasser et al. (2022), who highlight that orchestrating collaborative networks, especially for social innovation, is a complex task necessitating leadership that is distributed across various individuals and organizations.

While Wegner and Verschoore (2022) introduce the concept of "network leaders" as those responsible for performing micro-governance functions, they leave the specifics of this leadership profile undefined. Susha et al. (2023) complement this by detailing the convener's role in fostering trust and alignment, yet they also acknowledge that other actors must take on leadership responsibilities to sustain the collaborative effort. In the case studied, the importance of the orchestrator's role was evident. The orchestrator's involvement was pivotal in keeping the project on track, even during periods of difficulty. Orchestrators take on a more active role in managing the collaboration by coordinating activities, aligning resources, and ensuring

smooth and effective progress (Dhanaraj and Parkhe, 2006). They handle the ongoing operational management and strategic direction, making sure that the collaborative efforts are sustained over time.

The empirical results confirm that both roles are indispensable, while conveners establish the groundwork, orchestrators maintain and advance the collaboration's goals. Clear delineation of these roles and understanding which actor will assume each role is essential for the success and sustainability of data collaboratives. It becomes clear that the roles of conveners and orchestrators in micro-governance are distinct yet complementary. Therefore, the role of the conveners aligns with the micro-governance functions of designing and aligning, typically associated with the initial phase of collaboration. While, the orchestrator's role, in particular, is critical in managing the collaboration's day-to-day activities and ensuring its long-term viability. Thus, these findings advance knowledge about micro-governance, providing the opportunity to develop two new propositions:

Proposition 02a: The conveners primarily perform the micro-governance functions of Designing and Aligning to set clear expectations and create an environment conducive to effective cooperation.

Proposition 02b: The orchestrators primarily perform the micro-governance functions of Mobilizing, Integrating, Organizing, Arbitrating, Monitoring, Bridging, and Legitimizing to sustain an environment conducive to effective cooperation.

Wegner, Hölsgens and Bitencourt (2023) deepens this understanding by drawing parallels between orchestrators in collaborative technology networks and those in networks for social innovation, highlighting the unique demands of addressing social issues through collaboration. They argue that the complexity and collaborative nature of social innovation require a distributed leadership model, where multiple stakeholders, rather than individual leaders, orchestrate the network. The findings of this thesis are in line with Wegner, Hölsgens and Bitencourt (2023), indicating that key governance functions are carried out by various engaged participants who, even without formal leadership positions, play critical roles in the project's success. This distributed approach leverages the unique strengths and perspectives of different individuals, enhancing the collaborative effort by ensuring that all necessary functions are effectively performed. According to Strasser et al. (2022), the role of the orchestrator is crucial in empowering and supporting the network’s collective capacity to achieve

transformative impact. This also underscores the importance of commitment and alignment among those responsible for these functions, ensuring their actions consistently advance the collaborative goals. Given the discussions throughout this thesis, our analysis proposes that the micro-governance functions within any data collaborative are influenced by the specific actors and roles involved. While these functions may vary according to these factors, they are consistently present and should be executed by multiple leadership figures within the initiative to ensure effective governance and collaboration. Thus:

Proposition 03: The orchestrators distribute micro-governance functions among leading actors to sustain an environment conducive to effective cooperation

Klievink et al. (2018) emphasize that policymakers play a crucial role in data collaborations, particularly within data-sharing initiatives. They often act as facilitators, establishing the necessary conditions and frameworks for success. The evidence of this research support this perspective. The interviewees consistently identifying the policy-makers as crucial in motivating and committing all participants to the collaborative effort, creating an environment conducive to data sharing. In these data-sharing contexts, the policymakers' facilitation includes creating incentives for participation and clearly communicating the benefits of collaboration to all stakeholders, thereby fostering a conducive environment for cooperation. Interviewees highlighted the importance of government involvement in reinforcing interoperability and setting standards.

> In the data context, government creates basic legislation and sets standards…so government is an important actor in helping to reinforce that interoperability and the need for standards. (Interviewee 06)
>
> In countries where interoperability works, it works because there is government regulation that requires it. (Interviewee 02)
>
> I don't need to have a government actor involved, but it is very important. In Brazil, it's essential to provide direction in regulation. Through regulation, the parties must meet certain standards and compliance requirements, which helps accelerate the process. However, it's not mandatory. Yet, in the reality of Brazil, where 80% of healthcare is public, it ends up being essential. (Interviewee 05)

These insights underscore that while government involvement may not always be mandatory, it is often essential, especially in contexts like Brazil’s public healthcare system. According to interviews, the presence of policymakers ensures that data collaborations adhere

to necessary standards and compliance requirements, thereby facilitating smoother and more effective collaboration.

Furthermore, the study reveals a current immaturity in organizations regarding data culture, technological evolution, and trust in data sharing (Klievink et al., 2018; Ruijer, 2021; Susha et al., 2023). Thus, the role of the policymaker as a public entity participating in the initiative becomes even more critical, particularly in governance. The policymaker is not only a facilitator (Klievink et al., 2018), but also plays a pivotal role in overseeing the governance of the project, either entirely or in collaboration with other actors. This involvement is essential for ensuring that data is handled responsibly and securely, which is vital for building trust among participants. Additionally, the policymaker is perceived as the primary actor responsible for establishing interoperability standards and defining the rules of the game. These standards are crucial for the success of data collaborations, as they help mitigate technical and procedural discrepancies that could otherwise hinder effective collaboration. By being deeply involved in governance, the policymaker actor ensures compliance with these standards, thereby streamlining processes and reducing friction in data sharing.

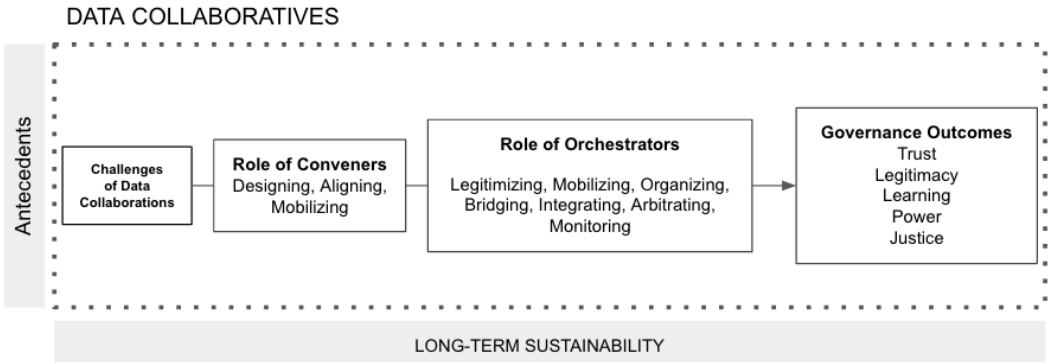
This perspective is consistent with Moore and Khagram (2004) discussion on legitimacy and public value creation. They argue that for public initiatives and collaborations to be effective and sustainable, they must achieve and maintain legitimacy through stakeholder support and alignment with public values. Policymakers, by setting standards, creating incentives, and ensuring compliance, contribute significantly to the legitimacy of data collaborations. Their role in governance helps ensure that collaborations are perceived as valid and valuable by all stakeholders, thereby enhancing their credibility and effectiveness. Moore and Khagram (2004) highlight that legitimacy involves not just meeting immediate operational goals but also addressing broader societal needs and expectations. In this way, the policymaker's involvement in data collaborations is pivotal in creating a stable, trustworthy environment that supports long-term success and public value creation.

The research advocates that the policymaker actor’s role in data collaborations is multifaceted and vital, particularly in its governance capacity. By taking an active role in governance, facilitating participation, ensuring secure data sharing, and setting interoperability standards, the policymaker actor helps create a stable and trustworthy environment for data collaborations. While successful initiatives without policymakers actor involvement are conceivable, the evidence indicates that the government's deep engagement in governance significantly enhances the likelihood of long-term sustainability and success in data collaborations.

Proposition 04: Policymakers perform the microgovernance functions of Alignment, Mobilization, and Integration to stimulate collaborative efforts toward broader public objectives

After conducting the case study, analyzing the evidence and the discussion presented in this section, the framework proposed at the end of the theoretical foundation section was refined and adjusted to better reflect the dynamics observed in practice. The updated figure builds upon the theoretical foundation presented earlier, integrating insights from the empirical findings to offer a more comprehensive view of the governance process in data collaborations. This revised model continues to draw from the Wegner and Verschoore (2022) framework, while incorporating the nuanced roles and additional considerations highlighted by Wegner, Hölsgens and Bitencourt (2023). It emphasizes the distinct responsibilities of both conveners and orchestrators in addressing the challenges and achieving the desired governance outcomes. The evolution of this model underscores the necessity of ongoing adjustments to governance strategies as data collaborations progress, ensuring that the micro-governance functions are effectively distributed and executed to maintain the collaboration's success.

Figure 03 - Governance Framework for Long Term Data Collaborations



Source: based on Wegner and Verschoore (2022) and Wegner, Hölsgens and Bitencourt (2023)

The second framework introduces a more nuanced division of roles between conveners and orchestrators within data collaboratives. In the original framework, the "Role of Conveners" was responsible for a broad set of micro-governance functions, including designing, legitimizing, aligning, mobilizing, organizing, bridging, integrating, arbitrating, and monitoring. The revised framework, however, separates these responsibilities, assigning the conveners a more focused role that includes only designing, aligning, and mobilizing. The additional functions—legitimizing, organizing, bridging, integrating, arbitrating, and

monitoring—are now attributed to orchestrators. This change reflects a clearer delineation between the strategic setup managed by conveners and the ongoing coordination and management tasks handled by orchestrators. The overall structure remains focused on achieving governance outcomes like trust, legitimacy, learning, power, and justice, within the context of long-term sustainability, but the distinction between conveners and orchestrators in the revised framework allows for a more specialized approach to micro-governance within collaborative networks.

6 CONCLUSIONS, LIMITATIONS AND AVENUES FOR FURTHER RESEARCH

The increasing recognition of the value of data sharing for purposes such as research, policy development, and innovation has led to the emergence of data-sharing initiatives for the common good (Susha et al., 2020; Klievink et al., 2018; Bartalucci, 2023). Despite growing evidence of their benefits and value, most initiatives remain limited, one-off experiences, struggling to achieve long-term sustainability (The Gov Lab, 2023; Susha and Gil-Garcia, 2019). This research upholds that to ensure the long-term sustainability of these initiatives, it is necessary to expand the knowledge and practices surrounding governance in data collaborations. A broad conceptualization of governance is insufficient for addressing the specific challenges and ensuring these partnerships thrive over time.

This thesis proposes that the challenges faced by data collaboration initiatives can be addressed by viewing them through the lens of downstream problems (Sørensen and Torfing, 2021). Specifically, the functions proposed by micro-governance will support the development of a collaborative environment (Wegner and Verschoore, 2022). This proposal is significant because data collaborations behave differently from other collaboration networks, and few studies have focused on governance, governance models, or the role of governance in sustaining data collaborations (Klievnik et al., 2018; Susha et al., 2018; Ruijer, 2021). Data collaborations present not only practical challenges but also remain a challenging area of study.

Drawing on qualitative research with an exploratory approach, this thesis conducted an in-depth case study of the Minha Saúde Digital (MSD) initiative. The case study provided empirical evidence to address the research question: What role does governance play in sustaining data collaborations for the common good? The findings were categorized based on challenges identified in the literature and framed through the lens of micro-governance. The study advances the field by proposing several propositions about governance in data collaborations.

The findings support Proposition 01: governance focused on downstream problems creates an enabling environment for the long-term sustainability of data collaborations. The MSD initiative exemplifies this, as it addressed challenges such as trust-building, alignment of stakeholder interests, and coordination over several years. This reinforces the argument that governance should not only set up initial structures but also adaptively address operational and emergent issues. The results demonstrate the operationalization of micro-governance functions within a data collaborative, empirically providing evidence that these functions act as drivers of effective collaborative governance.

Propositions 02a and 02b build upon and advance the propositions by Wegner and Verschoore (2022) distinguishing between the roles of conveners and orchestrators in performing micro-governance functions. Conveners primarily focus on Designing and Aligning to establish clear expectations and create an environment conducive to collaboration. In contrast, orchestrators perform a broader set of functions—including Mobilizing, Integrating, Organizing, Arbitrating, Monitoring, Bridging, and Legitimizing—to sustain cooperation over time. These findings build on and refine existing literature, offering a nuanced understanding of leadership roles within data collaborations. This thesis contributes to the theory by critically examining the roles of actors within data collaborations, specifically questioning and clarifying which micro-governance functions should be performed by conveners and which should fall under the purview of orchestrators. Aligned with the argument put forth by Wegner, Hölsgens, and Bitencourt (2023), who suggest that initiatives may need more than one leader over time and with Strasser et al. (2022) that emphasize that network leadership is a distributed practice involving various individuals and organizations that support transformative capacity development.

Furthermore, Proposition 03 emphasizes that orchestrators distribute micro-governance functions among leading actors to sustain effective collaboration. Evidence from the MSD initiative supports this assertion, demonstrating how distributed leadership can address complex challenges and maintain engagement over extended periods. This challenges traditional, centralized governance models by highlighting the need for shared responsibilities. Policymakers' Role in Collaborative Governance: Proposition 04 highlights the critical role of policymakers in performing the micro-governance functions of Aligning, Mobilizing, and Integrating to stimulate collaborative efforts. In the MSD case, policymakers contributed to fostering alignment between diverse stakeholders and mobilizing resources, underscoring their importance in bridging public objectives with collaborative governance.

In terms of contributions to the practice, this study provides valuable insights for leaders and participants in data collaboratives on fostering effective collaborative environments. Leaders must carefully assess the context to strike the right balance of governance functions. Additionally, having leaders with a strong understanding of governance and collaboration skills is essencial. In data collaboratives, these leaders must be equipped to navigate both the inherent challenges of collaboration and those that emerge from the innovative nature of the format.

Nonetheless this study has limitations that must be taken into consideration. Firstly, although Minhas Saúde Digital serves as a representative case of data collaboration, it is one of a very limited number of examples and should be considered unique within its specific context.

A larger number of cases could add more robustness to the results. Culturally, the case is also contingent upon the specific reality and technological maturity of the parties involved. While this study focuses on the health sector, it is important to note that other sectors may encounter different experiences, particularly regarding the maturity of the stakeholders and their understanding of their roles within the collaborative framework. Additionally, micro-governance functions may vary across different fields, potentially leading to diverse interpretations and implementations. A further limitation lies in the research timeframe. This thesis aimed to explore the role of governance in the long-term sustainability of a data collaborative. Although the MSD project was followed over several years, indicating potential longevity, the research remains temporally limited, as the case study has not yet reached its conclusion at the time of submission.

Despite its limitations, this research holds significant value. The field of data collaboration is still in its early stages of development, and examining it from an academic perspective can greatly enrich both theoretical understanding and practical application. Additionally, research on micro-governance functions and downstream problems remains an open field, offering ample opportunities for exploration. By combining these two paradigms, this study opens up new possibilities for future research, providing a foundation for further investigation into the complexities and nuances of data collaboration governance.

Considering the theoretical and managerial findings of this research, as well as its limitations, some topics for future research are suggested here, which can deepen knowledge on the topics covered. To further advance the understanding of data collaboration, future research should focus on examining cases that have been in existence for an extended period of time. Although this study was constrained by the scarcity of long-standing examples, it is acknowledged that there are currently evolving cases that can be monitored and analyzed over medium- to long-term durations. The academic research conducted in the first chapter of this thesis clearly indicates that markets outside of Brazil are more advanced in data collaboration, with a greater number of examples available globally (GovLab, 2023). Therefore, it is relevant to investigate the presence of micro-governance functions in a more mature contexts of collaboration and data sharing. Additionally, it would be beneficial to study how micro-governance functions manifest in data collaborations involving a diverse array of actors, including those with significant policymakers involvement or varying numbers of participants. Also, it is suggested that future research should delve deeper into the role of a policymaker participant. Specifically, should explore the level of maturity in relation to data culture and technology alongside policymakers participation. An important question is whether

policymakers involvement is necessary to ensure long-term sustainability given the lack of maturity. Additionally, it should be examined whether the need for a policymakers as a regulator diminishes as organizational maturity increases. Finally, while this research did not identify any new micro-governance functions beyond those already discussed in the literature, exploring more complex cases or collaborations with different hierarchical structures could provide valuable insights. Such studies would enhance the understanding of the dynamics and sustainability of data collaborations in various contexts.

REFERENCES

Ansell, C., and Gash, A. (2008). Collaborative governance in theory and practice. Journal of public administration research and theory, 18(4), 543-571.

Bardin, L. (2015) Análise de conteúdo. São Paulo: Edições 70.

Bartalucci, S. (2023). Designing data collaboratives' governance for their long-term stability: a key factor analysis through a multiple case study.

Bello-Orgaz, G., Jung, J. J., and Camacho, D. (2016). Social big data: Recent achievements and new challenges. Information Fusion, 28, 45-59.

Beyer, M., Laney, D. The Importance of 'BigData': A Definition, Gartner, Stamford, CT (2012).

Bianchi, C., Nasi, G., & Rivenbark, W. C. (2021). Implementing collaborative governance: models, experiences, and challenges. Public Management Review, 23(11), 1581-1589.

Blazquez, D., and Domenech, J. (2018). Big Data sources and methods for social and economic analyses. Technological Forecasting and Social Change, 130, 99-113.

Borzacchiello, M. T., and Craglia, M. (2012). The impact on innovation of open access to spatial environmental information: A research strategy. International Journal of Technology Management, 60(1-2), 114-129.

Bollier, D., and Helfrich, S. (Eds.). (2012). The wealth of the commons: A world beyond market and state. Levellers Press.

Bryson, J. M., Crosby, B. C., and Stone, M. M. (2015). Designing and implementing cross-sector collaborations: Needed and challenging. Public administration review, 75(5), 647-663.

Cabral, S., & Krane, D. (2018). Civic festivals and collaborative governance. International Review of Administrative Sciences, 84(1), 185-205.

Cai, L., and Zhu, Y. (2015). The challenges of data quality and data quality assessment in the big data era. Data science journal, 14, 2-2.

Cao, G., Duan, Y., & Li, G. (2015). Linking business analytics to decision making effectiveness: A path model analysis. IEEE Transactions on Engineering Management, 62(3), 384-395.

Carman, J. G., and Fredericks, K. A. (2010). Evaluation capacity and nonprofit organizations: Is the glass half-empty or half-full? American Journal of Evaluation, 31(1), 84-104

Creswell, J. W., & Poth, C. N. (2016). Qualitative inquiry and research design: Choosing among five approaches. Sage publications.

Davenport, T. H., & Bean, R. (2019). Big data and AI executive survey (2019). NewVantage Partners (NVP), Tech. Rep.

Dhanaraj, C., & Parkhe, A. (2006). Orchestrating innovation networks. Academy of management review, 31(3), 659-669.

Dulong de Rosnay, M., and Stalder, F. (2020). Digital commons. Internet Policy Review, 9(4), 1-22.

Dyer, J. H., Singh, H., & Hesterly, W. S. (2018). The relational view revisited: A dynamic perspective on value creation and value capture. Strategic management journal, 39(12), 3140-3162.

Eaton, B., Strassner, J., and Sørensen, C. (2017). Data-Driven Innovation: Big Data for Growth and Well-Being. OECD Publishing.

Ebrahim, A., & Rangan, V. K. (2014). What impact? A framework for measuring the scale and scope of social performance. California management review, 56(3), 118-141.

Edelenbos, J., & Klijn, E. H. (2007). Trust in complex decision-making networks: A theoretical and empirical exploration. Administration & society, 39(1), 25-50.

Eisenhardt, K. M. (1989). Building theories from case study research. Academy of Management Review, 14(4), 532-550.

Emerson, K., Nabatchi, T., and Balogh, S. (2012). An integrative framework for collaborative governance. Journal of public administration research and theory, 22(1), 1-29.

George, J. J., Yan, J., and Leidner, D. E. (2020). Data philanthropy: Corporate responsibility with strategic value? Information Systems Management, 37(3), 186-197.

GovLab, The. Data Collaboratives Explorer. Accessed (December 28, 2023) at https://datacollaboratives.org/explorer.html

Grossman, R. L., Heath, A., Murphy, M., Patterson, M., and Wells, W. (2016). A case for data commons: toward data science as a service. Computing in science and engineering, 18(5), 10-20.

Grubb, N., and Easterling, D. (2012). Harnessing the power of data: Data and analytics for the nonprofit sector. The Foundation Review, 4(3), 5.

Gupta, M., & George, J. F. (2016). Toward the development of a big data analytics capability. Information & Management, 53(8), 1049-1064.

Gupta, A., Panagiotopoulos, P., and Bowen, F. (2020). An orchestration approach to smart city data ecosystems. Technological Forecasting and Social Change, 153, 119929.

Hall, W., Shadbolt, N., Tiropanis, T., O'Hara, K., and Davies, T. (2012). Open data and charities.

Harrison, T. M., Pardo, T. A., and Cook, M. (2012). Creating open government ecosystems: A research and development agenda. Future Internet, 4(4), 900-928.

Head, B. W. (2022). Wicked problems in public policy: understanding and responding to complex challenges (p. 176). Springer Nature.

Hogan, M., Ojo, A., Harney, O., Ruijer, E., Meijer, A., Andriessen, J., ... and Malandrino, D. (2017). Governance, transparency and the collaborative design of open data collaboration platforms: understanding barriers, options, and needs. Government 3.0–Next Generation Government Technology Infrastructure and Services: Roadmaps, Enabling Technologies and Challenges, 299-332.

Hoffman, A. Boral, and D. Olukoya, “Insight Report Data Collaboration for the Common Good Enabling Trust and Innovation Through Public-Private Partnerships,” 2019. [Online]. Available: www.weforum.org

Janssen, M., Charalabidis, Y., and Zuiderwijk, A. (2012). Benefits, adoption barriers and myths of open data and open government. Information Systems Management, 29(4), 258-268.

Klievink, B., Van Der Voort, H., and Veeneman, W. (2018). Creating value through data collaboratives. Information Polity, 23(4), 379-397.

Klijn, E. H., van Meerkerk, I., & Edelenbos, J. (2020). How do network characteristics influence network managers’ choice of strategies?. Public money & management, 40(2), 149-159.

Kwon O, Lee N, Shin B (2014) Data quality management, data usage experience, and acquisition intention of big data analytics. Int J Inf Manag 34(3):387–394

Mayer-Schönberger, V., and Cukier, K. (2013). Big data: A revolution that will transform how we live, work, and think. Houghton Mifflin Harcourt.

McAfee, A., Brynjolfsson, E., Davenport, T. H., Patil, D. J., and Barton, D. (2012). Big data: the management revolution. Harvard Business Review, 90(10), 60-68.

McCrea, A. M. (2020). Can administrative capacity address wicked problems? Evidence from the frontlines of the American opioid crisis. Administration and Society, 52(7), 983-1008.

Mikalef, P., Pappas, I. O., Krogstie, J., and Giannakos, M. (2018). Big data analytics capabilities: a systematic literature review and research agenda. Information systems and e-business management, 16, 547-578.

Moore, M., & Khagram, S. (2006). On creating public value: What business might learn from government about strategic management.

Pappas, I. O., Mikalef, P., Giannakos, M. N., Krogstie, J., and Lekakos, G. (2018). Big data and business analytics ecosystems: paving the way towards digital transformation and sustainable societies. Information Systems and e-Business Management, 16, 479-491.

Perkmann, M., and Schildt, H. (2015). Open data partnerships between firms and universities: The role of boundary organizations. Research Policy, 44(5), 1133-1143.

Porter, M. E., & Kramer, M. R. (2011). Criação de valor compartilhado. Harvard Business Review, 89(1/2), 62-77.

Poel, M., Meyer, E. T., and Schroeder, R. (2018). Big data for policymaking: Great expectations, but with limited progress? Policy and Internet, 10(3), 347-367.

Provan, K. G., & Kenis, P. (2008). Modes of network governance: Structure, management, and effectiveness. Journal of public administration research and theory, 18(2), 229-252.

Rasche, A., Morsing, M., and Wetter, E. (2021). Assessing the legitimacy of “open” and “closed” data partnerships for sustainable development. Business and Society, 60(3), 547-581.

Rittel, H. W., and Webber, M. M. (1973). Dilemmas in general planning theory. *Policy Sciences*, *4*(2), 155-169.

da Rosa, J. R., Wegner, D., & del Ben, F. (2023). Role of Resource Asymmetry and Collaboration Time in the Governance of Cross-border Collaborative Networks. Journal of Borderlands Studies, 1–21.

Ruijer, E. (2021). Designing and implementing data collaboratives: A governance perspective. Government Information Quarterly, 38(4), 101612.

Saccol, A. Z. (2009). Um retorno ao básico: compreendendo os paradigmas de pesquisa e sua aplicação na pesquisa em administração. Revista de Administração da UFSM, 2(2), 250-269.

Schilke, O., & Lumineau, F. (2018). The double-edged effect of contracts on alliance performance. Journal of Management, 44(7), 2827-2858.

Shadbolt, N., Brett, A., Chen, M., Marion, G., McKendrick, I. J., Panovska-Griffiths, J., ... and Swallow, B. (2022). The challenges of data in future pandemics. Epidemics, 100612.

Simons, H. (2014). Case study research: In-depth understanding in context. The Oxford handbook of qualitative research, 455-470.

Stalla-Bourdillon, S., Carmichael, L., & Wintour, A. (2021). Fostering trustworthy data sharing: Establishing data foundations in practice. Data & Policy, 3, e4.

Stake, R. (1995). Case study research. thousand oaks, CA: Sage.

Stempeck, M. (2014). Sharing data is a form of corporate philanthropy. Harvard Business Review, 24.

Sørensen, E., & Torfing, J. (2016). Collaborative innovation in the public sector. Enhancing public innovation by transforming public governance, 115-116.

Sørensen, E., and Torfing, J. (2021). Radical and disruptive answers to downstream problems in collaborative governance? Public Management Review, 23(11), 1590-1611.

Strasser, T., De Kraker, J., & Kemp, R. (2022). Network leadership for transformative capacity development: roles, practices and challenges. Global Sustainability, 5, e11.

Susha, I., Pardo, T. A., Janssen, M., Adler, N., Verhulst, S. G., and Harbour, T. (2018). A research roadmap to advance data collaboratives practice as a novel research direction. International Journal of Electronic Government Research (IJEGR), 14(3), 1-11.

Susha, I., and Gil-Garcia, J. R. (2019). A collaborative governance approach to partnerships addressing public problems with private data.

Susha, I., Grönlund, Å., and Van Tulder, R. (2019). Data-driven social partnerships: Exploring an emergent trend in search of research challenges and questions. Government Information Quarterly, 36(1), 112-128.

Susha, I. (2020). Establishing and implementing data collaborations for public good: A critical factor analysis to scale up the practice. Information Polity, 25(1), 3-24.

Susha, I., Rukanova, B., Zuiderwijk, A., Gil-Garcia, J. R., and Hernandez, M. G. (2023). Achieving voluntary data sharing in cross-sector partnerships: Three partnership models. Information and Organization, 33(1), 100448.

Vangen, S., & Huxham, C. (2003). Nurturing collaborative relations: Building trust in interorganizational collaboration. The Journal of applied behavioral science, 39(1), 5-31.

Verhulst, S., and Sangokoya, D. (2015). Data collaboratives: Exchanging data to improve people's lives. Accessed (December 26, 2023) at: https://medium. com/@ sverhulst/data-collaboratives-exchanging-data-to-improvepeople-s-lives-d0fcfc1bdd9a.

Verhulst, S. G., & Young, A. (2017). Open data in developing economies: Toward building an evidence base on what works and how (p. 284). African Minds.

Verhulst, S. G. (2021). Reimagining data responsibility: 10 new approaches toward a culture of trust in re-using data to address critical public needs. Data and Policy, 3, e6.

Wegner, D., and Verschoore, J. (2022). Network governance in action: Functions and practices to foster collaborative environments. Administration and Society, 54(3), 479-499.

Wegner, D., Hölsgens, R., & Bitencourt, C. C. (2023). Orchestrating collaborative networks for social innovation: Orchestrators' roles in socially innovative initiatives. Technological Forecasting and Social Change, 195, 122786.

Woolthuis, R. K., Hillebrand, B., & Nooteboom, B. (2005). Trust, contract and relationship development. Organization studies, 26(6), 813-840.

Zuiderwijk, A., and Janssen, M. (2014). Open data policies, their implementation and impact: A framework for comparison. Government Information Quarterly, 31(1), 17-29

Zuiderwijk, A., Janssen, M., and Davis, C. (2014). Innovation with open data: Essential elements of open data ecosystems. Information polity, 19(1-2), 17-33.

APPENDIX A

ROTEIRO PARA ENTREVISTA SEMI-ESTRUTURADA	
Categorias de Análise	**Questões para o Entrevistado**
Caracterização	Quem é o entrevistado, experiências, principais atuações profissionais
Projeto/Resultado	Conte sobre a trajetória do projeto. Explorar participação, momentos, rotinas, etc.
Experiências Anteriores	Que experiência o entrevistado e os envolvidos tinham com projetos desta natureza anteriormente
Processos Busca e Identificação de Oportunidades	Como iniciou o projeto? Quem foi a tomada de decisão inicial? Como o projeto foi discutido antes de iniciar (papéis, responsabilidades, resultados esperados)?
Identificação dos Fatores Contextuais	Quais eram os incentivos para participação de cada ator? Na prática, os papéis definidos no inicio se mantiveram? Houve algum conflito durante o projeto, relacionado com definições feitas antes do projeto iniciar?
Identificação das Dificuldades Enfrentadas	Quais foram as principais dificuldades enfrentadas? (primeiro de forma geral, depois: em relaçao ao trabalho com os diferentes atores, em relação as entregas dos resultados, em relaçao as objetivos iniciais)
Processos Relacionados ao papel da Liderança	Qual o formato e qual deve ser o papel da liderança? Como eram os processos de tomada de decisão?
Processos Relacionados as Funções e Práticas	Durante o decorrer do processo, como ocorriam os alinhamentos necessários? Como se dava a comunicação e troca de informações? Eram realizadas atividades para engajamento e motivação? Em momentos de conflito, como eles eram resolvidos? Havia acompanhamento dos resultados no decorrer do projeto? De quem era o papel de realizar essas ações?
Identificação dos Resultados	Como foram os resultados da iniciativa? Foram dentro do esperado? Como eles foram comunicados? Houve comunicação?
Cenário Futuro	Existe previsão de continuar com o projeto ou outros spinoff?