

GOVERNANCE CHALLENGES IN OPEN GOVERNMENT DATA ECOSYSTEMS: A CASE STUDY FROM THE FINANCIAL SECTOR IN NORWAY

Christian R. Schultz
University of Oslo
Chrisc@ifi.uio.no

Alexander M. Kempton
University of Oslo
alexansk@ifi.uio.no

Abstract

In this paper, we examine how loose coupling between different stakeholders in an ecosystem creates data governance challenges regarding four topics: maturity, goals, data access, and data quality. Ecosystem and platform governance theory lacks real-life cases and is more focused on single-level than multi-level ecosystems. By examining three subcases in a data ecosystem in the Norwegian financial sector, we found that the complexity of an ecosystem is greater than what is portrayed by existing theory. This paper provides insights into ecosystems and their complexities and highlights important topics for understanding the governance challenges associated with loose coupling between stakeholders. Paths for future research are provided as a necessary step toward a better understanding of the governance challenges associated with emerging ecosystems.

Keywords: Data Governance, Open Government Data, Emerging Ecosystems, Case Study

1. Introduction

Scholars have found that open government data (OGD) encourages innovation [1][2]. Openness is important for generating value from data, and OGD is a crucial step toward sustainable innovative initiatives [3]. Many researchers believe that open data can revolutionize the creation of new services and products in both the public and private sectors [2]. However, few scholars have investigated OGD from the ecosystem governance perspective and how governance arrangements influence stakeholders in sharing and using OGD, thus affecting the potential creation of new services and products [2][4]. By considering theories on ecosystems [1][5] and governance [6], this paper explores how the loose coupling between stakeholders in an ecosystem creates governance challenges when OGD is utilized. Empirically, we focus on an emerging OGD ecosystem that encompasses both public sector

and private sector organizations in the Norwegian financial sector.

The idea behind OGD is that public sector data should be available in a convenient (ideally machine-readable) form as well as freely accessible, reusable, and redistributable by everyone [7]. Due to the large quantities of data possessed by governments, actors in the public sector can provide more data than private sector actors [8], who can provide open data but are often in possession of lesser quantities of data [2]. The availability of OGD is important because it facilitates the linking and reuse of data for the creation of innovative data-driven solutions [9].

Open data and its impact on the public have been an area of interest in research on information systems (IS) and eGovernment, which has broadly focused on OGD-driven co-created public services [10]. However, few scholars have investigated actual cases of data-driven value-creation using OGD and whether such cases have been successful or not [2]. In addition, most research has focused on the supply side of OGD—that is, the technical infrastructures necessary to generate and share data [11][12]. Sustained innovation based on OGD requires an ecosystem of public and private organizations that participate in value creation [1][13]. As an ecosystem by definition consists of a loosely coupled set of autonomous stakeholders [14], a well-functioning ecosystem requires governance structures and mechanisms that take this distributed nature of the ecosystem's stakeholders into account. More research is needed to understand the couplings between diverse stakeholders in an ecosystem and the governance challenges created by the stakeholders. To address this research gap and the scarcity of empirical studies on OGD ecosystems, we posed the following research question (RQ): What data governance challenges emerge in the relations between ecosystem stakeholders?

By investigating this research question, we sought to provide two key contributions. First, by examining and analyzing the empirical cases that leverage OGD, it is possible to find out how OGD affects the ongoing digital transformation and, therefore, how OGD is used.

Consequently, both scholars' and practitioners' perspectives can be broadened in terms of what governance domains are considered relevant when using OGD. The second contribution is a better understanding of the data governance process and how it affects the stakeholders in the studied ecosystem. Data governance has been discussed in existing literature [10][8], but we believe that more knowledge on data governance is needed. First, existing literature focuses on single-level ecosystems (e.g., direct transactions, such as in services) while overlooking other attributes, such as stakeholders who are not related to the actual transaction but who still influence the ecosystem [14]. Second, a study [15] showed that fewer than 15% of the ecosystems studied were sustainable, meaning that they did not gain a significant market position and were not financially sustainable. This, as well as the need for more real-life cases [2], means that more research is needed on the use of OGD in ecosystems in which data governance challenges are created by loosely coupled stakeholders.

In the following section, we position our research paper in the ecosystem literature and the data governance literature. Then, we discuss our methodology. Afterward, we present our case analysis to provide a broad view of the stakeholders studied in the emerging ecosystem and to show how governance challenges emerges. Finally, we discuss our contributions and future research avenues.

2. DATA GOVERNANCE IN ECOSYSTEMS

This article draws on two distinct but interrelated literature streams: ecosystems and data governance. The two literature streams are both highly relevant to our interest in researching data governance in an ecosystem. This interest stems from the lack of economic and social benefits achieved by open data initiatives [1]. There are several reasons for this, such as poor-quality data [16], mismatches between the needed and published data [17], and technical barriers [18].

An ecosystem is characterized by multiple autonomous stakeholders who are interdependent in terms of survival and performance and can be defined as "an economic community founded by organizations and individuals that includes customers, producers, competitors and other stakeholders" [14] [19] [20]. Such an ecosystem is important but poorly researched, as ecosystems in general are a complex topic [5], and scholars have focused on single-level rather than multi-level ecosystems, despite the innovation potential of ecosystems managed at multiple levels [14].

The stakeholders provide different value propositions, and the process for creating, capturing, and delivering value needs to be well orchestrated [19].

The ecosystem's structure reflects the interactions formed by complex networks [21]. From a business perspective, data ecosystems are a good means of explaining the economic and social benefits of open data initiatives [5]. A data ecosystem can be used as a tool for decision-making and planning, for localizing the relative positions of stakeholders in the ecosystem, and for facilitating access to the sharing and use of data.

Data governance in an ecosystem context.

Governance refers to comprehensive control that includes processes, policies, and structures. There are multiple types of governance, such as IT, information, data, and platform and social network service governance [22]. IT governance often addresses information and data governance, and information governance includes data governance [6]. Research on data governance has focused on "who holds the decision rights and is held accountable for an organization's decision-making about its data assets" [23]. In other words, existing research on data governance has predominantly investigated the organizational level.

When it comes to platform ecosystems, there is a consensus that roles, revenue sharing, trust, and control are key governance concepts [24][25]. These governance concepts should be implemented in data governance initiatives for platform ecosystems to encourage desirable behavior of all participating groups and to create value via the use of digital data [26]. Relevant issues and challenges should also be considered, such as unclear data ownership [27], the importance of the user contribution model [28], and invisible data use [22]. However, data governance has received little attention in platform ecosystem research [24].

For a platform to support innovations, platform architecture and its degree of openness are vital factors [29]. Openness, and thus the possibilities for future innovations, is governed by design rules [30]. The governance of these design rules is being implemented by platform authorities, who are responsible for managing both the architecture and the broader ecosystems of developers [29].

Focusing on the organizational level, Khatri and Brown [23] introduced a generic framework for supporting the design of data governance models, which includes the following five interrelated decision domains: data principles, data quality, metadata, data access, and data lifecycle. The framework provides a basic source of practices for data governance in an organization and highlights the importance of data governance from a practical standpoint in terms of what domain decisions are needed and which actor is responsible for each domain [23]. While this framework provides valuable insights, it does not focus on

governance issues in a platform ecosystem and thus does not contribute to research on data governance in the context of platform ecosystems [22].

To summarize, the change in the complexity of platform ecosystems and the lack of research on data governance in such ecosystems provide the basis for our analysis and aligns with our research question, hence highlighting the governance challenges emerging in relation with the ecosystem stakeholders.

3. METHOD

To explore how loose coupling between stakeholders in an ecosystem creates governance challenges, we chose an embedded case study as our approach. An embedded case study is a case study that contains more than one subunit of analysis [31]. We chose cases from the financial sector in Norway because this sector is mature in terms of how and to what extent data are used in real-life situations, which complies with the need for IS research [2]. We selected the following three cases as our subunits: a public organization (PublicOrg), a private start-up (FinStartUp), and a semi-private company (MiddleOrg).

We chose the PublicOrg case due to its role as a platform and a producer of OGD to other private stakeholders as well as the heavy load of technical debt PublicOrg bear that needs to be accounted for. PublicOrg is a front-runner in terms of OGD-driven initiatives in Norway and a role model for other public organizations in terms of creating and using OGD for value creation in both the public and private sectors.

We chose the FinStartUp case due to the company's role as a platform that provides innovative solutions while working with OGD as its main source of data without having bilateral agreements. FinStartUp is a small-sized company and is thus an agile stakeholder with minimal technical debt due to being born digital. They have created a Business to Business (B2B) platform that collects and purchases OGD, thus their role as a customer and a producer, to analyze them and provide their customers with insight into how other companies are doing financially and organizationally.

We chose the MiddleOrg case to understand the ecosystem from a multi-level perspective. The company works directly with PublicOrg and indirectly with FinStartUp. As MiddleOrg is a semi-private company, it is funded by stakeholders from both the public and the private sector. MiddleOrg's mission is to be the finance sector's infrastructure company, represent the banking sector and manage and provide the requests from the banking sector to PublicOrg. This kind of mediation is necessary to avoid any direct relationship between

private and public companies, thus eliminating the possibility of an individual private company being favored, which is against Norwegian law. The companies represented by MiddleOrg can be customers of FinStartUp, which is why the two organizations share digital data. We selected these cases due to their potential to provide rich insights into our RQ [32]. The structure of the studied ecosystem is presented in Figure 1.

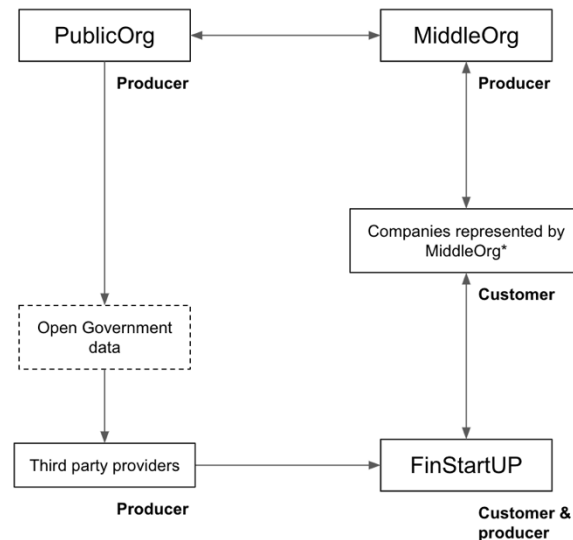


Figure 1. The ecosystem

The subunits PublicOrg and FinStartUp did not have an official agreement, and their relationship was not straightforward, as shown in Figure 1. The two organizations are connected through OGD, which is collected from the PublicOrg by third-party private providers, and through internal data from the companies that FinStartUp has a customer relationship with, represented by the subunit MiddleOrg. Competitors is essential in the ecosystem but direct competitors has not been focused on here. Knowledge of the complexity of the data ecosystem is essential for understanding how loose coupling between stakeholders in ecosystems creates data governance challenges.

3.1 DATA COLLECTION

We collected various kinds of evidence from our three subunits over a period of five months (December 2020–April 2021). Our approach combined semi-structured in-depth interviews with informants working within different domains from each of the three subunits, documents found online and received from the informants (news articles, presentations, and slides), and a demo presentation of the B2B platform created by

FinStartUp. Each interview lasted between 45 and 75 minutes, and each participant was asked questions from the following three main categories: working with digital data, the value of data, and collaborations in terms of data. These three categories was chosen as we sought to understand the emerging ecosystem (collaboration), how data played a role in their work (working with) and why (value) they worked with data. The question asked where designed by analyzing the literature with a focus on understudied fields, thus the emergence of the three categories, iterate using one test interview before finalizing the interview-guide. Table 1 summarizes our data sources.

	PublicOrg	FinStartUp	MiddleOrg
D	8 interviews	12 interviews	1 interview
a	<i>1 Scrum master</i>	<i>1 customer success manager</i>	<i>1 Program manager</i>
t			
a	<i>1 User experience researcher</i>	<i>1 Co-founder and CTO</i>	Documents
c			<i>Information online</i>
o	<i>1 Business developer</i>	<i>1 Co-founder and CPO</i>	
l	<i>1 Product manager</i>	<i>1 Co-founder and CEO</i>	
e			
c	<i>1 Senior advisor</i>	<i>1 Sales development representative</i>	
t			
e	<i>1 Information architect</i>	<i>1 Commercial manager</i>	
d	<i>1 Senior tax lawyer</i>	<i>1 Account executive</i>	
	<i>1 IT enterprise architect</i>	<i>1 AI trainer</i>	
	Documents	<i>4 Software engineers</i>	
	<i>Information online</i>	Documents	
	<i>Internal documents</i>	<i>Information online</i>	
		<i>Internal documents</i>	
		Demo presentation	
		<i>How the platform works and its key benefits for customers when it's used for auditing.</i>	

Table 1. Overview Empirical Data

3.2 DATA ANALYSIS

Each interview was transcribed and analyzed following Pan and Tan [33], with the data being collected, coded, and validated with the interviewees as well as using additional information, such as documents. To avoid validity issues, we contacted several informants to clarify the topics. Due to time constraints, only one person from MiddleOrg participated. To ensure the correctness of the statements and reduce validity issues, we triangulated the data to check the credibility of the interview statements [34].

Our analysis revealed a pattern, with four different data governance topics standing out. First, the maturity aspect was highlighted in several interviews from both PublicOrg and FinStartUp. Second, the goal of sharing data and thus the value of the data were described differently by interviewees from the three subunits. Third, data access was an important topic. Fourth, the subunits faced the same challenges regarding the quality of the data, but the means to tackle this issue differed between the public and the private sector. We will highlight the findings from the interviews later on.

4. FINDINGS

Data maturity. PublicOrg stated that, in general, they were mature in terms of seeing the possibilities for innovation of the data that they possessed but that some sections of the organization were lagging behind. The reason mentioned was the trouble in understanding the importance and value of data. PublicOrg stated that this was especially relevant for employees working with the legal aspects of data and data sharing. If changes in the data sharing process are relevant for the legal department, employees from the legal department can be part of the actual processes and change the direction, thus maturing accordingly.

“When the legal department is a vital part of the process, then they understand it. It is a manifestation of what they have in the documents and in their heads, which also makes them understand it.” (Information architect, PublicOrg)

According to the senior advisor in PublicOrg, if the legal department was not part of the plan, they would not see the value of data, and the gap between the legal department and, for example, the business developers would increase. The program manager also added that the differences between the employees working on the business side and the employees working on the legal side stopped many projects due to legal challenges. These challenges include differences in policy and practice due to the issue with data maturity, and were encountered by the government entity in charge of

PublicOrg, which, as described by the senior tax lawyer, created further issues:

“We don’t have a say in the laws being made so we can do our job, and the people in charge are from another department where they hire people who are good at law but don’t have the expertise in data and technology.” (Senior tax lawyer, PublicOrg)

For PublicOrg, expertise in data and technology and the potential value generated by the data were important. However, as mentioned by the IT enterprise architect, the focus from PublicOrg was more on different applications and IT systems. FinStartUp stored their data in the cloud and had no legacy systems, which made the organization more mature. The CTO and co-founder of FinStartUp explained that FinStartUp being a small stakeholder with no legacy led to a more robust organization in terms of changes. This meant that it was easier to change the direction if the organization, for example, changed its data model. FinStartUp’s CTO and co-founder compared FinStartUp to PublicOrg, stating that as the latter was a public organization, it needed to maintain everything from the old systems, which made them slower than FinStartUp. FinStartUp’s advantage resulted from the use of new technologies as well as the organizational culture. FinStartUp was mature because everybody was working in a hands-on manner with data and learned to make their own decisions.

“We want everybody to work independently, meaning that they take control of their own work and that everybody understands the importance of making their own decision. Our motto is ‘You are FinStartUp, make decisions.’” (CEO and co-founder, FinStartUp)

Goals. For PublicOrg, the main goal of working with and sharing digital data was not value creation. Rather, the important part was those different stakeholders, both from the public and the private sector, were getting value from the data. If socio-economic values are created, this raises the importance of sharing the data.

“If there are value to be harnessed from a socio-economic perspective, we will argue that the data should be shared, and if the sharing of the data can lead to more data being collected, that is positive.” (Senior advisor, PublicOrg)

This statement was shared by the the program manager from MiddleOrg, who stated that the companies represented by MiddleOrg could access the public data about their customers through PublicOrg rather than asking their customers if they could share the digital data. This was considered to be greatly valuable due to the time and resources saved and due to the centralization of information.

FinStartUp’s goal was to ease the work of their customers by eliminating manual tasks that the customers of FinStartUp encounters when auditing companies with multiple datasets. This was the main value of FinStartUp’s platform, and the organization’s purpose was to make their customers’ work processes more effective, as stated by FinStartUp’s CEO and co-founder. Shared digital data provided a better picture of the world and its connections for their customers. More specifically, the value resulted from combining internal and external data, which was also the source of new value, as stated by the CTO and co-founder of FinStartUp. The combination of data is also the main value for FinStartUp as they collect the public digital data that is legally available but can be hard to access.

“Public data is important because you can’t do much with only internal data.” (CTO and co-founder, FinStartUp)

The CPO and co-founder of FinStartUp claimed that in the EU, there is an estimated yearly sum of 1,000 billion NOK used by private and public organizations for anti-laundry checks and that FinStartUp could halve that sum using its platform. Besides creating monetary value, having a use case for actually using the data was also mentioned as an important purpose for FinStartUp.

“Use cases that actually use the data for something real is the real gold.” (CPO and co-founder, FinStartUp)

Data access. PublicOrg was eager to share the collected data with both the public and the private sector, but, as stated by the senior tax lawyer, they could not treat specific private actors with exclusivity due to Norwegian law. To avoid this, a partnership with a provider that represents a segment in an industry, like banking, was established, with MiddleOrg being the provider. This way of providing access to the open government data came at a cost, which involved knowing what to share and possessing the required technical know-how, as stated by both PublicOrg and MiddleOrg.

“In the beginning, we did not have a specific architecture or agreement on what data the companies needed, so the public and the private pulled in two different directions.” (Project manager, MiddleOrg)

To fix this issue, PublicOrg, alongside other public organizations, took charge in leading the project in terms of what could be shared and how it should be shared. The IT enterprise architect stated that this was the right move, but it also created new barriers, as PublicOrg always needed to work on aligning what the private sector wanted and what the public sector could share. As stated by the senior tax lawyer, the private companies in the financial sector had to follow strict laws, which limited what they could access. To

understand what they can access the business developer stated that domain knowledge was important.

In the case of FinStartUp, most of the digital data needed by FinStartUp was bought from third parties, as public data are hard to access due to their structure, which makes gathering data time consuming, as stated by the CPO and co-founder. Using third party actors was also a part of FinStartUp's main strategy when accessing data, as buying already structured data from third parties involves fewer access points. FinStartUp also stated that it had three ways of accessing data. The first one was through an Application Programming Interface (API), in which case FinStartUp connected their platform directly. The second way was by scraping web pages via an algorithm that retrieves the needed information according to the structure set up by FinStartUp. The third one involved collecting files, such as XML files, from other parties, such as customers. These data sources often involved media data, as the latter has a good structure in a readable format. Another reason, as stated by the CPO and co-founder, was that due to previous work experience with media companies and other large companies, the CPO and co-founder knew the people, which made it possible to know whom to approach and when regarding data access.

Data quality. Data quality was important for PublicOrg due to the organization's collected data being shared with other entities, both public and private, through APIs or collaborations. During these collaborations, PublicOrg established standard contracts to ensure that all the legal requirements were in place; otherwise, data quality could have been affected due to sources being unavailable for sharing or use in specific use cases. These contracts also ensured that the semantics were in place, which was described by PublicOrg as a previous challenge when working with private companies, as different names were being used for the same object. PublicOrg aimed to solve this challenge by adopting the three principles that need to be followed when sharing data according to the New European Interoperability framework: that the semantics are in place, that there is no legal breaches such as GDPR, and the technical solutions can handle the work needed. With PublicOrg being a public organization, they handled masses of sensitive data that they could not share. These data could only be shared with other public entities if there is an agreement and if the entities were not private companies.

“We have data that is called code 6 and code 7, which we can't share due to legal reasons.” (Senior advisor, PublicOrg)

The data collected, handled, and shared could contain errors if the original data were reported incorrectly. For PublicOrg, the quality of the data increased over time due to a more nuanced picture

emerging as a result of digital data being reused. There was also a higher awareness that data quality was harder to achieve as society changed, but where Norway has an advantage, compared to other European countries, due to the amounts of digital data collected from citizens by public organizations here in Norway.

“It was easier before when people had one job for thirty years. It is harder now as people are changing jobs so often. But this is easier to track here in Norway, as we're actually pretty good at collecting data compared to other countries I've been in contact with.” (Business developer, PublicOrg)

In FinStartUp, establishing the semantics was a challenging process but the organization tackled this issue by analyzing data against domains. This was done by adding context to words so that sentences would be less likely to be misinterpreted. The adding of context was also done by triangulating the data to get a better overview of the domain for each piece of text, as stated by FinStartUp. This method is known as the *relevance model* and was created to adjust the data quality. Even though FinStartUp were working to improve data quality through interpretation, the organization still believed that it had a long way to go to reach perfection.

“I don't quite feel that we have succeeded in interpreting the data so that we can allow the machine to do everything.” (AI trainer, FinStartUp)

The relevance model helps to improve the data quality and is based on domains or categories set by the data provider, which affects the quality of the data on the FinStartUp platform. These categories can be changed without warning, which then affects the data quality, as it becomes harder to make sense of the collected data. FinStartUp tackled this problem by implementing a surveillance system to alert the organization whenever changes were made, but it still needed to fix the problems that could occur. If the data were not adjusted before reaching the platform, the users of the FinStartUp platform could vote whether something was relevant or not, which increased the data quality due to better accuracy for the end user. These fixes mainly had to do with semantic challenges but could also be used for other barriers that FinStartUp ran into. As the finance sector is an international sector, FinStartUp relied on both national and international data, with language being a barrier. This was described by the AI trainer as a minor problem that is being solved as more data are being collected and analyzed.

5. ANALYSIS

In this section, we analyze our findings to explain how loose coupling between stakeholders in an ecosystem leads to data governance challenges. Also,

we elaborate on the four topics presented in the Findings section and discuss their importance for the governance perspective in relation to the three subcases.

5.1. THE FOUR CATEGORIES

Data maturity. Data maturity refers to the knowledge on how data can be used and the possibilities that the data have for both the organization and its employees. As discussed in the Findings section, data maturity was more important for PublicOrg than FinStartUp. As stated by the CTO and co-Founder of FinStartUp, the differences in technical know-how can explain the differences in how pressing the maturity problem is. As PublicOrg was a public organization that had existed for many years and had many areas of interest, their lack of interest in the usage of digital data across the whole organization, was one reason why the whole organization was not aligned.

PublicOrg's legal department stood out due to its lack of data maturity. But if the legal department is accounted for as a vital part of the process of using digital data in more projects, a change in data maturity can happen. This inclusion of the legal department leads to two interesting points. First, being accounted for as vital, and thus trusted, is important in any work situation and leads to better task engagement. Second, the legal department had to obey external rules, as it worked to make sure that PublicOrg followed the requirements for data use and sharing according to Norwegian law. Trust and control are key governance concepts [24][25] and should be implemented in a platform ecosystem to encourage desirable behaviors [26]. As long as control and trust are lacking in PublicOrg, the desired data maturity will not be achieved.

Goals. A goal is something that an organization wants to achieve from a value perspective in terms of working with and sharing data. PublicOrg worked toward socio-economic goals, while FinStartUp was working to create monetary value. This is one of the main differences between all public and private sector organizations. However, whereas FinStartUp attempted to understand its customers in the best possible manner due to the commercial interest inherent to a private company, PublicOrg did not experience the same pressure. PublicOrg had to deliver results, but it was harder to quantify the socio-economic value that the organization provided.

In an ecosystem, each stakeholder works to cut time and make data more accessible. This is an important point of discussion, as better accessibility and availability can lead to the linking and reuse of data to create innovative data-driven solutions and/or products [9]. Accessibility can unify the goals of both the public

and the private sector and thus enhance the value proposition of PublicOrg, FinStartUp, and MiddleOrg.

Data access. Data access refers to the role that an organization working with data has and thus to what degree the organization succeeds in its role. The discussion on data access is complicated in the emerging ecosystem studied due to the number of stakeholders. As PublicOrg both gives and enables access to data, there is a clear distinction between PublicOrg and FinStartUp, as the latter analyze and collect data to present information to customers. This distinction provides insights into each organization's approach to data access.

PublicOrg needs to clearly take the leadership role in a data-sharing relationship. This is due to the organization's role as a provider of OGD, and thus setting the criteria for the structure between the stakeholders in the emerging ecosystem studied. It is important to highlight this structure as it allows FinStartUp to access public data through third-party providers. The barrier of not well-structured digital data is not only a governance challenge but also a business opportunity for third-party providers.

Data quality. Data quality refers to the challenges that make it difficult to use digital data and to how these challenges are handled. When sharing data, the data quality is an important aspect, and semantics was something that affected both PublicOrg and FinStartUp. Regarding the subcases, issues to do with semantics were something that they were trying to tackle. FinStartUp used technology to decrease the chances of error and to increase the data quality, while PublicOrg was more traditional in the sense that it used legal governance in the form of contracts. This is an interesting difference between the two organizations, which may be related to maturity and how they work with partners. The difference may also be affected by how close the relationships are, as PublicOrg has projects with many parties in which data are shared, while FinStartUp has more customer relationships.

FinStartUp mentioned language as one aspect that could influence data quality. This issue was not mentioned by any of the PublicOrg informants, which may be due to the fact that PublicOrg is a Norwegian public organization working with national data. However, it can also mean that PublicOrg does less work that involves cross-country data sharing, which means that the data quality will not improve, in an international perspective, due to less amounts of data.

6. DISCUSSION

Our RQ (What data governance challenges emerge in the relations between ecosystem stakeholders?)

examines different stakeholders in an ecosystem as well as data governance challenges. In our research, we found that in the studied emerging ecosystem, there were stakeholders from both the public and the private sector with multiple roles, such as customers, producers, and competitors. Based on their sectors, these stakeholders worked toward different goals while dealing with similar governance challenges in relation to data access, data quality, and data maturity. By answering the RQ, we have contributed to the literature on both OGD and platform governance. Furthermore, the observations from our analysis also have practical implications for stakeholders, depending on whether the stakeholders are public or private organizations within an ecosystem, as the knowledge we provided can be used directly by both established and new entities in similar ecosystems.

We have contributed to the literature on OGD by addressing the reasons for the lack of economic and social benefits achieved by open data initiatives [1]. In terms of reasons, the existing literature has pointed out poor-quality data [16], mismatches between the needed and published data [17], and technical barriers [18], which we have elaborated on using the findings from our studied cases. As FinStartUp needed to create technical tools to understand the semantics due to mismatches between the received and needed data, the emerging ecosystem did not appear to be well orchestrated. The literature states that for stakeholders to provide different value propositions, such orchestration needs to be present [19], which contradicts our findings regarding the creation of economic and social benefits. This is highlighted in FinStartUp's value creation process, as they managed to create actual value for their customers. This value consisted of digital data, which was analyzed to create information used by FinStartUp's customers to gain new insight into potential new customers. This process of value creation extends the existing open government literature by discussing new economic benefits that have been achieved in terms of open data.

In addition to the open government literature, we have contributed to the platform governance literature by addressing how stakeholders in an ecosystem addresses governance challenges. The literature states that a platform ecosystem's openness is governed by design rules [30] and that these design rules are created by platform authorities, which are responsible for managing both the architecture and the broader ecosystems of developers [29]. According to our findings, PublicOrg was responsible for creating these rules, which strengthens the literature and creates challenges for PublicOrg due to unclear data ownership [27]. As FinStartUp used third-party providers for the data, we argue that this extends the literature by showing

who manages the broader ecosystems of developers and thus contribute to the creation of new challenges. The third-party providers contribute to this challenge by putting a price on their data due to their structuring work, even though the data are open. These findings extend existing literature which highlight openness as an important factor for innovations [29], where the findings show that openness also can create new business models, which is the case for the third-party providers as they exploit the open government data.

7. CONCLUSION

We sought to understand what data governance challenges that emerge in the relations between ecosystem stakeholders. We found that governance topics differ in terms of relevance among stakeholders in an ecosystem due to the stakeholders' different perspectives on governance topics and their different roles in the ecosystem. In other words, we found a complexity issue in the studied emerging ecosystem, as there were many stakeholders, both related and unrelated to one another, who affected the different governance topics. New stakeholders also increased complexity as boundaries needed to be set to cope with the governance challenges.

Future studies should investigate the dynamics between the different stakeholders from a more diverse perspective. More research on each stakeholder involved is needed, both from the individual perspective and from the ecosystem perspective. This would help understand the governance challenges created by the loose coupling between the stakeholders within the ecosystem and would provide new insights into how governance challenges are created and by whom, as opposed to the organizational-level framework proposed by Khatri and Brown [23]. We have shown the importance of understanding potential stakeholders in ecosystems and the governance uniqueness in and for different roles. We suggest that different roles need unique forms of governance in emerging ecosystems.

8. References

- [1] C. Bonina and B. Eaton, "Cultivating Open Government Data Platform Ecosystems through Governance: Lessons from Buenos Aires, Mexico City and Montevideo", *Government Information Quarterly*, 2020.
- [2] M. Olesk, K. McBride, T. Kalvet, and R. Krimmer, "Open Data as Enabler of Public Service Co-Creation: Exploring the Drivers and Barriers", *EGOV-EPART Conference*, 2017.

- [3] T. Jetzek, M. Avital, and N. Bjorn-Andersen, “Generating Value from Open Government Data”, *International Conference on Information Systems (ICIS 2013): Reshaping Society Through Information Systems Design*, 2013.
- [4] M. Janssen, Y. Charalabidis, and A. Zuiderwijk, “Benefits, Adoption Barriers and Myths of Open Data and Open Government”, *Information Systems Management*, 2012, pp. 258–268.
- [5] C. Azkan, F. Möller, L. Meisel, and B. Otto, “Service Dominant Logic Perspective on Data Ecosystems – A Case Study Based Morphology”, *In Proceeding of the 28th European Conference on Information Systems*, 2020.
- [6] F. Faria, A.C. Maçada, and K. Kumar, “Information Governance in the Banking Industry”, *Proceedings of the Annual Hawaii International Conference on System Sciences*, 2013, pp. 4436–4445.
- [7] D. Dietrich, *The Open Data Handbook*, 2009.
- [8] V. Bianco, V. Myllärniemi, M. Raatikainen, and M. Komssi, “The Role of Platform Boundary Resources in Software Ecosystems: A Case Study”, *Working IEEE/IFIP Conference on Software Architecture*, 2014.
- [9] European Commission, *Open Data: An Engine for Innovation, Growth and Transparent Governance*, 2011.
- [10] A. Zuiderwijk, M. Janssen, S. Choenni, R. Meijer, and R.S. Alibaks, “Socio-Technical Impediments of Open Data”, *Electronic Journal of e-Government*, 2012, pp. 156–172.
- [11] L. Danneels, S. Viaene, and J. Bergh “Open Data Platforms: Discussing Alternative Knowledge Epistemologies”, *Government Information Quarterly*, 2017.
- [12] G. Maccani, B. Donnellan, and M. Helfert, “Exploring the Factors that Influence the Diffusion of Open Data for New Service Development: An Interpretive Case Study”, *European Conference on Information Systems*, 2015 Completed Research Papers. Paper 127.
- [13] R. Medaglia, J. Hedman, and B. Eaton, “It Takes Two to Tango: Power Dependence in the Governance of Public-Private E-Government Infrastructures”, *ICIS 2017 Proceedings of the International Conference on Information Systems*, 2017.
- [14] P. Wang, “Connecting the Parts with the Whole: Toward an Information Ecology Theory of Digital Innovation Ecosystems”, *MIS Quarterly*, 2021, pp. 397–422.
- [15] R. Martin, L. Hen, L. Julien, and G.J. Michael, “How Business Ecosystems Rise (and Often Fall)”, *MIT Sloan Management Review*, 2019.
- [16] A. Vetro, L. Canova, M. Torchiano, C. Minotas, R. Iemma, and F. Morando, “Open Data Quality Measurement Framework: Definition and Application to Open Government Data”, *Government Information Quarterly*, 2016.
- [17] J. Gurin, C. Bonina, and S. Verhulst, “Open Data Stakeholders – Private Sector”, *The State of Open Data: Histories and Horizons. African Minds and the International Development Research Centre (IDRC)*, Cape Town, 2019, pp. 418–429.
- [18] P. Conradie and S. Choenni, “On the Barriers for Local Government Releasing Open Data”, *Government Information Quarterly*, 2014.
- [19] R.C. Basole, M.G. Russell, J. Huhtamaki, N. Rubens, K. Still, and H. Park, “Understanding Business Ecosystem Dynamics: A Data-Driven Approach”, *ACM Transactions on Management Information Systems*, 2015.
- [20] J.F. Moore, *The Death of Competition: Leadership and Strategy in the Age of Business Ecosystems*, Harper Business, New York, 1996.
- [21] D. Green and S. Sadedin, “Interactions Matter – Complexity in Landscapes and Ecosystems”, *Ecological Complexity*, 2005, pp. 117–130.
- [22] S. Lee, L. Zhu, and R. Jeffery, “Data Governance for Platform Ecosystems: Critical Factors and the State of Practice,” *Proceedings of the twenty First pacific Asia Conference on Information Systems*, Langkawi, 2017.
- [23] V. Khatri and C.V. Brown, “Designing Data Governance”, *ACM*, 2010, pp. 148–152.
- [24] M. Schreieck, M. Wiesche, and H. Krcmar, “Design and Governance of Platform Ecosystems – Key Concepts and Issues for Future Research”, *Proceeding of the Twenty-Fourth European Conference on Information Systems*, 2016.
- [25] A. Tiwana, B. Konsynski, and A. Bush, “Research Commentary—Platform Evolution: Coevolution of Platform Architecture, Governance, and Environmental Dynamics”, *Information Systems Research*, 2010, pp. 675–687.
- [26] K. Weber, B. Otto, and H. Osterle, “One Size Does not Fit All—A Contingency Approach to Data Governance”, *ACM Journal of Data Information Quality*, 2009.
- [27] S. Kaisler, F. Armour, J. Espinosa, and W. Money, “Big Data: Issues and Challenges Moving Forward”, *Proceedings of the Annual Hawaii International Conference on System Sciences*, 2013, pp. 995–1004.
- [28] G. Parker and M.W. Van Alstyne, “Platform Strategy”, *The Palgrave Encyclopedia of Strategic Management*, 2014.
- [29] C. Bonina, B. Eaton, and S. Henningsson, “Governing Open Data Platforms to Cultivate Innovation Ecosystems: The Case of the Government of Buenos Aires”, *Proceedings of the 39th International Conference on Information Systems (ICIS) Association for Information Systems*, 2018.
- [30] C. Baldwin and K. Clark, “Design Rules: The Power of Modularity”, *The Academy of Management Review*, 2000, pp. 130–133.
- [31] R. K. Yin, *Case Study Research: Design and Methods*, Thousand Oaks, CA, Sage Publications, 2003.
- [32] B. Flyvbjerg, “Five Misunderstandings about Case-Study Research”, *Qualitative Inquiry*, 2006, pp. 219–245.
- [33] S.L. Pan and B. Tan, “Demystifying Case Research: A Structured-Pragmatic-Situational (SPS) Approach

- to Conducting Case Studies”, *Information and Organization*, 2011, pp. 161–176.
- [34] R. Edwards and J. Holland, *What Is Qualitative Interviewing?*, Bloomsbury, London, 2013.