TALLINN UNIVERSITY OF TECHNOLOGY

School of Business and Governance

Ragnar Nurkse Department of Innovation and Governance

Sandra Metsis

CROSS-SECTORAL DATA COLLABORATION PLATFORMS

Master's thesis

Supervisor: Veiko Lember, PhD

Tallinn 2020

I hereby declare that I have compiled the thesis independently and all works, important standpoints and data by other authors have been properly referenced and the same paper has not been previously presented for grading.

The document length is 12109 words from the introduction to the end of conclusion.

Sandra Metsis ……………………………
                    (signature, date)
Student code: 122928HAGMM
Student e-mail address: sandra.metsis@eesti.ee

Supervisor: Veiko Lember, PhD:
The paper conforms to requirements in force

……………………………………………
(signature, date)

Chairman of the Defense Committee:
Permitted to the defense
…………………………………
(name, signature, date)

TABLE OF CONTENTS

ABSTRACT

The potential value of data has been an intriguing topic for decades, digital platforms have made their way to redefine the collaborative relationships and the importance of collaboration between sectors to provide value to the wider public cannot be underestimated. Cross-sectoral data collaboration platforms might have the potential to present a dynamic and comprehensive collaboration framework that enables realizing the data for the public good. Yet, there is a lack of knowledge on how to organize cross-sectoral platforms that strive to create public value from data. This thesis aims to fill the gap by creating a theoretical framework consisting of elements of collaboration, platform and data value creation. The framework will be validated through exploring success-factors of the real-life collaborative data platforms with different scopes. The research results in presenting the factors enabling data collaboration platforms development and the public value creation, which could ease the strategic planning of future cross-sectoral data collaboration initiatives and realizing the potential of data.

Keywords: cross-sectoral collaboration, data exchange, public value creation, platforms

INTRODUCTION

The potential and underutilization of data has been a burning discussion for more than a decade. In 2011, McKinsey Institute issued a report that stated that big data will become the basis of competition, productivity growth and enable new waves of innovation – as long as there are right policies and enablers in place (McKinsey Global Institute, 2011) and estimated that **connecting data across the industries and geographic boundaries could create roughly $3 trillion annual economic value by 2020** (McKinsey, 2019). But even if acknowledging the potential, most industries have still not come close to realizing the full potential of it.

Most of the excitement around the data in the business world focusses around its potential commercial value. Six of the most valuable private companies in the world rely on data (Statista, 2019). The value creation from data is an actual topic also in the public sector, but as there is no single approach for defining "public value" in the application of data (OECD, 2019), exploitation showcasing concrete value is challenging.

Public sector itself holds huge amounts of data and is an enabling the public value for example by providing access to Open Government Data (OGD). But data as an intangible asset does not provide value automatically when made available - even in this context the value is co-created with different stakeholder's collaboration (Bolivar, et al., 2019) and presented in co-creation of services and solutions.

The real value lies further than making the governmental data available, it lies in making the actual use of the data and involving the available resources across the sectors. **The private-sector input on providing data itself and expertise to benefit from it has been widely recognized as an essential element to solve global problems** and for example 2030 Sustainable Development Goals (United Nations, 2014). There is a need for going further from the one-time transaction-based data exchanges, and one-way data provision such as different "Open Data" portals and bring in more dynamic tools to enable to the collaboration.

To realize the potential of collaboration and data available across the sectors, several challenges need to be overcome. Starting from technical – developing enabling technical solutions, legal and standardization framework, educating skilled personnel; and ending with

building trust and good governance models. While challenges are growing in complexity, the technological change in public administration literature remains under-researched (Andrews, 2019).

The concept of Digital Era Governance (DEG), brought to life by Dunleavy, et al.(2008) has been one of the new paradigms looking into how the public sector should be reorganized in the digitalization era, bringing three central topics: reintegration, needs-based holism and digitization changes on the table. DEG answers the numerous calls (see for example (World Economic Forum, 2015) that emphasized driving towards networks and platform based social and economic models, and government based on **digital platforms that will redefine the relationship between government and the people**.

The development of public-private-academic **partnerships and platforms is seen as one of the critical success factors in exploiting the benefits of the Data Economy** (The World Bank Group, 2017), where partnerships practices are covering both - infrastructure and nontechnical aspects, including policy assessments and implications, public perception and awareness, data stewardship, financial models, business value propositions, competency, skill requirements, etc (*Ibid*).

Even though the long run trend of government increasing responsibility for socio-economic activities has started to reverse from the 1970s (Alford & O'Flynn, 2012, pg. 25) - it has not meant giving back the roles to private sector, but has led to more complex role division and coordination. Externalization has worked well in only certain circumstances. Challenging situations demand **new forms of governance**. This brings in **complicated processes in multi-stakeholder decision-making** (de Bruijn, 2018) that needs to mastered in the networks and often find their output on different digital platforms.

Public service platforms are a notable part of the public service digitalization, as the public sector is the world's largest service provider (PwC, 2007). Numerous public initiatives have been looking for a ground of innovation, collaboration, and ecosystem development in the area of platforms. Platforms that enable collaboration and developing innovative solutions to problems and to scale innovation are commonly spread (Daglio, et al., 2014.) Yet, **there is lack of information on how to organize cross-sectoral platforms.**

Platforms are nothing new – they are seen as an output for new forms of governance and digitalization. Literature about the platforms has grown significantly over the last decade. Still, many questions remain unanswered, and lack of agreement on definition and conceptualization – technical and non-technical – is challenging the creation of common understanding. The complexity of the phenomena offers many more sides to explore.

During the recent years, several policy documents (for example OECD, 2019 and European Commission, 2020) emphasize the need for cross-sectoral collaboration in exploiting data and finding better models to solve the challenges in society and deliver efficient public services, but the research regarding cross-sectoral data collaboration is still rather scarce.

Platforms as data collaboration model has been reflected in the research mostly as just one example of possibilities of presenting the outcomes, not providing deeper insights how to exploit the model for better collaboration and value creation out of data. Wider theory on the platforms, both as collaboration and infrastructure framework, could offer various insights that could be beneficial also realizing the data collaboration efforts and offer a good governance framework for stakeholders and data.

This work will be focusing on **cross-sectoral data collaboration platforms as a central research object,** exploring the triggers of this specific form of data collaboration and looking into how platforms enable the collaborative value creation from data. Hence, the following research question is formed:

What are the main enabling factors of cross-sectoral data collaboration platforms development and public value creation?

Existing literature of the platform concept, cross-sector data collaboration, value creation from data and enablers of the collaboration is gathered into a framework. A case study with three cases is developed to validate the model. The selected platforms are representing a variety of examples with different scopes: (1) regional data marketplace, (2) national sectoral data service platform and (3) global platform harnessing and sharing data for sectoral developments.

As a result, the framework of enablers of platform as a collaboration model for value creation in the cross-sectoral data collaboration is validated for better exploitation of the data trough collaboration platforms and potentially triggering additional cross-sectoral initiatives.

After the introductory chapter, Chapter 1 is providing a State of Art – overview of the rise of the importance of data, value creation and cross-sectoral collaboration in the context of data. The chapter gives an understanding of the essence of the platforms as a collaboration enabler and looks into the dynamics of the data collaboration governance and public value creation. Chapter 2 provides an insight of the research approach and cases selection. Cases are introduced in Chapter 3, cross-analysed based on the research questions in Chapter 4 and the research will concluded in the last chapter.

1. STATE OF THE ART

1.1. The rise and relevance of data

In 2006 British mathematician Clive Humby compared data resource with oil (Walsh, 2018) – referring how its value comes after refining the raw material into a product. While the comparison gives a somewhat good understanding of the essence of data, the scope of production and the possibilities cannot be compared with this limited mineral resource.

Referring to the potential of the data economy has only raised since 2006. The decade after we were presented the first actual applications and value of data, while American tech corporations – Alphabet, Amazon, Apple, Facebook and Microsoft managed to realize huge profits exploiting data (The Economist, 2017). The critical essence of unregulated fields in privacy, ethics and many legal issues were brought into the spotlight.

Ever since the time and money have been invested in the capturing, management, processing and stewardship of data, it has been leading the strategies of many private and public institutions (OECD, 2019). Notable 256% growth in data science jobs over last 5 years has been taking place (Indeed Hiring Lab, 2019) and according to data professionals skills gap indicator, there will be over 1.5 million unfilled positions related to data in European Union (EU) 27 member states by 2025 (IDC Italia srl and The Lisbon Council, 2019).

The McKinsey Institute has **calculated potential efficiency gains at EUR 250 billion per year for the European public sector** (Manyika, et al., 2011) by **data-driven approach** and potential for the **private sector to raise the productivity on average around 5%** has been detected by McAfee and Brynjolfsson (2012).

Even though the history of automated **data-driven decision making** in the public sector reaches to 19th century (Agar, 2003), the realization of this process **in the digital area** has been led rather by the private sector. The ways to monetize data and drive value have been exploited there in bigger scales (Opher, 2016). In 2019, OECD published a working paper of "A Data-driven Public Sector" recognizing data as an important asset to policy-making (van Ooijen, et al., 2019) and promoting data-backed service design as well as embedding good governance values.

The oppression to data has brought two important notions on the table, that are constantly used to explain the potential of strategies produced by the public sector. First of those notions is **Big Data**. It has uncertain origins, few mentions already in 1990s, but it became widespread in 2011 (Gandomi & Haider, 2014). Since then, "The Three V's" – standing for *Volume, Velocity, Variety* (Manyika, et al., 2011) have become a common framework in defining Big Data. Those three are the characteristics of data, but do not give the effect on any scale if analytics power – sophisticated and computer-powered, is not added to them. Without the analytics capability, the gathered (big) data is useless.

Exploiting Big Data has stated (for example (Ketter, et al., 2015) to have a potential to mitigate wicked problems (Rittel, 1973) that exceed the capacity of individual research groups – climate change and feeding the growing population being some of the most actual examples. Yet, the ambiguous understanding of the term itself and the potential has not reflected in many concrete actions. Also, the information asymmetry between the governance, regulatory institutions and technology companies is affecting if the problem can be defined as wicked and the solution is to be found (Andrews, 2019). As in this work, the data is looked as a more general resource and does not necessarily have the elements of "The three V's", the author does not see any need to reflect on the term any further.

The other relevant notion in data and public value creation context is **Open Data.** Lead by US Open Data Platform data.gov in 2009, many countries have followed to initiate national Open Data portals. They all carry values like transparency and openness, but the actual impact has been rather limited (see for example Worthy (2015) or Wang (2016) – and measuring this (tracking the uses of data sets) is complex. One of the common assumptions for the reasoning is the simple demand-offer mismatch that the world full of data is facing, in addition to legal, ethical and technological barriers. Simply having nor even "opening" the data is not serving the public interest.

Some authors have looked into approaches, of how the usage of Open Data can be simulated and scaled up in meaningful ways, bringing actual benefits, not only being provided for transparency. Ruijer and Meijer (2019) evaluated the **Open Government Data (OGD)** as an Innovation Process, and conclude that the importance of commitment to managerial aspects – both dealing with the organization and technical aspects, should be taken more into focus. While Open (Government) Data is often part of the data collaboration platforms, this work is

looking further from exploiting this and into the datasets that are obtained from both – public and private sector, and the value created by the platform collaboration.

Data-enabled algorithmic approaches and decision-making propose another set of challenges and risks. Those should be assessed carefully by analysing different tools and collaboration forms to adapt. Andrews (2019) research brought out challenges like selection error, algorithmic law-breaking, manipulation, algorithmic propaganda, algorithmic brand contamination and unknowns. The public value creation is strongly connected with the public trust, ethics, and legal framework which brings in the complexity to data governance. The essence of data – being an intangible asset, adds another layer of complexity.

EU has done notable actions to regulate the data economy, the rights of individuals and created obligations to organizations. General Data Protection Regulation (GDPR (European Parliament, Council of the European Union, 2016) that admits the protection of the personal data as a fundamental right, is one of the most prominent examples. There is also an effort made to create ethical borders and standards – two of the burning questions in the data economy.

On the other hand, the continent does not have any notable data-driven companies, such as the previously mentioned US corporations. EU Data Strategy vision sees the future in enabling legislative framework for the governance of common European Data Spaces and exploring the need for legislative action on issues that affect relations between actors in the data-agile economy (European Commission, 2020). The increase in **competitiveness of EU data economy is expected to come from the collaboration with the private sector**. This has been getting increasing attention from the policy-makers but not many realizations.

1.2. Value creation from data

There is a common understanding that data harvested from the general public should provide public value. The origins of the concept of "public value"- the value an organisation contributes to society, takes us back to 1997, when Moore (1997) wanted to find an equivalence to the private sector's shareholder value within public management. Benington and Moore (2011) proposed a strategic triangle to understand public value creation consisting

of the following: development of a clear public purpose; management of the authorizing environment; and development of the relevant capacity.

Some theories view government bodies as actors competing against each other, captured by the interest groups and serving rather limited circle, not a public as a whole (Jarman & Luna-Reyes, 2016, p. 8). As a response to that Talbot (2011) brought out the importance of the perspective of the general public, meaning that the public value should not be declared by governments, but by citizens – what they understand to have gained via five dimensions: trust and legitimacy; collectivism; security; personal utility; autonomy (Talbot, 2008, p. 18).

According to Kelly and Muers (2002) **public value has three building blocks – services, outcomes and trust,** while services are the "vehicle" delivering the value. Jarman and Luna-Reyes (2016, p. 9) have proposed that public value is a concept tied to an approach – the process towards defining it is as important as the value itself. If and how this process is connected with initiating the data collaboration is an interesting question to look into exploring the cases in the following chapters. While some examples exist, the overall acknowledged framework is missing and the public value measurement remains rather a conceptual model (Bigoni, et al., 2018) not actually applied.

European Union sees that **data can serve the public interest** in improving the awareness, give a better understanding in causes and variables, provide prediction capability, run more rigorous impact assessments and guide decision-making (European Commission, 2020, p. 19). The Data Collaboratives concept (Verhults & Young, 2017) brings in the collaborative perspective and adds improving public service design and delivery and contributing to knowledge creation and transfer between the sectors.

Those are rather abstract means, with the aforementioned public interest, improvements can be complex to measure. Adding the fact that in its essence data is not a tangible asset, makes the value measuring even more complex process - neither the private nor public sectors have yet developed a definitive model for measuring the value of data (OECD, 2019). But even if measurable contribution of platforms to enabling public value can be challenging to assess, platforms contribution to enabling the process can be explored.

The data value creation process is not linear and iterative, but seen as cyclical – benefitting from evolution and learning (Cordey, et al., 2010). Ooijen, C., B. Ubaldi and B. Welby (2019) stated that the improved management of Government Data Value Cycle (Figure 1) can support the data-based decision-making and lead to policy improvements. This concept offers a valuable input for public value creation framework from data, but is limited in the context of this work, as platforms are seen having a wider role in enabling the data-collaboration than being an enabler in sharing and providing an interface.

Figure 1. Government data value circle
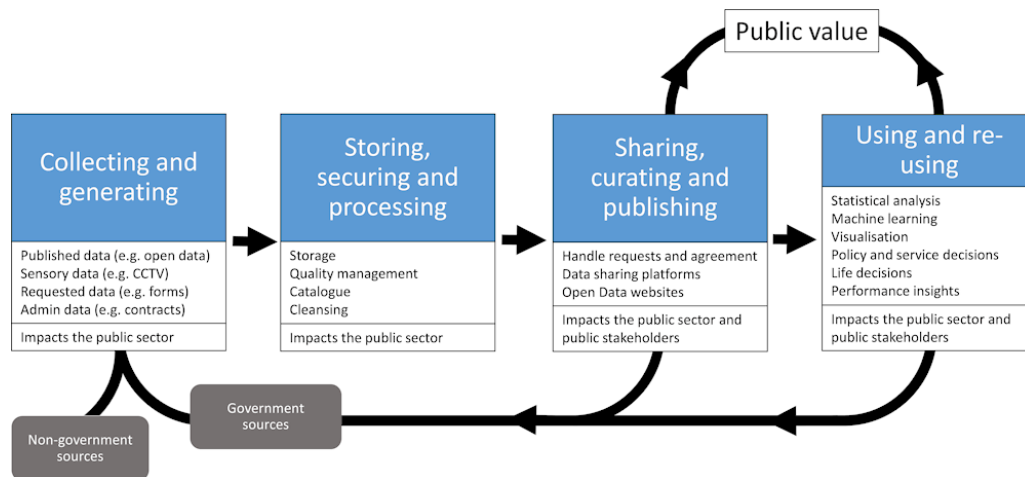
Source : van Ooijen, et al., 2019

The platform as a collaboration framework, offering a technical environment to facilitate the processes, can potentially affect all the steps in the value circle. This approach is further elaborated in Chapter 1.5 after looking into the concept of the platforms more thoroughly in the next chapter.

1.3. Public-private data collaboration

Collaboration among sectors is popular discussion matter for different scholars. While the importance, in essence, is agreed, the definitions and typologies are fragmented. One of the most popular terms in this context is the **Public Private Partnership (PPP).** It has many different interpretations, according to Linder (1999, p 41-48) six distinct uses of PPPs can be identified: management reform, problem reformation, moral regeneration, risk scattering, reformation of public services and power-sharing.

PPSs in science, technology and innovation are defined by OECD (2014, p 102) as: "*Any formal relationship or arrangement over fixed-term/indefinite period of time, between public and private actors, where both sides interact in the decision-making process, and co-invest scarce resources such as money, personnel, facility, and information in order to achieve specific objectives in the area of science, technology, and innovation.*"

Lachman (2000) distinguished collaboration types based on the relationship between the stakeholders: (1) partnership - close working relationships between teams; (2) collaboration - more complex, multiple organizations, shared purpose; (3) market (not a cooperative partnership, no trust needed, just a mean to make transactions).

One common involvement of external actors, including private for-profit companies, into public domain, is the public service delivery. Significant work has been addressed by for example Salomon (1981 and 2002), Osborne and Gaebler (1992), Kettl (1993), Sullivan and Skelcher (2002) etc. Involvement is taking various forms, such as large supply contracts and consulting, and it mainly involves public-procurement. Alford & O'Flynn, (2012, p. 19) bring out the following continuum of coordination modes of the service delivery: compulsion, supervision, classical contracting, negotiation, and collaboration. Collaboration is the form we are looking further into in the context of data.

Coming to the collaboration concepts that are exploiting the use of data - **"Data commons"** is one of the collaborative frameworks that is used in describing academy-industry collaboration. Grossman (2016) defined Data Commons in the context of Data Science as Service, that raises the research capability of academic institutions with the help of private sector by providing infrastructure and tools for data analyses (*Ibid,* p. 2).

Another definition for data collaboration is "**Data Clearinghouse**" (Lachman, 2000), which rather refers to a certain organization that deals with data as a mediator. "**Data Collaboratives**" - a broader concept that was introduced in 2015 (Ruijer & Meijer, 2019) to define the cross- sectoral collaboration on the (big) data field. The term is defined by Susha, et al. (2017, p. 2691) as: "*collaboration initiative aimed at data collection, sharing, or processing to address a societal challenge*". Most relevant definition in the context of this work, that will be taken as a basis for the rest of the thesis, is proposed by Klievink, et

al.(2018): "***data collaboratives are a form of cross-sector partnership for exchanging data and using it to generate public value***".

There is a growing tendency in the experimentation of the data collaboratives, and GovLab - one of the institutions exploring the concept, has gathered more than 150 examples of data collaboratives globally (GovLab, 2019). However, the research in this field is still rather limited. The focus has only recently shifted towards private sector (Bharosa, et al., 2013) in the works of digital government scholars, which data collaboratives concept is introducing and the need for new structures and practices (Susha, et al., 2018).

Studies have assessed the impact of the use of public sector data (open data) for the public interest. For example, by Carrera, et al. (2015) and Governance Models for Creating Public Value in Open Data Initiatives by Bolívar, et al. (2019). On the other hand, there are not many cases in Europe in relation to B2G data sharing put into studies, therefore, despite the perceived value of B2G data sharing, the practice remains limited. As actions to facilitate sharing data from business-to-business(B2B) in certain sectors are taken on EU level (for example, in banking integrating payment services (European Commission, 2015) and in common electricity market (European Commission, 2016) there potentially will be more initiated cases in following years.

This work will be exploring the dynamics of the public-private collaboration represented in one certain form - data collaboration platform, to understand which collaborative factors effect and enable the platforms development and how the infrastructure, information exchange and value creation from data are connected with the collaboration.

1.4. Typology of data collaboration

The distinction of cross sectoral data collaboration can be done in many levels, based on the literature and browsed examples. Starting from expected **outcome** that can present either direct **innovation** – input to specific research, service delivery etc. or open a marketplace allowing the **transactions** to happen. Nambisan (2019, p47) has proposed that the essence of collaboration platforms for wider public good can be categorized by desired outcome level as follows:
- Exploration (Defining problems and potential solutions, connecting stakeholders);

-         Experimentation (Develop solution prototypes, test, assessment of solutions);
-         Execution (Building solution templates, adoptions, facilitate collaboration, implement standards).

The European Commission strategy report on B2G data sharing stated that the main tools used for the B2G data sharing in EU are still **contractual agreements** (European Commission, 2020, p. 31). Even though, there is an understanding that this can be highly restrictive for the dynamic data exchange. Another aspect that can be restrictive in the context is the **duration of collaboration** - the models of public-private data collaboration can be roughly divided into long-term partnership based and one-time transaction based. The report (*ibid*) reveals that B2G data-sharing initiatives often take the form of standalone pilots and do not evolve to sustainable initiatives. Systematic usage therefor is limited, and collaborations tend to happen *ad hoc*. Also, the public procurement of data, as used often in public service delivery, is not seen as appropriate approach (European Commission, 2020, p. 20) in many cases. Mostly because the fact that data needs to be acquired timely and often constantly from various sources to have value.

Distinction by organizational arrangements within the collaboration - different decision-making and governance structures, and the level of engagement of the stakeholders, is proposed by Alford & O'Flynn (2012, p. 17): (a) Government decides, production of public value from data are shared, (b) Government decides, external party produces, (c) Deciding role and production is shared, (d) Deciding role shared, external party produces. GovLab (Verhulst et al., 2019) has looked also looked into **the institutional arrangements and operational dynamics** that have enabled private-sector data holders to collaborate with external parties to create new public value by analysing the **engagement level - the degree to which data supply and demand actors co-design the use of data assets.**

Another distinction can be carried out by **data streams presented** – either the data is moving from business-to-government (B2G), business-to-business (B2B); government-to-business (open government data), government-to-government(G2G), government-to-citizen (G2C), or the model is involving several streams. GovLab has also distinguished **6 collaboration practice areas** for the data collaboratives (Verhulst, et al., 2019): Research and Analysis, Public Interfaces, Trusted Intermediaries, Intelligence Generation, Data Pooling and Prizes

and Challenges. Platforms as a form of data collaboration can, in principle, be present in any of these practice areas.

Industry-academy collaboration in research is one of the most common data collaboration areas. Several authors have brought the challenges and opportunities of industry-academia collaboration into data era (for example: (King, 2019)), where traditional methods in social sciences may fail as it has become more difficult to complete the research without industry data, and even though tendencies from the governments are towards opening more datasets to the public, the industry creates more value and wishes to understandably monetize it. CGIAR – the global platform gathering the agricultural big data for research is an example of such collaboration that will be further looked into in Chapter 3.

Public interfaces are making the data of data-based tools or services available for the general public, and according to Verhulst, et al. (2019) include data platform and Application Programming Interfaces (APIs). The case of Metsään.fi explored below, presents an interface that offers services for the public. **Intelligence generation** stands for internally develop data-based resources that are released to the public. **Data pooling** is a unified presentation of datasets. And making data available to the external parties to develop solutions based on that is often done via competitions that include **prizes and challenges**.

The (trusted) **intermediaries** have been one way of reducing complexity by outsourcing the collaboration logic, that results as a service (Aulkemeier, et al., 2019). Service providers act as domain experts for particular collaboration scenarios (Aulkemeier, 2017) (Sherer & Adams, 2001). Different approaches and terms are used for the intermediators. Perkmann (2015) suggested a term "boundary organization" to facilitate the data and goals for all the stakeholders. The intermediaries can be representing either public or private sector. One advantage seen on relying on intermediaries is the better possibility of discovering new business opportunities, likely for the private sector-initiated intermediaries who are devoted to this role and as the for-profit goal is to monetize their services.

Data marketplace can also be considered as a form of trusted data intermediator. Data Markets as such are currently rather undeveloped. Only a few of them exists today (e.g. Dawex, Nallian, Deutsche Telekom's Data Intelligence Hub) while data brokers usually act

as intermediaries providing processing services when it comes to commercial offerings of data (European Commission, 2020). The lack of marketplaces and not well-defined market, is one of the reasons why regular procurement of data is difficult. The challenge could be also in the essence of data – changing in time.

Hautamäki A. (2018) has brought out how marketplaces in the private sector are used in coordinating market transactions in a rather efficient way, but public-sector platforms cannot be built by transposing mechanical models of the private sector. This is due to the market logic of public services which differs from open markets. One of the case-studies will involve data marketplace– Copenhagen City Data Exchange – to understand the reasons behind the initiation of this collaboration model from the public sector perspective and looks into the dynamics of the collaboration on this type of platform.

Another distinction of different data collaborations are the different technical solutions– in what state the data is presented and whether the data publishing is done via Application Programming Interfaces (APIs) that generate the access for data, but does not visualize it or platforms that provide the access to data or its end products usually via web portal or mobile application. Data brings in other many technical variables, such as accessibility – **open and restricted access**; presented data types (raw, pre-processed, etc); data governance principles of assuring security and privacy. These distinctions are too scattered to bring them into the typology of the collaboration, but the most acknowledged aspects will be involved in the next chapter.

The typology of the different distinctions of the data collaborations is presented in Table 1 below and how and if the aspects are relevant in the context of this work will be further elaborated in subchapter 1.5.

Figure 2. Data Collaboration different typologies

Figure 2. Data Collaboration different typologies
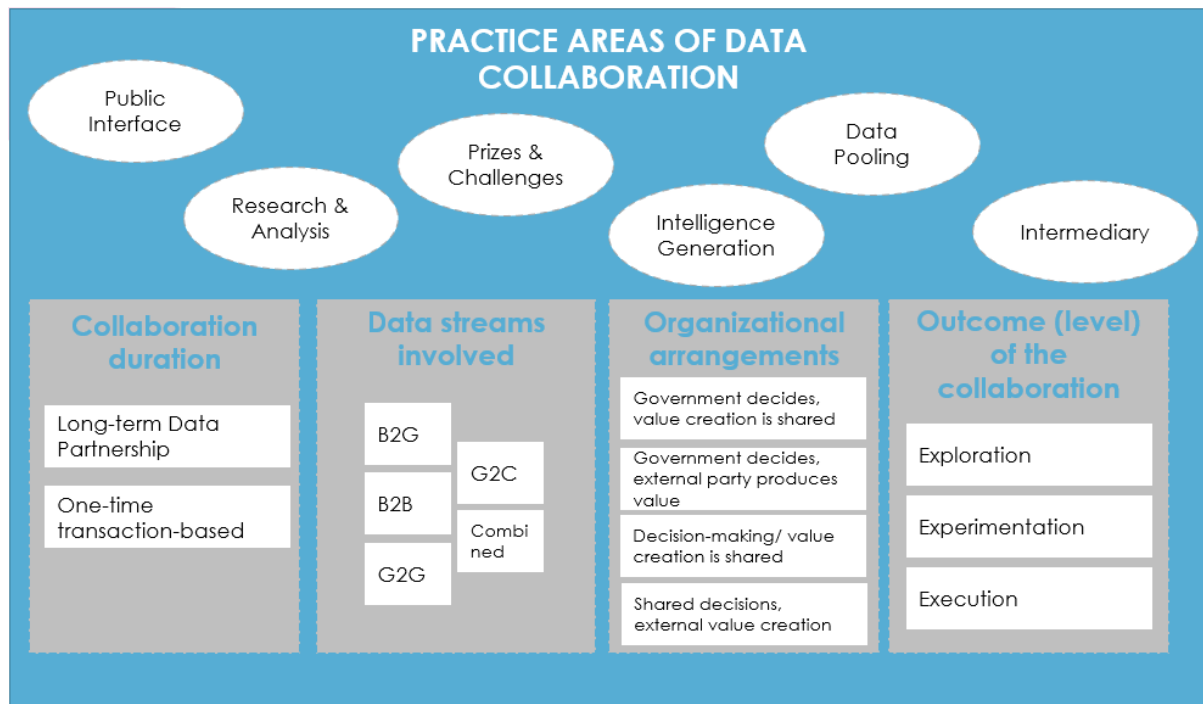
1.5.   The concept of platforms

Platforms have been popular in research and studied from various perspectives - organizational, structural, functional and technical. Dominant topics are related to technical infrastructure and value creation in business settings. Platform in data collaboration has mostly seen as an interface to share the results. This chapter expands the term and argues that platform, when looked as a wider concept can be beneficial for the collaboration and collaborative value creation.

The platform concept can be seen complementary for the idea of collaborative governance regimes (CGRs) proposed by Emerson, et al. (2012) and defined as a (Emerson & Nabatchi, 2015, p. 18): "*type of public governance system in which cross-boundary collaboration represents the predominant mode for conduct, decision making, and activity between autonomous participants who have come together to achieve some collective purpose defined by one or more target goals*". Collaborative governance brings public and private stakeholders together in collective forums with public agencies to engage in consensus-oriented decision making (Ansell & Gash, 2008). Connections between collaborative governance and consensus-based decision-making in the chosen data platform cases will be looked into in Chapter 3.

Ansell and Gash looked into collaborative platforms as a strategy meant for collaborative governance, and defined "collaborative platforms*"* as organizations which in the broadest sense are structured frameworks for promoting collaborative governance (Ansell & Gash, 2018). Nambisan (2019) defines platforms as a particular set of practices and systems that are needed for collaboration. Platform-based approaches are offered to simplify inter-organizational collaboration in the private sector by several authors, such as for example van Hillegersberg, et al. (2012) and van Heck and Vervest (2007). Platforms can be an attractive form of collaboration, as they offer a large amount of levels and opportunities of interaction between parties- both horizontally between organizations and vertically - for example, shared services in the public sector (Klievink, et al., 2016).

On a more general view - an environment that allows an open, flexible and demand-driven collaboration has been referred to as **digital ecosystems**, often associated with digital platforms. The platforms are seen necessary to enhance the quick connect capability and transform (business) networks into digital ecosystems (Aulkemeier, et al., 2019).

Digital platforms are defined by Klievink, et al. (2016) and Kenney & Zysman (2016) as: "*multisided digital frameworks or socio-technical artefacts, powered by the connectivity and enabled by data, that shape the terms on which participants interact with one another*". The enablement of the cloud computing and algorithms has created an entire platform-based ecosystem in the private sector – often also referred to as "platform economy".

Markus and Bui (2012) see IT-based platforms as tools for interoperability (in processes and data) for interactions among the organizations, in particular industrial communities, that they define as inter-organizational coordination hubs. A collaboration platform in business is a category of software that enables the work process with a goal to foster innovation by incorporating knowledge management into processes (Magi & Maruping, 2019).

The platforms often tend to be in essence multi-party networks, that require the involvement of multiple parties, whom the public sector can have a separate dynamic relationship. As for many state services and public value creation process, there is a need for multiple sources of information, this brings another layer of complexity – the dynamics of governing multiple

parties involved in data collaboration. A digital platform has seen by Hautamäki A. (2018) as a new way to coordinate the actions of a great number of actors in society.

The platform approach here can also be linked with the of **Government as a Platform (GaaP)** framework realizing the opportunities of the linked world and Web 2.0 in governance and adding platform thinking to government technology projects and its wider role in society (O'Reilly, 2010). This framework is about the co-creation and the importance of the information produced by and for the citizens, and learning from the private field, but does not look into the dynamics in cross-sectoral platforms as we aim to do in the following pages.

Klievink (2015) is one of few authors who has looked into collaborative value creation configured in the form of platforms that deal with the information exchange – he explored collaborative platforms as a tool for outside-in transformation for the government. His work looked into governance and infrastructure of the platform, exploring how government can exploit the platforms to transform. He emphasized that collaborative public-private platforms necessitate combining technical and governance instruments (Klievink et al., 2016), expanding the platform as an enabler of the collaboration that this work is looking further into.

Combining the previously mentioned distinctions, this work defines Data Collaboration Platforms as following: **Data Collaboration Platforms are multisided digital frameworks, shaped by the interactions of stakeholders, powered by the connectivity and enabling value creation from data.** Next chapter will take the concept of exploiting platforms in value creation from data further – looking into enablers of collaboration and value creation on platforms.

1.6.   Collaboration and value creation on data on platforms

Data collaboration between public and private sector is a rather novel topic, only recent years some research on the enabling factors have been published. Klievink, et al. (2018) presented an analytical lens for understanding on how data collaboratives are enabled by three main factors - starting conditions, context and leadership, while the outcomes come through trust-building and collaboration process. World Economic Forum and McKinsey (2019) brought out enablers of public-private data collaboration that are related to stakeholder alignment,

good governance, unbiased insights, support tools and ensuring the long-term economic sustainability. This list provides both - elements of collaboration and value creation enablers. In the context of this work, they should be looked separately for more detailed understanding of the data collaboration platform and the enabling factors of its development.

During the writing of this thesis European Commission has published an assessment of the general obstacles for data collaboration, where the additional layer is presented specifically with the B2G data sharing issues (European Commission, 2020, pp. 31-36): (1) Lack of structures to support data transfer (main tools contractual agreements); (2) Shortage of dedicated functions in the field (professionals); (3) Economic barriers (such as entry cost, monopolistic pricing) (4) Fragmented landscape – in some EU member countries horizontal or sector-specific legislation for data sharing with the public sector; and (5) Legislation differences in the Member States.

Platforms could potentially provide a framework that supports overcoming several of those obstacles and enable the cross-sector data collaboration. Platforms' dynamic architecture could facilitate services or tools to support the collaboration and enable value creation. Sectoral platforms can potentially help to achieve the cross-border collaboration and reduce economic barriers. Well-functioning and user-friendly processes for the data gathering and processing could help to overcome the issues with the lack of professionals on the field. If and how this is present in the real-life and what role platforms play in overcoming the obstacles will be further looked into while exploring the cases.

Coming back to the research question– **What are the main enabling factors of cross-sectoral data collaboration platforms development and public value creation?** The author hereby proposes a three-layer framework (Figure 3) where data, collaboration and platform are separated into layers, but still strongly interconnected and affecting each other. Concentrating on the **enabling factors of collaboration**, other layers will also need attention. The proposed framework is based on the collaborative factors presented in existing research, aims to further decipher the platforms in data collaboration context and is the basis for analysing the cases in the next chapters.

Figure 3. Data collaboration platform assessment framework
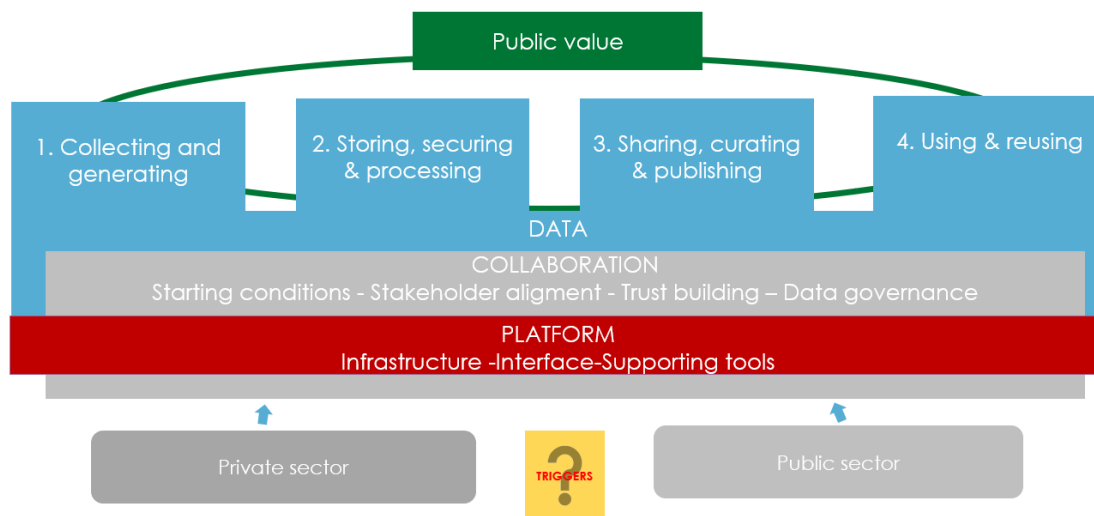
Source: Author, based on van Ooijen, et al. (2019)

The work is further exploring the triggers of the data collaboration platform through analysing the objectives to understand the benefits of the platform model and its enablers through the eyes of initiators. Achieving the objectives is closely tied with the selection of the output - how the results are presented and/or measured. Platforms in most cases offer (public or restricted) interfaces to present the results, therefore, their contribution to sharing and publishing in the circle of the value creation from data is most visible. But looking deeper in the processes gives us an understanding of **how the platform collaboration affects and enables the value creation process in whole**.

Based on the abovementioned research and initial analyses of the cases, the **collaboration aspects** that appear the most relevant for enabling the developing the cross-sectoral data collaboration platforms and should be validated are **supportive preconditions, stakeholder alignment, trust-building and governance.** Mentioned aspects influence on enabling the successful development are looked into, while keeping the research open to any additional collaborative aspects that have demonstrated influence on the cases.

Starting or preconditions are considered to be any prior collaboration or other existing relations that collaborators have, in addition to any pressure from peers (government for example). Ansell and Gash (2007) referred that this connected to initial trust level, resource asymmetries and incentives for participation. Their existence should create an environment

that helps to resist the opportunistic behaviour that is known to obstruct international IT endeavours that platforms are (Premkumar, et al., 2005). Starting conditions affect the collaborative process and the trust-feedback loop therein (Klievink, 2015).

In theory, all public sector-initiated platforms are launched to achieve public good. Even if the value is clear for the public, the motivations of external providers are more complicated. The concept based on French & Raven (1959) and presented by Alford & O'Flynn (2012, p. 57) has proposed that the external powers propensity depends on (1) willingness (fundamental drivers) and (2) ability (depends on facilitators, the ways government can make contribution easier). Platform as a facilitator could potentially offer tools to raise the ability for the external providers to provide data for the set value with less effort. The ensured value provision for all the stakeholders, while creating public value, is part of the **stakeholder alignment** that should be looked into.

Klievink, et al. (2016) has brought out that too much attention to the structure of information (data) compared to limited attention to **the governance structure** of the platforms can lead to delay in transformation, which can deduct the value of the results. The cases will look into if and how this issue is solved, exploring the models and their effect of **organisational arrangements** throughout the value creation process from data - where the decision-power lies and how is the stakeholder alignment and responsibility of the value creation divided. Which tools and processes are set up, or facilitated on the platform to enable the collaborative value creation? In theory, the more institutional arrangements are agreed and on place, the easier it should be to build up trust and data exchange.

Overall governance of the platform is also related to the understanding the **legal context** where the platform operates and how it is communicated trough platform for the trust-building. Closely related to this is good **data governance** that is seen to benefit in common goals and mandate of the participating organizations, who produce, access, share or reuse common datasets (OECD, 2019). The presence of specific procedures within the platform collaboration that are assuring the good data governance and its implementation for the public good should be therefore detected and their impact assessed.

The value creation from data throughout four Data Value Cycle steps: 1. Collecting and generating; 2. Storing, securing and processing; 3. Sharing, curating and publishing; 4. Using

and reusing, are looked through both layers – platform and collaboration. The cases will look into how the detected preconditions enable the value creation process, if and which support platform as a framework can offer in the process.

To conclude, the work will assess if any of the typologies (presented in Figure 1) are dominant within the platform collaboration, and what is the reasoning behind in such cases. This would help to understand the platforms better in general data collaboration context and could provide beneficial input in planning future collaborations.

2. RESEARCH METHODOLOGY

The focus of this thesis is to explore the platforms enabling cross-sectoral data collaboration in real-world settings, looking into existing cases. Therefore, **qualitative research method** was chosen (Golafshani, 2003) (Babbie, 2014). Qualitative research, in most usual cases, aims to describe the form of what exists, examine the reasons behind that, associations between the stakeholders and appraise the effectiveness of the existing practice to develop theory and strategy for further actions (Ritchie, 2013).

One of the aims of the qualitative research method is to look at the research object from the perspective of the participants and develop the empirical knowledge via discoveries made by this exploration. Understanding a specific phenomenon – platform as a data collaboration tool and the answers for the questions presented in Chapter 1 can, therefore, be found through the method.

Case studies are one of the preferred strategies to research a contemporary phenomenon in a real-life context, while there is limited research on the relevant topic (Yin, 2008) (Eisenhardt, 1989). The focus of the case study research is to explore and understand the existing dynamics within one setting (Eisenhardt, 1989) – in our case data collaboration platform. Multiple sources of information are used, during the case study research to build additional theory around the topic.

The research question is presented trough "what" question, which according to (Yin, 2003) favours methods of archival analysis and surveys. As the research in this context is scattered, there is limited data for the archival analyses which limits the number of potential survey respondents. The case study method is not relying on prior empirical quantitative and qualitative data. Therefore, it allows the analysis of a small sample applicable to "what" questions (Eisenhardt, 1989).

Yin (2008) refers to four different strategies based on two dimensions while working with case studies: (1) single case versus (2) multiple cases – considering the number of cases examined; and (3) holistic versus (4) embedded – referring to the unit of analysis. To better understand the dynamics of the platforms in data collaboration, this work has chosen **multiple and holistic case study strategy.**

Initial research questions were defined to find the focus and were improved while more information was gathered. Defining the sample of the cases, was followed by the data collection by multiple methods, including interviews, reports, webpages, presentations and media-articles.

The selection criteria of the cases was based on **theoretical sampling** (Eisenhardt, 1989) as there is limited literature available on the topic and this approach does not require theoretical propositions in advance. The aim was to **find polar types and provide a maximum amount of interesting observations** but keeping in mind that the aim of the work also is providing insights for further emerging cases, therefore the cases represent the common use cases of cross-sectoral platforms.

Following characteristics of the cases were taken into consideration while selecting:
- The platform should be beneficial to the wider community of people directly or indirectly;
- The platform should include elements of cross-sectoral collaboration;
- The platforms' main usage field is data exchange;
- Sufficient information should be available for the initial analyses based on the research question and aim in the adoption of the practical considerations opted by Dul and Hak (2008).

Data, in its essence, is not limited to state borders and international collaboration is common among the data initiatives. This, on the other hand, does not reflect the deficit of the regional or local collaborations. On the contrary – we can see the collaboration platforms on different levels and with different scopes. To ensure that the depth of the examination would be sufficient, selected cases are represented from regional, national and global level. That allows us to look at the enablers of collaboration development and value creation on different scopes and have a comprehensive overview of the factors useful for a wider audience. The short description of the cases is brought out below in Table 1.

Table 1. Descriptions of selected cases

Case	Description
CDE	CDE was a collaborative project between the Municipality of Copenhagen and Hitachi. The purpose of the collaboration was to examine the possibilities of creating a marketplace for the exchange of data between public and private organizations.

Metsään.fi	The platform is utilizing forest data gathered from private and public sources via offering datasets and services to (private) forest owners, service providers and the wider public.
CGIAR	The CGIAR Bid Data Platform aims to increase the impact of agricultural development by embracing big data approaches to solve different development problems.

In order to explore the cases thoroughly to find the learnings and distinctions needed to make conclusions and adapt those into the context, a variety of different techniques in data collection were utilized. **The primary data about the cases for the core of the case studies was collected via analysing documents, reports and web pages**. As the available material and analysis of the cases was one of the preconditions of choosing them, those materials (containing analyses, strategy documents, statements and explanations on the web-pages) were exploited to build the bases for the case studies. All of the used data sources are presented under References.

Semi-structures interviews (SSI) were carried out with the representatives of platforms to get more unique insights for the research questions and supplement the document analysis. Three (3) interviews were conducted in spring 2020 via means of telecommunication. To support answering the two-sided research question interviews contained two blocks – a collaborative platform enabling aspects and platform framework as an enabler throughout the 4-step Data Value Cycle. The questions are brought out in Appendix 1 and the list of interviewees is presented in Appendix 2.

The analyse overlapped with data collection to speed the process and improve the aspects included to fit the right focus. Interview questions were shaped based on the initial data found during the research and cross-case comparison and search for causality was carried out through data collection via both- sources of web-data and interviews. The final step was linking the findings into the theoretical framework, adding value to work done so far regarding the platforms as data collaboration enablers.

3. CASE STUDIES

The following chapter will look into the actual cases of different data exchange platforms, to test the proposed framework (Figure 3), bring out best practices within the and modify the components if appears necessary.

3.1. City Data Exchange (CDE) – regional marketplace for data

The explorative and experiential aims were mixed in The City Data Exchange (CDE) launched by Municipality of Copenhagen. The collaboration was triggered by the objective to develop a marketplace for all data users and providers, aiming to diversify the available information which can be leveraged for unforeseen business or societal ventures (Municipality of Copenhagen and Capital Region of Denmark, 2018). In addition, the Municipality aimed to showcase Copenhagen as an innovative city by creating the widely accessible data source for the creation of the data-driven services- both in the private and the public sphere (Interview 1).

This CDE operations were concluded by 2017, but the team has been concentrating on sharing the learnings and since CDE brought attention to several challenges, it gives an interesting perspective to explore within this work. For example, as a result of several barriers, **the collaboration platform was assessed to have a fairly low impact on the project** (Municipality of Copenhagen and Capital Region of Denmark, 2018). Many of the disabling factors were present specifically related to collaboration, such as the preconditions and trust building.

CDE initiation was a cross-sectoral collaboration itself and has been branded as a collaboration ever since. The organization and a technical platform were established in collaboration with Hitachi (a Japanese multi-national corporation). The participation in the project strongly aligns to their mission and strategy (Hitatchi, 2016). The decision-making related to platform development was carried out in collaboration - Hitachi and the Municipality were meeting regularly to discuss the outcomes and shape the platform taking the information collected into account (International Data Corporation (IDC) and the Lisbon Council , 2016).

Even though the expertise for the platform development and administration was outsourced, the municipality has invested to have the data related expertise developed inhouse. Dedicated institution - Copenhagen Solutions Lab - part of the Technical and Environmental Administration in Copenhagen Municipality is dealing with intelligent technologies to create data-driven solutions (Copenhagen Solutions Lab, 2020).

Municipality of Copenhagen report (2018) divides CDE into three layers: (1) a **collaboration** between the different partners on supply, and demand of specific data; (2) a **platform** for selling and purchasing data aimed at both public, and private organizations; and (3) an effort to establish further **experience** in the field of data exchange between public, and private Organizations. The third layer, could be considered successful, even though the platform that did not find its way to the market.

Data was collected through several different sources - cell phone tracking, wireless connection counting, camera image counting, traffic sensors, visual surveying, ticket purchases, etc – gathering more than 140 datasets in its first year (Municipality of Copenhagen and Capital Region of Denmark, 2018). It included the collection from the open sources, but also from the private sector – institutions like Danish Technological Institute, Saxo, GoBike, EverImpact, and Dexi.io (Municipality of Copenhagen and Capital Region of Denmark, 2018, p 12).

Regarding private companies, often the case was that companies share once but they have no interest in further sharing and they are not willing to share data for free (International Data Corporation (IDC) and the Lisbon Council , 2016). This is since an elevated effort in terms of resources and time are encountered and lack of skills and motivation was overwhelming (Interview 1).

The data exchange between the seller and purchaser was nearly always needed to be done case-by-case. Therefore, the model was not sustainable nor scalable, even if a large variety of datasets were offered. All the data that uploaded was aggregated and made anonymous. Yet, the ethical questions were holding back many organizations to share their data to the platform (Interview 1). The platform was presenting the necessary ethical and legal backgrounds, but was not used to present any information or case studies for encouraging the data collecting.

Even if legal aspects are in order, concerns raised about the public (customers and media) option. Success-cases showing the public value were missing.

Workshops and stakeholder interviews (~1000 people) were carried out during the project (Interview 1). It was often the case that data seekers do not actually know which kind of data is useful for them. Also, it was brought out that that data requested was fairly specific and not readily available - therefore expensive to offer and not justifying the business case (Municipality of Copenhagen and Capital Region of Denmark, 2018). Demand-side from both – private and public sector, was fragmented and no one-size fits-it-all models could be implemented to data offer. Every case needs a different combination of data-sets and combining the offered sets by each of them was needed.

The processed and shared data on the platform was attractive to several international start-up companies, who saw the business opportunity and were looking to use the data to work on new solutions (Interview 1). In addition, the learnings can lead to building successful cases and there are already similar models taken into practice, like Amsterdam Data Exchange (AMDEX)(City, 2019) for example, that already seems to put more effort in design the approach before launching the service. Different use cases and other opportunities or trust-building are presented at AMDEX.

3.2. METSÄÄN.FI national platform for forest data

The of Metsään.fi platform aims to provide efficient use of the forest resource data collected in connection with the Finnish Forest Centre's (FFC) regional planning. The strategic objectives in the background of the Metsään.fi according to the online service concept created in the FFC 2010 (Valone, et al., 2019) were the following:

- to improve the possibility to run errands with the authorities online;
- to make the utilisation of forest resource data that has been collected with public funds more efficient;
- to advance the implementation of forest political objectives;
- to implement the Forest Centre's role as a bridge-builder between forest owners and forest sector actors.

The initial technical development of the platform was carried out by a private company (CGI) who won the competitive tendering. The service development and further refining has since

then been carried out by FFC and notable development within in-house data stewardship has been carried out - there are several technical and interoperability specialists working with the platform and strategical approach developed for the data sharing (Interview 2).

The main stakeholders on the platform are private forest owners, service providers, and FFC. To offer value to all the stakeholders involved, surveys on customer expectations before launching the service and usability tests were conducted (Valone, et al., 2019)**. Involvement of stakeholders in the development of the platform has been ongoing**. FFC offers the advisory service for all the Metsään.fi users. The stakeholders can provide their feedback via the advisory service contact. Furthermore, workshops are organized now and then, where the future development needs are gathered (Interview 2).

As FFC has been operating in the forest sector for decades**, there is a pre-existing collaboration** between the stakeholders, which gives a better starting position and is helpful in trust-building needed in data collaboration (Interview 2). Metsään.fi platform is an example of providing a compelling view into the public value of increasing the availability of data collected from private sector - Finland's privately-owned forests. A value provided for the private sector parties - overview for the forest owners, the business case for service providers, supports the collaboration further and is a precondition for data gathering.

Metsään.fi is a collaboration platform beyond sharing the data, offering the services based on data and collaboration opportunities for the private forest owners and service providers, while forest owners are in control of sharing their data (Interview 2). The forest owners also have the opportunity to update the data on the platform (*Ibid*).

FFC obtains data from the public sources and private sector through legislation, and by purchasing (Interview 2). Collecting forest resources data from privately owned forests in Finland is one of the statutory tasks for the FFC (Metsäkeskus, 2020).

There are two main outputs for the data on Metsään.fi –data-based services for the forest owners (restricted access – needs login) and open forest data presented on the platform. There is a legal regulation for making the forest data available for the forest owners. Metsään.fi-e Services offer the latest information directly to forest owners on their properties (Interview 2). Specifically, people who own forest property in Finland can conduct business related to

their forests through the platform. It connects owners with related third parties, including providers of forestry services. Moreover, the portal saves service providers the cost and effort of visiting sites to obtain the latest data on which to base planning. It also contains up-to-date contact details of forest owners.

FFC and Metsään.fi operate under the steering of the Finnish Ministry of Agriculture and Forestry. The ethical data governance is based on the Finnish Forest Act and other European legislation, for instance, the GDPR etc. (Interview 2). There is no separate data legislation needed for Metsään.fi. But FFC is leading a steering group for the forest data standardisation that aims to help further developing and improving the collaborative services (*Ibid*).

Standardisation should help to overcome one of the biggest challenges in data governance - integrating several different sources of information that offers all data simultaneously (Valone, et al., 2019)*,* which contributes to the up-to-datedness of forest resource data and further improves the quality. Another particular challenge that the interview (2) brought out is the need to constantly renew the platform framework due to the technical challenges as the technology is getting outdated.

3.3. CGIAR – Global platform for big data in agriculture

Consultative Group on International Agricultural Research Consortium (CGIAR) has launched the Big Data platform to harness the capabilities of (big) data to accelerate the research and development and conquer the challenges faced in agri-food sector (CIAT and IFPRI, 2016). The CGIAR platform is mainly focussing on the agriculture, but because of the linked nature of the field, it includes datasets across the sectors. In comparison to other cases, CGIAR platform is triggered with the research aim.

The CGIAR and the big data platform is led by International Center for Tropical Agriculture (CIAT) and International Food Policy Research Institute (IFPRI) and concentrating together all the 15 CGIAR Research Centers and 12 Research Programs, alongside 70 external thought partners ranging international institutions, universities to private companies (CGIAR, 2020).

During the Platform development, there was a consultation phase that included technology companies, agribusiness start-ups, etc. All the stakeholders are invited to the annual Convention to engage with the 6 technical communities of practice and around areas of research interest (Interview 3). There is an emerging collaboration in data sharing and data standards and in digital innovation/tech transfer within the stakeholders thanks to platforms (*Ibid*).

This case has the biggest example of involved private companies as data providers. GIAR is a global agricultural innovation network itself, and this means that the previous connections were supporting the launch of a data collaboration platform. Interview 3 stated that: "*CGIAR continues to have a reputation as a neutral public interest player, for the most part, and this enables us to weigh in on how the industry as whole could be working better. So, I suppose that reputation is built on prior collaboration and this has enabled us make alliances for data discovery*." The value proposition is very much driven by use-case or research interest. The discovery of a partner's data alongside all the open data from CGIAR is seen attractive to some partners, especially those that have open data policies they adhere to (Interview 3).

The governance model of **CGIAR distributes the strategy, governing and advisory models among several entities** to make sure that the diversity of stakeholders is represented (CGIAR, 2020). The platform serves as a data intermediary between CGIAR centers and other partners and has data use agreements that defaults towards openness of the products resulting from data use and sets up a framework for collaboration (Interview 3).

GARDIAN, the Global Agricultural Research Data Innovation & Acceleration Network, is the CGIAR flagship data harvester (CGIAR, 2020) and seen as the main enabler of a data exchange as a platform (Interview 3). GARDIAN enables the discovery of datasets, publications and repositories across all CGIAR Centres to enable the reuse and innovations based on data. CGIAAR has its **Data Management Strategy published** with the following main principles (CGIAR Research Program on Climate Change, Agriculture and Food Security (CCAFS), 2015): Accessibility and Open Access ; Ease of use; Facilitating increased collaboration; Ethical use and sharing of data; Provision of support for data generators ; Ensuring that credit and visibility go to data generators; Adherence to international standards for data storage. To support the implementation of the strategy the

Data Management Support Pack (documents, templates and videos covering the different aspects of data management) is made available on the platform (CCAFS, 2020).

CGIAR Platform is guided also by the principle of making data FAIR – Findable; Accessible, Interoperable; Reusable (Wilkinson, 2016)) The primary focus in the early years of the Platform was to get CGIAR data resources organized and FAIR. The Platform has been working on establishing the infrastructures, tools, and approaches to make CGIAR data visible and usable, and has developed FAIR Metrics and downloadable Guidelines.

The platform has enabled access to shared services in support of CGIAR research and offers open-access to its datasets and set of services, and capacity building include courses and webinars. It offers a series of services ranging from services for metadata access via easy-to-use APIs that can be used by any institution. Full access to GARDIAN's source code is also available.

The platform provides **datasets and insights**, that can be implemented into action. In addition, **trainings** and best practices are provided. Here we see also an example of another data collaboration form initiated on the platform – Inspire Challenge. This hackathon type of yearly event provides an opportunity for the outside CGIAR researchers to partner with research centres in CGIAR to solve big problems with their ideas(CGIAR, 2019).

4. RESEARCH RESULTS AND DISCUSSION

The cases above were explored to validate the proposed three-layer framework and enabling aspects within platform environment in data collaboration process – presented in the Figure 3 in subchapter 1.6.

The wider **objective** of all the presented cases has been making use of data as a resource - involving the available data from the public and the private sector. This being the first and foremost reason the cases are considered to be data collaboration platforms. However, the concrete aims of the cases vary from accelerating the agricultural research to showcasing the innovativeness of a region via data marketplace. While sectorial platforms – like Metsään.fi have a narrower scope of data gathered and offered, then in CDE case attracting private sector from various industries and failed to put the demand and offer together in this fragmented context. On the other hand, CGIAR demonstrates the successful fulfilment of objectives even though it contains the biggest amount of data sources and bold goals.

Exploring three different initiation-model of data collaboration platforms – public sector lead Metsään.fi; private sector lead CGIAR and CDE which was sort of a dynamic solution, initiated by the public sector but largely also executed by the private sector –shows no direct effect in the context of this work, which does not mean that there is no effect in a general and especially in value creation perspective.

The **triggers** of the collaboration initiation between the public and private sector are mainly related to lacking input of data or expertise. Meaning, cross-sectoral collaboration is often represented in building the platform, in data gathering and value creation. Public-private partnerships are often already part of initiating the data exchange platforms, as the required technical knowledge to deal with large databases is not available in the public sector, and might not be reasonable to develop.

While the technological aspects are not the focus of this work, they are an important aspect of the platforms. Having a **trustworthy infrastructure** is one of the preconditions of developing data collaboration platform, and keeping the technology from outdating can be one of the biggest challenges for the platform initiators, as it was brought out in the interviews (for example with Metsään.fi representative (Interview 2). Outdated technology or

non-user-centric design can harm the potential of the platform as a collaboration tool and value creation enabler. If users (data providers or service end-users) face difficulty finding the information they need or struggle providing data – the engagement will be low and collaborative actions may degrade. On the other hand, as presented in the CDE case, **if the other preconditions are not set then the platform, together with its infrastructure, has rather low impact on the collaboration.**

Starting or **preconditions** both - technical and non-technical, prove to be one of the main enablers of the data collaboration platforms. Trustful relationship between the stakeholders is one of the most notable preconditions for successful development of the data collaboration. **Trust** is not self-evident and requires either proven previous experience in the field and previous relationship (partnership) with the stakeholders. Metsään.fi and CGIAR and the institutions behind the collaborative platforms have years of proven sectorial expertise and strong previous collaboration with the stakeholders, while in CDE case this was more complicated and it faced severe challenges in trust-building.

Achieving **stakeholder alignment** and involving stakeholders, both – providers of the data and the end-users on the platform to service design, allows to validate the value offered to all parties. **The effort to involve the stakeholders has been carried out in all cases**, but via different means and varying in depth and thoroughness. The importance of doing that already prior to the launch, to bring value to all stakeholders, was emphasized especially by both the CGIAR and Metsään.fi representatives. While specific value is important in all the cases, providing an understanding of the wider perspective – the public value creation – is seen crucial to be achieved in long-term collaborations.

Values shall be aligned and data sharing should be with a reasonable effort for all the providers – this can be made easier with the user-centric approach on the platform. While some public initiated platforms, like in Metsään.fi, can lie on the legislation that can be used as a basis to collect the data also from the private parties, it is not always the case and if the value creation also involves business cases (CDE), the private sector tends to be extra cautions. Private companies wish to control the usage of their data for strategic concerns. In CDE this required case-by-case approaches and was considered one of the reasons why the platform resulted unsuccessful and broker services were seen a better solution. Here, one can argue if better preconditions could result in the success of the platform on not, but **surely**

there are models of collaboration in data exchange where the platform is not the most suitable solution.

Another important aspect of success that was highlighted by the platform initiators was that stakeholder alignment and trust-building should come together with making the **processes in multi-stakeholder decision-making transparent** and line with the values. In addition to legislation behind the framework, the platform enables trust-building via publishing the structure of decision-making and procedures - like in CGIAR platform case where several published documents regarding decisionmakers and -making be found.

The value creation from data is strongly connected with the trust, ethics, and legal framework which brings in the complexity to **data governance**. **Assured ethical and transparent governance of the data is utmost important as reflected by all the cases**. Platform is used as a tool to present the good data governance principles and build trust. Considered that the trust level is also related much with the level of control that partners can have over the data, clear goals and purpose for data and governance that assures this is also needed.

A set of ethical standards is not good by itself setting morally good outcomes, but is likely to facilitate the actions towards this (Floridi, 2014). The same goes for the platform. There are a couple of theoretical frameworks for ethical governance in the era of data-driven decision-making algorithms, such as FATEN (standing for fairness, autonomy, trust, education, non-maleficence) proposed by Nuria (2019) or FAIR – Findable; Accessible, Interoperable; Reusable (Wilkinson, 2016)) used by the CGIAR, what acknowledges the fact that when algorithmic decisions affect thousands of people, important ethical dilemmas arise – regarding the control and responsibility, where human should always be the core. It is crucial to build trust and common ethics in the among the whole data ecosystem to enable more collaborative data initiatives.

Platforms can "translate" those ethical principles to be relevant in their field, but in most of the cases, there is a need for combining the general principles with legislation and tailored privacy policy – which is appropriate for this specific platform. Apart from the understanding of the regulations, it is also important that data holders fully understand the potential value created to justify the actions taken for the users.

All the explored data collaboration platforms objectives reflect the aim to create public value. Even if the scope the value creation from data differ greatly even in the small sample that was explored here - starting from creating more innovative municipality, and ending with research towards ending the hunger and offer better public services for forest owners. The methods remain similar and **platforms are demonstrating the potential to support all the 4 steps of the value creation** in explored cases by enabling value-creating interactions through its collaboration enablement and with the dedicated infrastructure, which presents a framework of rules set by the platform governors.

Platforms pose an environment that in addition to providing an interface to collect data, enables collaboration for sharing it in a different format– from processed data to visualizations, set of services and build the trust. But this does not come without challenges that need a lot of resources, both time and finance - starting from building a user-centred platform to keeping the technology up to date and assuring the security while dealing with such a delicate asset as data.

Platforms can be, in principle, present in any of the practice areas in the typologies (presented in Figure 1). The cases explored involved combined data streams and showcased that **platform structures and collaboration seem to work better in long-term partnerships** (transactions within CDE vs legislative approach of Metsään.fi) that are enabled more in the environment where value creation is shared, and value is provided to all stakeholders. Decision-making should reflect the interest of all stakeholders.

CONCLUSION

This work aimed to understand the main enabling factors of cross-sectoral data collaboration platforms development and public value creation by exploring a polar example of cases that reflect popular data platform representations.

The theoretical contribution of this study is presenting a framework where **three layers** of data collaboration platform - **collaboration, platform and the data** are presented separately, but are strongly intertwined, and which can be used to conceptualize a collaborative data platform. In a practical perspective, the work presented an overview of the crucial elements that have to be achieved to develop a well-working data collaboration platform that can be implemented in the future initiations.

Validating the proposed framework via case studies provided an answer for the research question - What are the main enabling factors of cross-sectoral data collaboration platforms development and public value creation? **The main enabling factors of the data collaboration platforms development in the context of this work are the supportive starting conditions, successful stakeholder alignment and trust.** The cases emphasized that successful data platforms are built either on the "earned" trust via previous actions on the field (Metsään.fi) or the previous collaboration between the stakeholders across sectors (CGIAR). **Strategy of earning trust and involving stakeholders from the development phase should be part of each data collaboration platform development.** A sufficient stakeholder involvement process could help to address the lack of cooperation or experience. Up-to-date and secure platform infrastructure with the supportive legal and ethical framework is another important point to be addressed. The need for dedicated resources should not be underestimated and should be kept in mind while planning the data collaboration platform.

The main enabling factors of public value creation on the platforms in the context of this work are the interface and set of supportive tools. Via showcasing experience (Metsään.fi) and publishing legal and strategical basis for the data collaboration (CGIAR) it enables trust and demonstrates the value created**. The value is created on platforms through interactions and making data-based services available.** A strategic approach, involving the end-users and stakeholders who provide value is strongly connected to the success.

It is important to emphasize that **platforms do not create value by themselves**. They are an enabling tool for data governance, stakeholder's management and throughout the data-creation value chain – enabling the collection, storing, sharing and using the data. Platforms present an environment that allows to demonstrate the motivation, legal aspects and data governance principles to a wider public – and therefore build trust and transparency. **Platform's role is, therefore, amplifying the efforts and communicate them to stakeholders**.

Although the polar sample provided a good overview, exploring only three cases and two aspects within this complex topic can be considered one of the main **limitations of this study.** There are several topics and connections that this work was unable to address, due to time and space limitations and which provide an interesting basis for the **further research**.

While platforms enable better overcoming of several challenges proposed by the B2G data-sharing – missing structure of data sharing and lack of professionals in data by creating the enabling environment and concrete structure. Several challenges – like economic barriers and legislation remain to be solved in a wider scale. Private sector motivations to share the data specifically on the platforms, and how the available trust-building tools – such as different published statements and data governance guidelines affect their willingness to share data is one of the topics that would pose an interesting further research.

The effect on the previous experience and relations from the private sector data providers perspective would be another interesting topic to look into. One more important point to explore more specifically would be what are the means in the platform collaboration to overcome the missing preconditions – such as the previous relations among the stakeholders.

The rise in platforms and other forms of data collaboratives shows that there is a clear understanding of the value that lies in data and collaboration. This work provided some insights into the enablers of data collaboration platforms, but due to the depth of this topic, it requires more exploration. Generally, a deeper look into different dynamics of collaborative actions in the field of data is required to fully understand how to benefit the most from this complex asset.

LIST OF REFERENCES

Agar, J., 2003. *The Government Machine.* s.l.:MIT Press.

Alford, J. & O'Flynn, J., 2012. *Rethinking Public Service Delivery: Managing with External Providers.* s.l.:Palgrave Macmillan, Basingstoke.

Andrews, L., 2019. Public administration, public leadership and the construction of public value in the age of the algorithm and 'big data'. *Public Administration - Special Issue: Symposia: Leadership for Public Value and Integration, Functional Differentiation and Problem-Solving in Multilevel Governance,* 97(2), pp. 296-310.

Ansell, C. & Gash, A., 2008. Collaborative Governance in Theory and Practice. *Journal of Public Administration Research and Theory,* 18(8), pp. 543-571.

Ansell, C. & Gash, A., 2018. . Collaborative Platforms as a Governance Strategy. *Journal of Public Administration Research and Theory,* 28(1), p. 16–32.

Aulkemeier, F., Iacob, M.-E. & Hillegersberg, J., 2019. Platform-based collaboration in digital ecosystems. *Electronic Markets,* Volume 29, p. 597–608.

Aulkemeier, F. I. M.-E. &. v. H. J., 2017. An architectural perspective on service adoption: A platform design and the case of pluggable cross-border trade compliance in e-commerce.. *Journal of Organizational Computing and Electronic Commerce,* pp. 325-341.

Babbie, E., 2014. *The Basics of Social Research.* s.l.:Chapman University.

Benington, J. & Moore, M., 2011. *Public value: Theory and practice.* London: Palgrave Macmillan.

Bharosa, N. et al., 2013. Tapping into existing information flows: The transformation to compliance by design in business-to-government information exchange. *Government Information Quarterly Volume 30, Supplement 1,,* 30(1), pp. S9-S18.

Bigoni, M., Bracci, E. & Deidda Gagliardo, E., 2018. Measuring public value: a conceptual and applied contribution to the debate. *Public Money & Management,* 38(7), pp. 503-510.

Bolívar, R., Bwalya, M. & Reddick, K. J., 2019. Governance Models for Creating Public Value in Open Data Initiatives. *Public Administration and Information Technology.*

Carrera, W., Chan, W. S., Fischer, S. & Steenbergen, E., 2015. *Creating Value trough Open Data. A study on the Impact of Re-use of Public Data Resources,* s.l.: European Open Data Portal.

CCAFS, 2020. *Data Management Support Pack.* [Online]
Available at: https://ccafs.cgiar.org/data-management-support-pack
[Accessed 05 2020].

CGIAR Research Program on Climate Change, Agriculture and Food Security (CCAFS), 2015. *CCAFS - Data Management Strategy,* s.l.: CCAFS.

CGIAR, 2019. *About The Platform for Big Data in Agriculture.* [Online]
Available at: https://bigdata.cgiar.org/about-the-platform/
[Accessed 8 11 2019].

CGIAR, 2020. *About GARDIAN.* [Online]
Available at: https://gardian.bigdata.cgiar.org/about.php#!/
[Accessed 3 2020].

CGIAR, 2020. *About the platform.* [Online]

Available at: https://bigdata.cgiar.org/about-the-platform/

[Accessed 04 2020].

CGIAR, 2020. *Governance.* [Online]

Available at: https://www.cgiar.org/how-we-work/governance/

[Accessed 05 2020].

CIAT and IFPRI, 2016. *CGIAR Big data coordination platform.* [Online]

Available at:

https://cgspace.cgiar.org/bitstream/handle/10947/4450/2.%20Big%20Data%20platf

orm%20CGIAR%20Resubmission.pdf?sequence=1

Copenhagen Solutions Lab, 2020. *Who are we?.* [Online]

Available at: https://cphsolutionslab.dk/en/what-we-do

[Accessed 05 2020].

Cordey, C., Woods, M. & Collier, P., 2010. From value chain to value cycle: The role of risk

management and ICT. *SSRN Electronic Journal.*

Daglio, M., G. D. & Kitchen, H., 2014.. *Building Organisational Capacity for Public Sector*

Innovation. Paris, s.n.

de Bruijn, H. &. t. H. E., 2018. *Management in Networks.* s.l.:Routledge.

Dunleavy, P. & Margetts, H., 2015. *Design Principles for Essentially Digital Governance..* Sand

Fransisco, American Political Science Association, Annual Conference .

Dunleavy, P., Margetts, H., Bastow, S. & Tinkler, J., 2008. *Digital Era Governance: IT*

Corporations, The State, And E-Government. s.l.:Oxford University Press.

Eisenhardt, K. M., 1989. Building Theories from Case Study Research. *The Academy of*

Management Review, 14(4), pp. 523-550.

Emerson, K. & Nabatchi, T., 2015. *Collaborative Governance Regimes.* s.l.:Georgetown University Press.

Emerson, K., Nabatchi, T. & Balogh, S., 2012. An Integrative Framework for Collaborative Governance. *Journal of Public Administration Research and Theory,* 22(1), pp. 1-29.

European Commission, 2015. *Directive (EU) 2015/2366 of the European Parliament and of the Council of 25 November 2015 on payment services in the internal market,* s.l.: s.n.

European Commission, 2016. *DIRECTIVE OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on common rules for the internal market in electricity,* Brussels: s.n.

European Commission, 2019. *Guidance on private sector data sharing.* [Online]
Available at: https://ec.europa.eu/digital-single-market/en/guidance-private-sector-data-sharing
[Accessed 29 10 2019].

European Commission, 2020. *A European strategy for data.* [Online]
Available at: https://ec.europa.eu/info/sites/info/files/communication-european-strategy-data-19feb2020_en.pdf

European Commission, 2020. *Towards a European strategy on business-to-government data sharing for the public interest,* s.l.: s.n.

French, J. & Raven, B., 1959 . *The bases of social power.* s.l.:Univeristy of Michigan, Institute for Social Research.

Gandomi, A. & Haider, M., 2014. Beyondthehype:Bigdataconcepts,methods,andanalytics. *InternationalJournalofInformationManagement,* Volume 35, pp. 137-144.

Golafshani, N., 2003. Understanding Reliability and Validity in Qualitative Research. *The Qualittative Report,* 8(4), p. 12.

Goldsmith, S. & Eggers, W., 2004. *Governing by Network: The New Shape of the Public Sector.* s.l.:A Brookings Institution Press and Ash Center for Democratic Governance and Innovation .

GovLab, 2019. *Data Collaboratives Explorer.* [Online]

Available at: https://datacollaboratives.org/explorer.html

[Accessed 29 10 2019].

Grossman, R. L. H. A. M. M. P. M. &. W. W., 2016. A Case for Data Commons: Toward Data Science as a Service. *Computing in Science & Engineering,* 18(5), pp. 10-20.

Hautamäki A., O. K., 2018. Digital Platforms for Restructuring the Public Sector. *Collaborative Value Co-creation in the Platform Economy. Translational Systems Sciences,* Volume 11, pp. 91-108.

Hitatchi, 2016. *Hitachi Sets Data Free to Give Copenhagen Smart City Insights.* [Online]

Available at: https://www.hitachivantara.com/en-us/news-resources/press-releases/2016/gl160518.html

IDC Italia srl and The Lisbon Council, 2019. *D2.6 Second interim report. The European data market monitoring tool: key facts and figures, policy conclusions, data landscape and quantified stories,* s.l.: IDC Italia srl and The Lisbon Council.

Indeed Hiring Lab, 2019. *Data Scientist: A Hot Job That Pays Well, Indeed Hiring Lab.* [Online] Available at: https://www.hiringlab.org/2019/01/17/data-scientist-job-outlook

International Data Corporation (IDC) and the Lisbon Council , 2016. *Opening Up Private Data for Public Interest.* [Online]

Available at:

http://datalandscape.eu/sites/default/files/report/Story_1_New_format.pdf

[Accessed 05 2020].

Jarman, H. & Luna-Reyes, L., 2016. *Private Data and Public Value Governance, Green Consumption, and Sustainable Supply Chains.* s.l.:Springer International Publishing.

Jørgensen, T. & Bozeman, B., 2017. Public Values: An Inventory. *Administration & Society,* 39(3), pp. 354-381.

Kelly, G. & Muers, S., 2002. *Creating Public Value—An Analytical Frameworkfor Public Service Reform. ,* s.l.: London: Cabinet Office Strategy Unit...

Kenney, M. & Zysman, J., 2016. The Rise of the Platform Economy. *Issues in science and technology,* 31(3), pp. 61-69.

Ketter, W., Peters, M., Collins, J. & Gupta, A., 2015. Competitive Benchmarking: An IS Research Approach to Address Wicked Problems with Big Data and Analytic. *Social Science Research Network..*

Kettl, D., 1993. *Sharing Power: Public Governance and Private Markets.* s.l.:Brookings Institution Press.

Klievink, B., 2015. *Understanding Public-Private Collaboration Configurations for International Information Infrastructures.* s.l., s.n., pp. 170-180.

Klievink, B., Bharosa, N. & Tan, Y.-H., 2016. The collaborative realization of public values and business goals: Governance and infrastructure of public–private information platforms.. *Government Information Quarterly, 33(1),,* p. 67–79..

Klievink, B., Van Der Voort, H. & Veenman, W., 2018. Creating value through data collaboratives: Balancing innovation and control. *Information Policy,* 23(4), pp. 1-19.

Lachman, B. E., 2000. *Public-Private Partnerships for Data Sharing: A Dynamic Environment.* s.l., Rand.

Lim, C. et al., 2018. From data to value: A nine-factor framework for data-based value creationin information-intensive services. *International Journal of Information Management,* Volume 39, pp. 121-135.

Linder, S. H., 1999. Coming to Terms With the Public-Private Partnership: A Grammar of Multiple Meanings. *. American Behavioral Scientist, 43(1),* p. 35–51.

Magi, M. & Maruping, L., 2019. Unleashing Innovation With Collaboration Platforms. *MIT Sloan Management Review.*

Manyika, J., M. Chui, B., Brown, J. & Bughin, R. D. C. R. a. A. B., 2011. *Big Data: The next frontier for innovation, competition, and productivity.* [Online]
Available at: https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/big-data-the-next-frontier-for-innovation
[Accessed 11 2019].

Margetts, H. &. N. A., 2017. Government as a Platform: What can Estonia Show the World?.

Markus, M. & Bui, N., 2012. Going Concerns: The Governance of Interorganizational Coordination Hubs. *Journal of Management Information Systems ,* 28(4), pp. 163-197.

Maru, A. et al., 2018. *Digital and Data-Driven Agriculture: Harnessing the Power of Data for Smallholders.* Rome, Global Forum on Agricultural Research and Innovation.

McAfee, A. & Brynjolfsson, E., 2012. Big Data: The Management Revolution. *Harvard Business Review*, October.

McKinsey Global Institute, 2011. *Big data: The next frontier for innovation, competition, and productivity,* s.l.: McKinsey Digital.

McKinsey, 2019. *Collaborating for the common good.* [Online]
Available at: https://www.mckinsey.com/business-functions/mckinsey-

analytics/our-insights/collaborating-for-the-common-good

[Accessed 12 2019].

Metsäkeskus, 2020. *Collecting Forest Resources Data from Privately Owned Forests in*

Finland. [Online]

Available at:

https://metsakeskus.maps.arcgis.com/apps/MapSeries/index.html?appid=e5a7453e

717544c6b8a377afd900864f

[Accessed 05 2020].

Moore, M., 1997. *Creating Public Value: Strategic Management in Government.* s.l.:Harvard

University Press.

Municipality of Copenhagen and Capital Region of Denmark, 2018. *CITY DATA EXCHANGE –*

Lessons learned from a public/private data collaboration. [Online]

Available at: https://cphsolutionslab.dk/content/2-what-we-do/3-data-platforms/3-

city-data-exchange/1-learnings-from-the-city-data-exchange-project/city-data-

exchange-cde-lessons-learned-from-a-public-private-data-

collaboration.pdf?1527149474

Nambisan, S., 20019. Platforms for Collaboration by Satish Nambisan. *Emergency*

Management | Innovation. Stanford Social Innovation Review, p. 44–49.

O'Reilly, T., 2010. Government as a Platfrom.. *MIT Press Journals, 6(1)..*

OECD, 2014. *OECD Science, Technology and Industry Outlook 2014.,* s.l.: OECD Publishing.

OECD, 2019. *The Path to Becoming a Data-Driven Public Sector.* [Online]

Available at: https://www.oecd-ilibrary.org/governance/the-path-to-becoming-a-

data-driven-public-sector_b45f4d51-en

Opher, A. C. A. &. O. A., 2016. *The Rise of the Data Economy: Driving Value through Internet of Things Data Monetization [Insights]..,* s.l.: IBM Corporation.

Osborne, D. & Gaebler, T., 1992. *Reinventing Government: How the Entrepreneurial Spirit is Transforming the Public Sector.* New York: Penguin Books.

Perkmann, M. &. S. H., 2015. Open data partnerships between firms and universities: The role of boundary organizations.. *Research Policy, 44(5),* p. 1133–1143.

Premkumar, G., Ramamurthy, K. & Saunders, C. S., 2005. Information Processing View of Organizations: an Exploratory Examination of Fit in the Context of Interorganizational Relationships. *Journal of Management Information Systems, 22(1),* pp. 257-294.

PwC, 2007. *The road ahead for public service delivery. Delivering on the customer promise,* s.l.: PwC.

Rhodes, R., 1997. *Understanding governance.* Buckingham: Open University Press.

Ritchie, J., 2013. *Qualitative Research Practice: A Guide for Social Science Students and Researchers (Second edition)..* s.l.: SAGE Publications Ltd.

Rittel, H. W. a. W. M. M., 1973. Dilemmas in a general theory of planning. *Policy Sciences, 4(2),* p. 155–169.

Ruijer, E. &. M. A., 2019. Open Government Data as an Innovation Process: Lessons from a Living Lab Experiment.. *Public Performance & Management Review,* p. 1–23.

Salomon, L., 1981. Rethinking public management: third-party government and the changing forms of government action. *Public Policy,* 29 (3), p. 255–75..

Salomon, L., 2002. *The Tools of Government: A Guide to the New Governance.* Oxford: Oxford Univeristy Press.

Sherer, S. & Adams, B., 2001. Collaborative commerce: The role of intermediaries in e-collaboration. *Journal of Electronic Commerce Research,* pp. 66-77.

Statista, 2019. *The 100 largest companies in the world by market value in 2019.* [Online]
Available at: https://www.statista.com/statistics/263264/top-companies-in-the-world-by-market-value/

Sullivan, H. & Skelcher, C., 2002. Working Across Boundaries: Collaboration in Public Services. *https://www.researchgate.net/journal/0031-2290_Parliamentary_Affairs,* p. 271.

Susha, I., Janssen, M. & Verhulst, S., 2017. *Data Collaboratives as a New Frontier of Cross-Sector Partnerships in the Age of Open Data: Taxonomy Development.* s.l., s.n.

Susha, I. et al., 2018. A Research Roadmap to Advance Data Collaboratives Practice as a Novel Research Direction. *International Journal of Electronic Government Research,* 14(3), pp. 1-11.

Talbot, 2008. *Measuring Public Value - a competing values approach.* s.l.:Work Foundation.

Talbot, C., 2011. Paradoxes and prospects of 'public value'. *Public Money & Management,* 31(1), pp. 27-34.

The Economist, 2017. The world's most valuable resource is no longer oil, but data.. *The Economist*.

The World Bank Group, 2017. *INTERNET OFTHINGS The New Government to Business Platfrom - A REVIEW OF OPPORTUNITIES, PRACTICES, AND CHALLENGES,* Washington: The World Bank Group.

United Nations, 2014. *A World That Counts: Mobilising the Data Revolution for Sustainable Development.* [Online]
Available at: http://www. undatarevolution.org/wp-content/uploads/2014/11/A-

World-That-Counts.pdf

[Accessed 29 10 2019].

Valone, M. et al., 2019. *Finland's model in utilising forest data,* Helisnki: PTT.

Walsh, T., 2018. *2062: The World that AI Made..* s.l.:Black Inc.

van Heck, E. & Vervest, P., 2007. Smart business networks: How the network wins..

Communications of the ACM, 50(6), pp. 28-37.

van Hillegersberg, J., Moonen, H. & Dalmolen, S., 2012. Coordination as a service to enable

agile business networks.. *The dynamics of global sourcing. Perspectives and practices

,* p. 164–174.

van Ooijen, C., Ubaldi, B. & & Welby, B., 2019. *A Data-driven Public Sector - OECD Working

Paper - OECD.* [Online]

Available at: https://www.oecd.org/gov/digital-government/working-paper-a-data-

driven-public-sector.htm

Wang, H.-J. &. L. J., 2016. Adoption of open government data among government agencies.

Government Information Quarterly, 33(1), p. 80–88.

Verhulst, S. G., Young, A., Winowatan, M. & Zahuranec, A. J., 2019. *Leveraging Private Data

for Public Good: A Descriptive Analysis and Typology of Existing Practices.* [Online]

Available at: https://datacollaboratives.org/static/files/existing-practices-report.pdf

[Accessed 02 2020].

Verhults, S. & Young, A., 2017. *Five ways data collaboratives create public value.* [Online]

Available at: http:// datacollaboratives.org/introduction.html#section5/5c

[Accessed 9 11 2019].

Wilkinson, M. D. D. M. A. I. J. A. G. A. M. B. A. B. N. B. J.-W. S. L. B. d. S. B. P. E. B. J. B. A. J. C. T. C. M. D. I. D. O. E. S. E. C. T., 2016). The FAIR Guiding Principles for scientific data management and stewardship.. *Scientific Data, 3(1),* p. 1–9.

World Economic Forum, & McKinsey & Company, 2019. *Data Collaboration for the Common Good - Enabling Trust and Innovation Through Public-Private Partnerships [Insight Report]..* [Online]
Available at:
http://www3.weforum.org/docs/WEF_Data_Collaboration_for_the_Common_good-pdf

World Economic Forum, 2015. *Global Agenda Council on the Future of Software & Society. Deep Shift 21 Ways Software Will Transform Global Society,* s.l.: World Economic Forum.

Worthy, B., 2015. The Impact of Open Data in the UK: Complex, Unpredictable, and Political.. *Public Administration, 93..*

Yin, R., 2003. *Case Study Research Design and Methods.* s.l.:SAGE Publications.

Yin, R. K., 2008. *Case Study Research: Design and Methods (4th edition)..* s.l.:SAGE Publications, Inc..

APPENDICES

Appendix 1 – Interview questions

Collaborative platform enabling aspects

1. **What are the objectives of the platform are translated to measurable goals and/or strategic plan how to achieve them?**

2. **Who are the stakeholders of the platform what is the value proposition of the platform for them?** Has there been any previous collaboration present and do you see this effecting the collaboration?

3. **How do you assess the importance of trust building within the stakeholders?**

4. **Was there prior collaboration with the stakeholders? What is the level of stakeholders' involvement into platform (service) development?** What are the means of involving the private sector? To what extent are the different stakeholders represented in the decision-making?

5. **Could the platform be defined as collaboration enabler?** Are there any (communication) tools or other means that are enabling interactions between the stakeholders on platform?

6. **How is good (ethical and secure) data governance insured on the platform?** Are there common principles used (based on the legislation) or in-house strategy developed? What are the main challenges that have been faced in implementing good data governance?

Platform framework as an enabler throughout 4-steps Data Value Cycle:
1. collecting 2. storing/processing 3. sharing/publishing 4. using/reusing

1. **What are the data collection methods used?** Is the data collection carried out through the platform? Are there any tools on the platform supporting the process?

2. **Are the data processing capabilities integrated into the platform** or provided by external tools/providers?

3. **What outputs for sharing data are enabled by the platform** - services, visualization, anything else?

4. **Data using/reusing.** Is there any additional usage for gathered data? Is it shared with other national or international platforms/service providers? Does it provide input to policy making? Anything else?
    How is the usage of the data promoted on the platform - Are there showcases of the data-based value creation presented? Any support provided on the platform to support the business cases built on the data provided?

5. **Do you see any other means how platform enables delivering the objectives and creating value?**

6. **Are there any particular challenges that platform framework provides (either technical or organizational)?**

Appendix 2 – List of Interviewees

Institution	Time and place of the interview
Interview 1	
Representative of the City of Copenhagen	Interview carried out via phone on May 4th 2020 13:00-14:00
Interview 2	
Representative of the Metsään.fi (Finnish Forest Centre)	Interview carried out via Google Meet on May 4th 2020 10:30-11:30
Interview 3	
Representative of the CGIAR	Interview carried out via Google Meet on July 14th 2020 11:00-11:35

Appendix 3 - Non-exclusive licence

A non-exclusive license for reproduction and for granting public access to the graduation thesis[1]

I, Sandra Metsis (*date of birth*: May 6, 1988)

1. Give Tallinn University of Technology a permission (non-exclusive license) to use free of charge my creation CROSS-SECTORAL DATA COLLABORATION PLATFORMS supervised by Veiko Lember, PhD.

1.1. to reproduce with the purpose of keeping and publishing electronically, including for the purpose of supplementing the digital collection of TalTech library until the copyright expires;

1.2. to make available to the public through the web environment of Tallinn University of Technology, including through the digital collection of TalTech library until the copyright expires.

2. I am aware that the author will also retain the rights provided in Section 1.

3. I confirm that by granting the non-exclusive license no infringement is committed to the third persons' intellectual property rights or to the rights arising from the personal data protection act and other legislation.

---

[1] *The non-exclusive license is not valid during the access restriction period with the exception of the right of the university to reproduce the graduation thesis only for the purposes of preservation.*