Government Information Quarterly 40 (2023) 101763

Contents lists available at ScienceDirect

Government Information Quarterly

journal homepage: www.elsevier.com/locate/govinf

An ecosystem perspective on developing data collaboratives for addressing societal issues: The role of conveners

Iryna Susha [a,b,*], Tijs van den Broek [c], Anne-Fleur van Veenstra [d], Johan Linåker [e]

[a] Utrecht University, Copernicus Institute of Sustainable Development, Faculty of Geosciences, Princetonlaan 8a, 3584 CB Utrecht, the Netherlands
[b] Örebro University, Department of Informatics, School of Business, Örebro 701 82, Sweden
[c] Vrije Universiteit Amsterdam, Department of Organization Sciences, De Boelelaan 1105, 1082 VA Amsterdam, the Netherlands
[d] Netherlands Organisation For Applied Scientific Research TNO, Anna van Buerenplein 1, 2595 DA The Hague, the Netherlands
[e] Lund University, Department of Computer Science, Ole Römers Väg 3, Lund, Sweden

ARTICLE INFO

Keywords:
Data collaborative
Data ecosystem
Convener
Innovation

ABSTRACT

With the open and big data movement in full swing, data sharing becomes more ubiquitous and more often crosses sectoral boundaries. The promise of data to help address societal issues and foster innovation requires public organizations to work together with businesses and researchers. Data collaboratives whereby actors collaborate to share and use data for public good gain increasing interest. Most of these collaborations, however, tend to be one-off, small, and limited in impact due to a complex web of legal, technical, ethical, commercial, and organizational challenges. Initiators of data collaboratives, termed as conveners, can potentially alleviate some of these concerns by playing various roles in developing a more sustainable data ecosystem for the data collaboratives. Our study investigates what convener roles are perceived to be critical in developing data collaboratives. By drawing on data ecosystems thinking, we developed a framework of convener roles and sub-roles which we further used to analyze four cases in the Netherlands and Sweden. We conclude that connecting role and learning catalyst role are critical at the initiation stage, while stimulating and mediating roles emerge as future critical roles as the data ecosystem develops. We further identified convener meta-roles that are associated with particular data ecosystem structures (keystone-centric, marketplace-based, intermediary-based, and platform-centric). Our research can be instrumental to actors leading the efforts of creating such data ecosystems as it provides insights on the needs and resources that can be leveraged to stimulate development and innovation.

1. Introduction

With the open and big data movement in full swing, data sharing becomes more ubiquitous and more often crosses sectoral boundaries. The promise of data to help address societal issues and foster innovation requires public organizations to work together with businesses and researchers (de Juana-Espinosa & Luján-Mora, 2019; Gupta, Mejia, & Kajikawa, 2019; van den Broek & van Veenstra, 2018). Such data partnerships, also termed as data collaboratives, whereby actors collaborate to share and use data for public good, have lately attracted increasing interest among innovation scholars.

Data collaborative is a concept which emerged in 2015 (Verhulst & Sangokoya, 2015) and attracted academic attention as a new frontier of data-driven innovation research (Susha, Janssen, & Verhulst, 2017a). The idea of a data collaborative rests on the premise that sharing and reusing data across different sectors, especially untapped sources of private sector data, has great potential for developing innovations to address a variety of public problems. Data collaboratives have been defined as "cross-sector collaboration initiatives aimed at data collection, sharing, or processing for the purpose of addressing a societal challenge" (Ibid., p.2691). One of the key characteristics of a data collaborative is its cross-sectoral nature of collaboration which is underpinned by the idea that many societal issues can be tackled more effectively if actors from different sectors pull their resources, data, and data expertise together to develop social innovations.

The repository of cases of data collaboratives[1] around the world

* Corresponding author at: Utrecht University, Copernicus Institute of Sustainable Development, Faculty of Geosciences, Princetonlaan 8a, 3584 CB Utrecht, the Netherlands
E-mail address: i.susha@uu.nl (I. Susha).

[1] https://datacollaboratives.org/explorer.html

https://doi.org/10.1016/j.giq.2022.101763
Received 5 November 2021; Received in revised form 7 July 2022; Accepted 18 September 2022
Available online 14 October 2022

compiled by The Gov Lab at New York University has grown from 23 cases in 2016 to over 200 in 2021. This shows steadily growing interest to and experimentation with this novel approach to partnering in the age of data. Most of these collaborations, however, tend to be one-off, small, and limited in impact (World Economic Forum, 2019). This could be attributed to a complex web of legal, technical, ethical, commercial, and social risks associated with data collaboratives and to the lack of trust among individuals, institutions, and governments when it comes to data sharing (Ibid.). For example, the functioning of a data collaborative depends on successful sharing of required data. Yet in practice data owners, both public and private, are often hesitant to share their data for various reasons (van den Broek & van Veenstra, 2018). Organizations find it increasingly difficult to balance the benefits and risks that such data sharing poses.

Initiators of data collaboratives, termed as conveners, can potentially alleviate some of these concerns by playing various facilitating roles in developing a more sustainable data ecosystem for the data collaboratives. Conveners can provide a data infrastructure for sharing data and an organizational structure for collaboration to diverse stakeholders (Gupta, Panagiotopoulos, & Bowen, 2020). They can also connect supply of data and data expertise with demand and with concrete issues to be solved (Bonina & Eaton, 2020). For instance, World Bank provides a platform[2] that offers centralized IT architecture, legal assistance, partnership management support, multidisciplinary expertise, and a repository of derived data products to organizations engaging in data collaboratives. Similar kinds of initiatives are emerging locally as well. Recently, Gupta et al. (2020) demonstrated in a case study of the London's smart city initiative how a convener can play an important role by orchestrating the accumulation and diffusion of resources and capabilities in a data collaborative.

However, for conveners it is not always clear which role they should assume at which point in time to stimulate the development of the data collaborative in the most effective way. Data collaboratives as a new phenomenon might require a particular approach to convening. In contrast to open data initiatives, data collaboratives are characterized by the voluntary nature of data sharing, are not limited to government data, and may be more closed in terms of data and collaboration itself (Susha et al., 2017a). The uniqueness of data collaboratives is manifested in how public and private interests are combined to create both value and tension (Klievink, Van Der Voort, & Veeneman, 2018). The tensions occur because the various goals of participating organizations might clash, and the organizations need to balance between control over the data and the degree of openness needed to create value in the collaboration (Ibid.). These tensions lead to additional coordination challenges for the convener, such as aligning incentives of proprietary data holders and the goals of (public) problem owners and negotiating the issues of data re-use and control (Susha, Janssen, & Verhulst, 2017b). Hence, our research poses the following research question: **What convener roles are perceived to be critical in developing data collaboratives by the different actors?**

Our intended contribution is two-fold. *First*, existing literature on data collaboratives does not provide substantial insights regarding the roles that conveners of data collaboratives can assume. Literature on open data initiatives can provide some answers regarding how such initiatives can be set up and developed (Bonina & Eaton, 2020), yet its applicability to data collaboratives is limited due to the aforesaid reasons. Therefore, we need to adopt a broader theoretical perspective – drawing on data ecosystems thinking and applying insights from innovation ecosystems literature – in order to develop a framework of convener roles applicable to data collaboratives. *Second*, we aim to make a contribution by investigating in more detail how the criticality of convener roles varies in data collaboratives. As mentioned earlier in the Introduction, data collaboratives have so far been mostly implemented as one-off initiatives limited in scale, thus it is often challenging to sustain and scale up these projects (European Commission, 2020; Susha & Gil-Garcia, 2019). This points to the relevance of investigating how conveners of data collaboratives might help move these projects beyond the initial phase. Convener roles needed for setting up data collaboratives may not be sufficient for scaling them up. From innovation ecosystems research it is known that some convener roles are more critical at the initiation stage and other – at a later development stage (Dedehayir, Mäkinen, & Ortt, 2018). In our research we investigate this further in the case of data collaboratives.

This paper is structured as follows. In section 3 we discuss the concept of data collaboratives from an ecosystem perspective and review relevant research around it. In section 4 we present our theoretical framework of the roles that conveners can play in a data collaborative. In section 5 follows the description of our multi-case study research design and in section 6 we report on the findings from four case studies. Sections 6 and 7 contain our comparative analysis and reflections about the cases and the contributions that our research makes to the topic.

2. Theoretical framework

We approach our research question from a data ecosystem perspective (Gupta et al., 2020) since it offers a holistic view of actors and their interactions in a data collaborative. We thus use the data ecosystem perspective as a theoretical lens for analyzing the phenomenon of data collaboratives.

An ecosystem is "an evolving, self-organizing system of feedback and adjustment among actors and processes" (Dawes, Vidiasova, & Parkhimovich, 2016, p.6). Typically, an ecosystem has certain properties, such as cyclical nature, dynamism, evolution, sustainability, demand-supply relationship, and embeddedness in a local context (Harrison, Pardo, & Cook, 2012). Oliveira, Barros Lima, and Farias Lóscio (2019) proposed an integrated definition of a data ecosystem: "a loose set of interacting actors that directly or indirectly consume, produce, or provide data and other related resources (e.g., software, services, and infrastructure); each actor performs one or more roles and is connected to other actors through relationships, in such a way that actors by collaborating and competing with each other promote data ecosystems". Heimstädt, Saunderson, and Heath (2014) further point out the circular flow of resources and dependencies between suppliers, intermediaries, and users as properties of data ecosystems.

Data ecosystems can emerge in the scientific, governmental, and/or industry domains and focus on innovation, interaction, and/or transaction among ecosystem members (Gelhaar, Groß, & Otto, 2021). Research on data ecosystems has been fragmented and has mostly focused on open data ecosystems. Research describes (open) data ecosystems using a number of elements, such as participants and their roles, data resources and tools, design, context, and interdependencies and interactions (Oliveira & Lóscio, 2018; Zuiderwijk, Janssen, & Davis, 2014). Recently data ecosystems have been studied that are intended to collect and process large and heterogeneous datasets, termed as big data, pointing at its distinct elements (Lnenicka & Komarkova, 2019), design mechanisms (Modgil, Gupta, Sivarajah, & Bhushan, 2021), and governance challenges (van den Broek & van Veenstra, 2018) of such ecosystems.

In an (open) data ecosystem the actors play the following main roles: data suppliers, data intermediaries, and data consumers (Heimstädt et al., 2014). This view holds true also for data collaboratives as the minimal value chain therein is also about matching data supply and data demand (Susha et al., 2017a). Additionally, some authors suggest that there is an additional role of "a central actor taking the initiative within networked systems organized to achieve a specific goal related to innovation" (Harrison et al., 2012). The actor who initiates and coordinates the development of the data ecosystem can be termed as **convener** of the data ecosystem. A convener can be defined as "a catalytic agent bridging unaware, unsure or skeptical actors to explore the

[2] https://datapartnership.org

possibilities of collaboration" (Kalegaonkar & Brown, 2000, p.9). The term convener is widely used in studies of inter-organizational collaboration. In the literature we also find other terms that capture the function of initiating and developing the data ecosystem: intermediaries, leaders, bridging agents, champions, catalysts, stimulators etc. We choose to use the term convener in our study because we are specifically interested in actors who initiate and coordinate the development of a data ecosystem. Yet below we also review research discussing the activities of convening but using other terms.

The role of convener of data ecosystems is deeply embedded in strategic ecosystem thinking (Harrison et al., 2012) The strategic ecosystem thinking entails a planned approach to data ecosystem development. It requires identifying the main actors who form essential components of the ecosystem, understanding the nature of transactions between them, recognizing what resources are needed by each actor, and observing the indicators signaling the health of the ecosystem (Harrison et al., 2012). Furthermore, such strategic data ecosystem thinking is framed by three concerns: intentionality, value creation, and sustainability (Ibid.). We thus adopt the assumption that data ecosystems can be intentionally created, developed, and managed for the purpose of achieving a managerial and policy vision (Mercado-Lara & Gil-Garcia, 2014). For instance, taking open data ecosystems as an example, since 2009 governments around the world engaged in cultivating open data ecosystems as part of the open government discourse. The goal of the government-led open data programs is creation of economic and social value, yet the benefits are often not easy to pinpoint and measure and vary across the actors of the ecosystem. This further complicates the issue of whether the ecosystem should be further developed or discontinued. In the open data literature, the convener role is typically attributed to government organizations (Martin, Turki, & Renault, 2017). In data collaboratives this is not always the case, therefore we adopt a broader view. Research on open data ecosystems further detailed the convener roles (e.g., Martin et al., 2017), yet these roles are predominantly focused on stimulating data reuse, whereas in data collaboratives the convener is also tasked with stimulating voluntary data sharing (Susha, Rukanova, Ramon Gil-Garcia, Hua Tan, & Gasco, 2019) and aligning the goals and interests of the different organizations (Ruijer, 2021). Therefore, we look beyond open data research and seek insights from broader research fields, such as business and innovation ecosystems, collaborative innovation literature, and partnerships research.

2.1. Convener roles in data ecosystems

To detail which roles conveners of data ecosystems can play, as a starting point we build on the framework of Agogué et al. (2017) who formulated four core roles that conveners play in innovation ecosystems: (i) connecting actors; (ii) involving, committing, and mobilizing actors; (iii) solving, avoiding, or mitigating potential conflicts of interests; and (iv) (actively) stimulating the innovation process and innovation outcomes. As this categorization is rather generic, we elaborated it by adding insights from other aforesaid literature streams, including open data research. For instance, in collaborative innovation and partnerships literature we found sub-roles that were not explicitly articulated in innovation ecosystems research (S3 – creating urgency, S4 – ensuring openness). As a result, we developed a framework (Table 1) detailing the roles and sub-roles of conveners in the context of data collaboratives. In the following we discuss the identified roles and sub-roles in more detail.

Connecting role is one of the most cited roles across the different literature streams. In this category we grouped together the connect and involve roles from the categorization by Agogué et al. (2017) since other literature sources mentioned these activities alongside each other. Thus, this role is about connecting, mobilizing, and involving relevant actors of the ecosystem (C1) (Dedehayir et al., 2018; Agogué et al., 2013; Ruijer, 2021). Additionally, the activity of matching the demand and supply in the ecosystem is mentioned (C2). For instance, Holzmann et al. (2014) discuss the role of innovation intermediaries in multi-sided markets, whereby innovation intermediaries are tasked with the matching process between demand and supply.

Table 1
Framework of convener roles and sub-roles in facilitating data ecosystems.

ROLE		SUB-ROLES	REFERENCES
Connecting role	C1	Connecting actors, mobilizing and committing actors	Iansiti and Levien (2004), Agogué, Yström, and Le Masson (2013); Dedehayir et al., 2018); Bonina and Eaton (2020); Ruijer (2021)
	C2	Matching demand and supply	Susha (2020); Holzmann, Sailer, and Katzy (2014); Dedehayir et al. (2018); Bonina and Eaton (2020)
Mediating role	M1	Mitigating potential conflicts of interest, building trust, managing expectations, crafting a shared vision	Agogué et al. (2013), Gupta et al. (2020); Dedehayir et al. (2018); Dorado and Vaz (2003); Bakici, Almirall, and Wareham (2013); Ruijer (2021)
	M2	Providing governance tools and coordinating structures	Gupta et al. (2020); Dedehayir et al. (2018)
Stimulating role	S1	Stimulating the innovation process and outcomes	Agogué et al. (2013); Dedehayir et al. (2018)
	S2	Identifying sources of advice, support, and funding for the innovation process	Lee (2014)
	S3	Creating the urgency for innovation and providing inspiration	Crosby, Hart, and Torfing (2017)
	S4	Ensuring openness and transparency in the data ecosystem	Gupta et al. (2020); Ruijer (2021)
Learning catalyst	L1	Providing structure or space for joint knowledge creation and mobilization of a collective body	Elmquist, Ollila, and Yström (2016), Agogué et al. (2013)
	L2	Providing research and expertise on the issue, providing best practices and resources for collaboration	Stadtler and Probst (2012)
	L3	Legitimizing and diffusing knowledge	Gupta et al. (2020)
Infrastructure provision	I1	Provide an initial set of resources (platform, tools, expertise)	Iansiti and Levien (2004); Dedehayir et al. (2018); Bonina and Eaton (2020)
	I2	Leverage capabilities across the network	Iansiti and Levien (2004); Dedehayir et al. (2018)
	I3	Making the creation of products/services by third parties more efficient	Iansiti and Levien (2004); Dedehayir et al. (2018)

Mediating role is realized by conveners mitigating potential conflicts of interest, building trust, (Dorado & Vaz, 2003), managing expectations (Ruijer, 2021), and crafting a shared vision, including across the government silos (M1). Also, providing coordinating structures and governance tools is mentioned in the literature (M2) (Dedehayir et al., 2018; Gupta et al., 2020). The recent literature on data ecosystem governance puts much emphasis on these activities and their crucial influence on the development of data ecosystems, without necessarily mentioning the role of convener (Lis & Otto, 2020; Lis & Otto, 2021; van den Broek & van Veenstra, 2018).

Stimulating role is concerned with stimulating the innovation process and outcomes (S1) (Agogué et al., 2013; Dedehayir et al., 2018) and identifying sources of advice, support, and funding (S2) (Lee, 2014). Additionally, we find insights from other relevant literature which also discusses creating the urgency or momentum for innovation and mobilizing the necessary organizational resources (S3) (Crosby et al., 2017)

and ensuring openness and transparency in the data ecosystem (S4) (Gupta et al., 2020; Ruijer, 2021).

In addition to the three roles above, we find one more role discussed in the literature on public sector innovation and public-private partnerships. *Learning catalyst role* is about providing expertise and resources to the parties to define the problem, as well as formulating best practices and offering trainings (L2) (Stadtler & Probst, 2012). Research by Agogué et al. (2013) further delves into the thesis that conveners can also take on a more active role, which they described as "an architect of collective exploration". In this sub-role the convener provides structure for collective exploration of complex issues (L1) and space for joint knowledge creation through mobilization of a collective body (L2) (Elmquist et al., 2016). Additionally, the sub-role of legitimizing and diffusing knowledge in the ecosystem is mentioned (L3) (Gupta et al., 2020).

Furthermore, one more role is discussed in the open data literature and more broadly business ecosystems literature. *Infrastructure provision* role means providing an initial set of resources, such as a platform, tools, expertise for resource exchange among parties, for the overall functioning and development of the ecosystem (I1) (Dedehayir et al., 2018; Iansiti & Levien, 2004). This role leverages important resources and capabilities across the network (I2). In this role conveners simplify the task of connecting network participants to each other thereby making the creation of new products or services by third parties more efficient (I3) (Ibid). Bonina and Eaton (2020) in their study of open data ecosystems investigated several types of resources that conveners can provide, namely informational tools, software tools, contractual rules, and social tools (e.g., networking events).

2.2. Criticality of convener roles in data ecosystems

The initiation and development of data ecosystems can follow different pathways; therefore, it is safe to presume that not all roles and sub-roles are critical at all times. From innovation ecosystems research it is known that some convener roles are more critical at the initiation stage and other – at a later development stage. For instance, Dedehayir et al. (2018) argue that at the initiation stage of particular importance is for the convener to forge partnerships between actors in the network and to set up ecosystem governance coordinating interactions and resource flows between the actors. In our research we investigate this further in the case of data collaboratives.

Furthermore, the criticality of convener roles can depend on the structure of the ecosystem. Data ecosystem structure refers to the different kinds of relationships, interaction, and organization of the ecosystem actors (Gelhaar et al., 2021). Dedehayir et al. (2018) suggest that the way actors are organized in a data ecosystem has implications for the activities that the convener should undertake. Lis and Otto (2020) confirm this by saying that the positions of organizations in the data ecosystems and the kind of interdependences between them influence the kind of actions required to govern the data ecosystem (which is one of the convener roles). In our research we aim to test these initial insights from the literature empirically in our case studies.

In the literature we find different approaches to conceptualize and categorize data ecosystem structures. Several data ecosystem taxonomies have been recently proposed (Curry & Sheth, 2018; Gelhaar et al., 2021; Lis & Otto, 2021; Oliveira et al., 2019) which all include data ecosystem structure as one of the dimensions. Some authors distinguish different ecosystem structures based on how control over data is distributed in the data ecosystem, e.g. if there is a keystone actor on whom all other actors depend. Namely, Lis and Otto (2021) describe four archetypical structures, such as hierarchy, market, network, and bazaar; this line of argument is rooted in theories of inter-organizational organizing, as for instance applied in van den Broek and van Veenstra (2018) and in Susha, Janssen and Verhulst (2017b). Since at early stages of development of data collaboratives issues of control are still being negotiated, it can be challenging to distinguish between these types. Therefore, we turn to another set of frameworks which conceptualize structure through *interdependencies between actors in the exchange of data in the ecosystem*. Oliveira et al. (2019) and Gelhaar et al. (2021) develop similar categorizations following this line of thinking and propose the following data ecosystem structures: keystone-centric, intermediary-based, platform-centric, and marketplace-based. These structures are not mutually exclusive and 'hybrid' versions are possible (for instance, a platform can also include data intermediaries). Yet, the main idea of the four archetypes is to show what the actors depend on in the flow of data in the ecosystem. We explain the four types in more detail below.

First, a *keystone-centric structure* is a structure in which stakeholders are organized around a focal stakeholder (Oliveira et al., 2019). In a keystone-centric structure, the central actor is responsible for providing much of the data in the data ecosystem (Gelhaar et al., 2021). For example, the central government may act as a keystone stakeholder providing government data that is used by multiple data intermediaries and data users (Dawes et al., 2016; Koster & Suarez, 2016; Koznov et al., 2016; Lee, 2014). The central actor in a keystone structure has no complete control over the other actors in the data ecosystem, e.g., members may freely join or leave the data ecosystem.

Second, in a *platform-centric structure*, a shared platform defines the organizational structure of a data ecosystem (Oliveira et al., 2019). A platform's responsibility is to provide an infrastructure to support the provision and consumption of data. The main rationale of a platform is to reduce the costs of providing data by minimizing transaction costs between actors and minimizing interoperability and usability problems between actors (Gelhaar et al., 2021). Platforms may offer data services to the actors and allow third-party developers to create and offer new services (Zeleti & Ojo, 2016).

Third, an *intermediary-based structure* is a structure that depends on the presence of data intermediaries in order to generate value from data (Oliveira et al., 2019). A data intermediary is an actor that does not provide data directly but facilitates the use of data for other actors. In an intermediary-based structure, data providers and data users are organized around these intermediaries, that can be domain specific.

Last, a *marketplace-based structure* is a structure that provides the required infrastructure, business models, rules, and services for transactions of data and software between actors (Gelhaar et al., 2021). In general, marketplaces include a technical platform which connects data providers and data users, and they can also enable the sale of data, services, and applications (Oliveira et al., 2019).

In our research we aim to investigate the criticality of convener roles in connection with these different data ecosystem structures.

3. Method

Our research was designed as an embedded multiple case study (Yin, 2018) conducted according to the qualitative research tradition based on interpretive assumptions of the world (Klein & Myers, 1999; Walsham, 1995). Interpretivist case study research is focused on understanding the phenomenon in question through accessing the meanings that the subjects of study assign to them (Orlikowski & Baroudi, 1991, p. 5). The primary source of data for interpretivist case studies are interviews that enable researchers to access case participants' views and interpretations of actions and events (Walsham, 1995).

We selected cases based on theoretical sampling (Eisenhardt, 1989). We chose to sample our cases based on the four archetype data ecosystem structures discussed in section 3: keystone-centric, intermediary-based, platform-centric, and marketplace-based. We consider these archetypes of theoretical importance, since the role of the convener is part of and shaped by the data ecosystem structure. Thus, we included four cases which fill in these theoretical categories (Table 2). Such case selection strategy is aimed at extending emergent theory (Eisenhardt, 1989) and investigating convener roles in various contexts in order to find differences and similarities. It is however important to note that some of our cases may exhibit features of several data

Table 2
Selected cases.

Archetype data ecosystem structure	Case	Description	Convener
Platform-centric	AI Innovation of Sweden 2019 –	AI Innovation of Sweden is a national initiative that aims to accelerate the implementation of AI through the sharing of knowledge and data, co-location of competences, and collaboration projects. It provides a data infrastructure (Data Factory), together with processes and frameworks for data access, to make data exchange between actors easier by minimizing transaction costs and interoperability and usability problems.	AI Innovation of Sweden Consortium funded by public and private organizations
Marketplace-based	Amsterdam Data Exchange (AMdEX) 2018-	The Amsterdam Data Exchange (AMdEX) is an initiative of Amsterdam Economic Board that aims to provide broad access to data for researchers, companies and citizens. The AMdEX is intended as a data market on which public and private parties can buy and sell data.	Amsterdam Economic Board (foundation that connects public, private and knowledge organizations)
Keystone-centric	Data Ecosystem of Statistics Netherlands 2019-	The Data Ecosystem is an initiative of Statistics Netherlands (CBS) to better match the offering of data products and services by businesses, academic institutions, and public organizations with the actual needs of local government organizations in the Netherlands (municipalities and provinces). As the main data provider, CBS is the driving force of the ecosystem.	Statistics Netherlands ('Centraal Bureau voor de Statistiek', CBS)
Intermediary-based	JobTech Dev 2018-	JobTech Dev is a data ecosystem focused on improving digital matchmaking and guidance services related to the Swedish labor market. One of the data sets shared in the ecosystem regards job ads relating to the Swedish labor market. For this data set, SPES functions as an intermediary by collecting the job ads from 10 large ad providers and publishing the ads through a public job ad API.	Swedish Public Employment Service (SPES)

ecosystem structures, yet we assigned them to one type based on which features in our view dominated. This was done based on an initial screening of publicly available case information and how the initiatives presented themselves. Additionally, the cases fulfilled the following criteria:

- A case whereby a convener has recently initiated a data collaborative among diverse stakeholders
- Cases represent diverse domains and contexts
- Availability for interviews and document studies

Within each of the four initiatives we selected two projects for further investigation as embedded subcases (Yin, 2018), see Table 3 below. Zooming in on the level of subcases allows for a more contextualized investigation of the convener roles. We selected the subcases from different domains within the four data ecosystems. The subcases are further described in the discussion of results.

The data for our research was collected between March, 2020 and February 2021. The main source of data were in-depth interviews with key stakeholders in each case (21 in total) (Table 4a). For every case we interviewed not only the conveners of the ecosystem but also respondents from main partner organizations in the selected subcases in order to obtain their views on the role that the convener plays. In each case when selecting whom to interview on the convener side we ensured that we interviewed respondents in leadership function who have different competences and position in the respective organizations. Thereby we ensured that different perceptions of convener roles are included in our analysis. The interview protocol is provided in the Annex. It was structured in such a way so that to include questions about the background of the case, the perceived roles of the convener from the perspective of the convener themselves and from the perspective of the subcase actors, and the expected or anticipated evolution of these roles. We thus aimed to capture the backward- and forward-looking perceptions of the respondents, similar to other research aiming to elicit expectations (Sun & Medaglia, 2019; Mergel, Edelmann and Haug, 2019). Following an interpretivist approach, we deliberately kept the interview questions open-ended (instead of providing the interviewees with the categories of convener roles from our theoretical framework) so that to allow respondents more freedom in highlighting the critical issues in their own words. In the interviews we asked questions about roles perceived to have been critical so far and roles that the respondents anticipated to become critical in the future.

To analyze the interview data, we followed several steps in coding the data. First, we coded the data by reading the transcripts and assigning codes inductively (open coding). We also assigned whether the excerpt related to the present or the future convener roles (temporal codes). Then we related the developed codes to the categories from Table 1. We however remained open and let the data speak for itself meaning that when no suitable categories from the framework could be applied, we created new categories inductively. This approach is illustrated in Table 4b, and the results are presented in Tables 5–-8.

The following step was to aggregate the insights of the different groups of actors in each case (convener, actors in subcase 1, and actors in subcase 2). In order to do so, we chose to represent the criticality of roles by means of the following three categories:

- primary role – role that was perceived as critical by all groups of actors,
- secondary role – role that was perceived as critical by more than one group of actors, and

Table 3
Selected subcases within each of the data ecosystems.

Initiative	Selected subcases	Partners	Domain
AI Innovation of Sweden	AI Impact Lab	Örebro University, Region Örebro	Healthcare
	National space data lab	National Space Agency, Luleå University, RISE Institute	Environment
AMdEX	Smart city data market	Start-up, City of Amsterdam, and Police	Public safety and mobility
	Energy data market	Port of Amsterdam and Excess Material Exchange	Energy
Data Ecosystem of Statistics Netherlands	Datalab Gelderland Oost (GO)	Six municipalities in Gelderland province, Berenschot Consulting	Local government
	Labour market data	Etil research group	Labour market policy
JobTech Dev	Labour market matchmaking	Various job ad providers	Labour market
	Labour market statistics	Statistics Sweden, job ad providers	Statistics services

Table 4a
List of conducted interviews.

Case	Respondent's role	Respondent's position	Interview date	Code
AI Innovation of Sweden	Convener	Scientific director of AI Innovation Sweden	4 March, 2020	AI1
		Director of operations of AI Innovation of Sweden	19 August 2019	AI2
	Subcase actor – AI Impact Lab	Director of AI Impact Lab at Örebro University	27 February, 2020	AI3
		Business developer at Örebro University	18 March, 2020	AI4
		Development manager at Region Örebro	24 March, 2020	AI5
	Subcase actor – National Space Data Centre	Head of Innovation at National Space Agency	19 March, 2020	AI6
		Manager for business development at RISE	15 April, 2020	AI7
AMdEX	Convener	Project and technical lead of AMdEX	7 October 2019	AM1
	Subcase actor – Crowd management data market	Entrepreneur at a start-up	12 February, 2020	AM2
		Project manager at City of Amsterdam	24 August, 2020	AM3
	Subcase actor – Energy data market	Business developer at Port of Amsterdam	3 March, 2020	AM4
		Technical consultant at Axians	10 June, 2020	AM5
		Director of Excess Materials Exchange	8 September, 2020	AM6
Data Ecosystem of Statistics Netherlands	Convener	Director Data Services, Research and Innovation	1 April, 2020	DE1
		Director Strategy and Board Advisor	9 April, 2020	DE2
	Subcase actor – Datalab GO	City manager Bronckhorst municipality	13 November, 2020 (joint interview)	DE3–1
		Managing director Berenschot Consulting		DE3–2
	Subcase actor – Regional labour market data	Commercial director Etil research group	14 October, 2020	DE4
JobTech Dev	Convener	Product owner at Swedish National Employment Service	16 November, 2020	JT1
		Project manager at Swedish National Employment Service	18 November, 2020	JT2
	Subcase actor – Labour market matchmaking	Manager at an ad provider	5 February 2021	JT3
	Subcase actor – Labour market statistics	Manager at Statistics Sweden	10 December, 2020	JT4

Table 4b
Example of coding the interview data.

Interview excerpt	Code	Temporal codes	Category from the framework
"To make the transition to such a self-sufficient initiative, CBS considers itself important in developing the organization structure and governance, including financial matters and determining who will steer and make decisions concerning memberships of the ecosystem"	Development of organization and governance structure to achieve a 'self-sufficient' collaboration in the future	Future	M2 Providing governance tools and coordinating structures
"Their networking capacity will be the thing that increases the most to make sure that if you have something within AI and want to do something that would be the natural place to start to see who should I work with, what should I do etc. to get an insight into what the capacities in Sweden are"	Growing networking capacity and becoming the main contact point	Future	No existing category applicable, new created: C3 Connecting disparate actors and initiatives in a one-stop-shop

- emerging role – role that was perceived to be critical in the future by more than one group of actors.

To fill these three categories of role criticality we engaged in another round of coding, at a higher level of abstraction, whereby we synthesized the formulations of roles by the convener and the subcase actors from Tables 5-8. For example, in the AI Innovation case the connecting role was perceived as critical by all groups of actors; thus, the label "Central hub and networking partner" represents the sub-roles as described by the convener and the actors in subcases 1 and 2 (first row in Table 5). Further results of this analysis are presented in Table 9.

The following step was to conduct within-case analysis by producing detailed descriptions for each of the four cases. This enabled us to gain rich familiarity with each case and allowed for the unique patterns of each case to emerge (Eisenhardt, 1989). The within-case analysis is presented in sections 4.1-4.4 and focused on identifying convener roles that were perceived as critical (presently and in the future) by the convener on the one hand and by the subcase actors on the other hand.

The within-case analysis was followed by the cross-case search for patterns which we carried out by looking for within-group similarities coupled with inter-group differences (Eisenhardt, 1989). To be exact, we searched for similarities across the four cases and explored the differences based on which data ecosystem structure the case represented. In doing so, we continued coding and further integrated the primary, secondary, and emerging roles into convener meta-roles (last row of Table 9). The cross-case comparison in presented in section 6 and our key observations of the emerging patterns are summarized in three propositions. Our interpretive case studies were aimed at generating rich insight and contributing to theory as possible forms of generalizations from case study research (Ibid.).

4. Findings

4.1. The case of AI Innovation of Sweden

This case is an example of a platform-centric data ecosystem structure since it provides a data infrastructure (Data Factory), together with processes and frameworks for data access, to make data exchange between actors easier by minimizing transaction costs and interoperability and usability problems. The case also shows elements of an intermediary-based data ecosystem structure connecting AI experts with domain specialists and problem owners and stimulating the innovation process.

Table 5
Comparison of perceived convener roles in the AI Innovation case in the present and future (default font – present, *italic font* – future).

	Convener	Subcase actors 1	Subcase actors 2
ROLES			
Connecting role	Hub, meeting place (C3) *Continue to be the main hub for AI innovation (C3)*	Unified front, common denominator (C3) Access to the right partners and talent (C1) *Liaising with and incentivizing the big players (C1)*	Network and channels to reach out to AI community (C1) *Growing networking capacity and becoming the main contact point (C3)*
Mediating role		*Strengthening ties between different nodes directly (M1)*	Acting as project manager (M2)
Stimulating role		*Advocating and promoting for AI innovation with government and funders (S3)* *Turning focus more to the needs of society and humanity in general (S3)*	*Help securing funding and promoting the subcase (S2)*
Learning catalyst	Sharing of ideas, solutions, data through seminars and events (L1)	*Enabling collective learning and expertise across the network (L1)*	Organization of events (L1)
Infrastructure provision	Data Factory (I1) Office space for nodes (I1)	*Evolving Data Factory into a Model Factory (I1)*	

AI Innovation of Sweden[3] is a national initiative officially launched in February 2019 and primarily funded by the Swedish innovation agency Vinnova. The initiative was set up "*to bring together players from academia, public sector, and industry together around common problems to help spread data and to work together on common infrastructures, common labs and projects*" (interview code AI1). Politically speaking, the initiative emerged from the sense that "*we needed to do something together and concentrate efforts within AI in Sweden, especially AI that can get out to society to really provide innovation*" (AI1).

Prior to the formation of the AI Innovation of Sweden, the collaboration space among stakeholders was fractured, bi- or multi-lateral, and project-based. AI Innovation of Sweden thus aimed "*to provide a common arena for all players*" across application areas who would like to work with AI-driven innovation. The convener perceives that the absolute main function of AI Innovation of Sweden is to serve as "*a hub, a meeting place, and a sharing place of experiences, ideas, data, solutions and so on*" (AI1). To realize this vision, AI Innovation of Sweden as convener provides an organizational infrastructure to its network of partners, such as office space for 'nodes' in different regions of Sweden, as well as organizing seminars, events, networking activities. Besides, AI Innovation of Sweden put in place a data infrastructure called Data Factory that can be used for storing and sharing data within the network and aims to connect and coordinate existing data sharing initiatives. According to the convener, the unique value proposition of Data Factory is that it also provides processes and a framework to access data which makes is easier for prospective users. The convener considers that AI Innovation should not be viewed as a federation of local initiatives but as an initiative for the whole of Sweden, with local presence.

The convener perceives that in the future AI Innovation will realize its vision and continue to serve as a national hub and the main contact point for AI-driven innovation in Sweden. According to the interviewee, they have been extremely successful so far having been able to attract 80 partners in the building phase of the initiative. As the project was moving to the execution stage, the convener highlighted that the main challenge had been "*to line up everyone's expectations with what is possible to do and how to do it within the resources available*" (AI1).

The first subcase – AI Impact Lab – is a co-funded collaboration between Örebro University and Region Örebro launched in September 2019 and aimed at accelerating the uptake of AI within the region of Örebro. The Lab drives pilot projects with engineering competences focusing on the public sector and, as a secondary activity, runs student challenges to discover talent. Due to their expertise, the AI Impact Lab was approached by AI Innovation of Sweden and put on a fast track to become a node (which happened in December 2019).

According to the subcase respondents, "*the greatest value of AI Innovation is that it doesn't fragment Sweden too much*" and creates "*a unified front*" and acts as "*a common denominator*" to give Sweden a chance to catch up with the rest of the world in AI (AI3, AI5). For an initiative from a relatively small region, such the AI Impact Lab, being part of the national network has added more visibility and has been extremely valuable, because their partners do not need to choose and know that they are in sync with AI Innovation of Sweden. It has also been valuable to be able to "*tap into the network of talent*" that AI Innovation gathered around itself (AI5). However, the respondents pointed out that it was still very early to draw conclusions about any direct benefits of being part of the network.

In a future-looking perspective the expectations of the subcase actors are that they will be able "*to drive pilots and projects with much more force and much more speed because they can join forces with the right partners*" from AI Innovation of Sweden (AI3). It is the connecting role of the AI Innovation that could enable them to connect with other geographically dispersed partners. The second aspect where AI Innovation could deliver value according to the subcase actors is enabling learning among partners and sharing lessons learnt to help avoid making the same mistakes and build collective expertise. The communicating role is also seen as important in that AI Innovation spreads the ideas and knowledge developed in Sweden (AI3). From the subcase actors' perspective, AI Innovation of Sweden could accelerate progress by pushing the 'big players' (like Google) for more leadership. Besides, in the future it will be necessary to better strengthen the ties between the different nodes spread across the country (AI5). At the time of the interview the subcase did not use the Data Factory as there was no need for it yet, but they anticipate using it in the future for instance for the organization of challenges and to help the region connect to this resource by for instance donating some data to the Factory from local partners. From the subcase actor perspective, the Data Factory should also become a 'Model Factory' because "*it is not only the data that matters so much, it's also the models that are generated and the ability for us to do verification*" (AI3). While at present AI Innovation serve as "an engine" powering the innovation process and the building of the ecosystem, in the future they could become more of an advocate promoting AI research and innovation with the government and funding agencies and communicating the results of AI work in Sweden (AI4). Another desired evolution of AI Innovation, according to the respondent, is, once they have served the needs of the members, to turn more to the needs of society and broader questions that humanity faces (e.g., energy and sustainability). In fact, seeing more implementation and showing what can be done with AI for the society is perceived as the most important from the subcase perspective (AI4).

The second subcase – National Space Data Centre – emerged from the need of the user community (mainly government organizations but also developers of services) to have the data, processing, and different methods in one place in order to lower the thresholds and make is easier

[3] https://www.ai.se/en

to use space data in Sweden. The subcase is driven by the National Space Agency, in collaboration with other government agencies who are potential users of the data (such as e.g., the Swedish Forest Agency, Swedish Agency for Marine and Water Management, Swedish Environmental Protection Agency) and with the RISE research institute (who provides the capability to handle big amount of data). AI Innovation of Sweden has a formal role in the subcase as a project leader and provided, through Vinnova, the initial funding for the National Space Data Centre.

Overall, the subcase actor (NSA) is positive about the collaboration with the AI Innovation of Sweden and assessed the initiative was very successful and has a lot of potential for the future. The key value that the convener provides to the subcase, according to the interviewee, was the network and the channels to reach out to the AI community. Surprisingly, there was also a downside mentioned: there was not a lot of awareness about what AI Innovation of Sweden was among agencies and that it was unable to enter in contractual agreements with other organizations because it was not a legal entity (AI7). The respondent from RISE emphasized that the main value of AI Innovation of Sweden has been their function as the project coordinator. RISE already had an established network throughout Sweden and had been driving the AI agenda and data science projects for years, whereas AI Innovation of Sweden was a new player building up their network and developing their vision. However, the assistance with the organization of the space data hackathon was seen as very helpful (AI7).

The expectations of the subcase actors are that in the future "*their networking capacity will be the thing that increases the most to make sure that if you have something within AI and want to do something that would be the natural place to start to see who should I work with, what should I do etc. to get an insight into what the capacities in Sweden are*" (AI6). Another important aspect mentioned was that the most pressing need of the subcase was securing funding and that being part of the AI Innovation of Sweden could help with that by promoting and showcasing why the subcase is important.

The case findings are summarized in Table 5 which presents critical present and future roles from the perspective of the convener and the subcase actors and relates them to the roles and sub-roles as presented in our theoretical framework in Table 1. We introduced a new sub-role based on the analysis of this case: Connecting disparate actors and initiatives in a one-stop-shop (C3).

The foregoing analysis summarized in Table 5 above shows that the role of the convener differs in the eyes of the subcase actors and the convener themselves. Overall, there is an agreement that the most critical role of the AI Innovation of Sweden (at the time of the interviews) was their connecting role. Yet, the user case actors and the convener emphasized different sub-roles associated with the connecting role: a hub vs access to the right partners vs access to AI community.

The value of other roles which the convener tried to pursue, such as being the learning catalyst and providing the infrastructure, was less critical at the time for the subcase actors. There are also stark differences between the subcases as to the perceived role of the convener. The first subcase, AI Impact Lab, saw more value in being part of AI Innovation of Sweden and expressed much hope about the possible evolution of the convener role. For the second subcase, the National Space Data Centre, the major need and expected future role of AI Innovation would focus on securing funding and continuing with their connecting role. This points to the fact that conveners have to deal with diverse expectations from their subcase partners when shaping their mission and goals for the future. The subcases had different current and future needs which would have to be addressed by different convener roles. Also, the offered data infrastructure and the data sharing expertise was not at the time of this research of crucial value to the interviewed subcase actors. The space data case had their own data facilities which formed a part of the Data Factory, and the AI Impact Lab's data needs were met in other ways.

4.2. The case of Amsterdam data exchange

This case shows elements of a marketplace-based structure because the Amsterdam Data Exchange[4] (AMdEX) is intended as a data market on which public and private parties can buy and sell data. The Amsterdam region strives to be a leading region in data-driven innovation in Europe by 2025. The Amsterdam region particularly aims to lead in terms of its digital infrastructure, capacity to build innovative data-driven applications, and trust in a digital society.

As part of this strategy, the Amsterdam Economic Board (AEB) initiated AMdEX as a joint (open) data market in 2017. AMdEX connects three types of stakeholders: data providers, data hubs and data users together. AMdEX would offer these stakeholders facilities to process and deliver data on stakeholders' own conditions. AMdEX design principles are user-driven (e.g. data owners remain in control of their own data), transparent, secure, and scalable. AMdEX aims to be a neutral platform that sets the rules for trust in data exchange between those actors. The desired outcome of AMdEX data market is to offer a platform to identify, collect, store, combine, process and capitalize on (open) data. Furthermore, the AMdEX platform aims to offer value-added services related to software, models and hardware (e.g. sensors) to participating stakeholders on the AMdEX platform, allowing other (private) organizations to join AMdEX to offer their data services on the platform.

Initially, the founders of the AMdEX platforms aim at three main roles: technological infrastructure, rules for trust and cooperation, and matching supply and demand. First, AMdEX offers an infrastructure to stakeholders to access and operate the data market. AMdEX plans to connect with generic storage providers and other external infrastructures (e.g. cloud services), offers secured data sharing facilities, supports connectivity and bandwidth between services and provides software. Second, AMdEX aims to provide governance mechanisms that should ensure trust and cooperation between stakeholders. AMdEX represents a legal framework for data exchange, e.g. standardizing and scaling model agreements between multiple stakeholders on the AMdEX platform. Rules include data quality assurance, data access, connectivity, transparency, ownership/control, etc. This framework would replace the conventional vis-à-vis data contracts that are normally set between a data provider and supplier. As a convener, AMdEX envisions to audit and enforce the platform rules, and act as a neutral mediator to resolve possible conflicts between stakeholders. Last, AMdEX has a role as matchmaker between stakeholders on the platform. Specifically, this means that AMdEX offers services to select, filter, annotate and connect data providers and users.

The first subcase – the Smart City project – was started at the end of 2018. A start-up, the city of Amsterdam, the local police organization, and a soccer stadium started to develop a project that aims to combine software (e.g. data mining and machine learning), hardware (e.g. sensors and cameras), data from multiple sources (e.g. security, traffic and crowd management data) to ensure the mobility, hospitality and safety of guests during large events (e.g. soccer matches). The initial geographical scope of the project was the direct area of the soccer stadium during large events, which could be scaled up to smart city services to the entire city when successful. This scaling-up to the entire city became particularly urgent to enable better crowd management during the Covid-19 crisis. The combination of multiple data sources from various private and public parties makes it a complex and unique subcase.

AMdEX adopted this project as an informal smart city subcase. Initially, AMdEX organized and joined brainstorm sessions, workshops, networking and knowledge exchange sessions, during which AMdEX facilitated these sessions to further work of the subcase concepts and advised on how to share data between consortium members. Also, they helped to inform and mobilize stakeholders within and outside the

[4] https://towardsamdex.org

Table 6
Comparison of perceived convener roles in the AMdEX case in the present and future (default font – present, *italic font* – future).

	Convener	Subcase actors 1	Subcase actors 2
ROLES			
Connecting role	Matchmaker between data providers and data users (C2)	Supporting business development by expanding the social network of stakeholders (C1) Mobilizing and committing top management support for the subcase (legitimization) (C1)	Supporting business development by expanding the social network of stakeholders (C1)
Mediating role	A legal framework for data exchange (M2)	*Standardize the legal requirements and contracting between parties (M2)* *Mitigating legal risks and lowering the transaction costs of sharing data (M2)*	*Technical and legal standardization of data-exchange between parties (M2)* *Trusted-third partner that guarantees a neutral governance of the platform, minimizing the risk of power imbalances or abuse (M2)*
Stimulating role		Providing knowledge and inspiration (e.g. via workshops) to further develop the concepts behind the subcases (S3) *Stimulate innovation process by helping organizations to work towards proof of concepts (S1)*	*Stimulating the innovation process by acting as a venture capitalist funding the emergence of subcases as start-ups (S2)*
Learning catalyst		Providing structure or space for joint knowledge creation and mobilization of a collective body (L1) *Providing research and expertise on the issue, providing best practices and resources for collaboration (L2)*	Think tank or sparring partner, providing expertise to the subcase (L2) *Engage and manage an online community of developers (L1)*
Infrastructure provision	*Platform infrastructure for stakeholders to access and operate the data market (I1)*		*Offering secure data storage (I1)*

perimeter consortium to get involved in the project: *"They predominantly think along with us to further detail the subcase, and to inform and mobilize stakeholders"* (AM2). The interviewees indicated that AMdEX's current role was focused on building and sharing knowledge and helping the subcase to expand its network. They characterized this role as a *"sparring partner"* (AM2 and AM3), "*advisor*" (AM2) and a *"project office"* (AM2) that helped to initiate new subcases. One of the interviewees expect that this initial advisory role is a first important step to better understand how a technical data platform should be designed to add value to the subcases: "*They purposefully choose to first build their own expertise, and advise subcases, and as soon as they better understand what is needed, they will take the step to design the technical platform"* (AM2). The role has slowly changed over time with AMdEX developing data services for the subcases, such as filtering and transferring large and sensitive datasets. While AMdEX has resources in the form of a small team of (technical) experts, advisors and an external business developer, the initiative lacks resources to fund a technical platform that would support data exchange in the smart city subcase. Hence, the vision of AMdEX as a full-fledged platform that enables organizations in the Amsterdam region to share and commercialize data remains an idea. One of the interviewees noted: *"The future role is unclear"* (AM3), indicating AMdEX's role ambiguity and search for an identity: will they become a public, private or public-private platform? This ambiguity is reflected by the different innovation approaches the interviewees see for AMdEX: entrepreneurial and emergent versus a more planned strategy.

The interviewees from the subcase foresee potential added value from the AMdEX platform in the future. First, making bilateral agreements about sharing data between parties currently is a cumbersome and inefficient process. As a platform, AMdEX could standardize the legal requirements and contracting between parties. As one of the interviewees put it: "*We need to make new contracts for bilateral data exchange, which is inefficient. We need to reinvent the wheel each time, and it would be of great help if this could be standardized*" (AM3). Additionally, AMdEX could mediate between parties when conflicts related to data exchange occur. Second, the interviewees indicated that organizations are reluctant to share data due to risks related to data protection regulation, loosing competitive advantage, lack of knowledge on data preparation and sharing, and security. For example, one of the interviewees indicated: *"Commercial parking service companies near the airport are reluctant to openly share data on how many parking spots they have still available"* (AM2). However, many of these risks are in practice easy to be mitigated. AMdEX could act as a trusted-third party that lowers the threshold for these reluctant organizations by mitigating those risks and lowering the transaction costs of sharing data. Third, interviewees hope that AMdEX will act as an innovation platform for emerging data collaborations, bringing actors together and helping organizations to work towards proof of concepts. While one of the interviewees regrets that the promise of AMdEX as a technical platform has not been kept so far, they indicate that the technical deployment and management could be done by a commercial IT-vendor, and AMdEX could keep their knowledge, innovation and advise role. One of the interviewees compares the future role of AMdEX with the Amsterdam Exchange (AEX), *"with the task to ensure data exchange between parties goes well and innovative tools and data services are available, while commercial parties could deploy and manage the technical platform"* (AM2).

The second subcase – the Energy Data market – aims to develop an open energy data market in the Amsterdam region, starting with a pilot in the Amsterdam port area. The rationale of this subcase is that the capacity of electricity in the Amsterdam area may reach its limits, particularly for businesses that demand more and more capacity (e.g. due to automation of production processes and logistics). Stakeholders simply have the strong need to ensure business continuity: *"The goal is not to merely decrease electricity costs, but to ensure that the lights remain on*" (AM4). Hence, stakeholders aim to develop a local energy data market that optimizes demand (e.g. forklifts) and supply (e.g. solar panels) of electricity in the Amsterdam region. The subcase existed before the AMdEX initiative: it was started in 2016 at the innovation center of the Port of Amsterdam.

Initially, the energy data market was designed as a decentralized blockchain architecture, in which all electricity-consuming or producing actors in the network may enter in smart contracts based on forecasting (e.g. based on weather data). After a successful small-scale pilot, the initiators founded a coalition with a blockchain and an automation consultancy firm to further develop the concepts and architecture behind the local energy market. At this time, one of the initiators became involved with the AMdEX initiative and suggested the local energy data market as a subcase.

The expectations of AMdEX's roles that the subcase actors had deviated from the roles envisioned by the convener. While AMdEX strongly mainly focused on the infrastructure role and mediating role, the initiators felt that focusing on these roles was too early in the development process of AMdEX. They highly valued a more network/matchmaking role, bringing together entrepreneurs that are motivated to further develop their data market, and provide funding and specific

technical or legal expertise when required by the entrepreneurs. Interviewees experienced role and identity ambiguity of AMdEX as convener: *"AMdEX was all the time searching of its identity"* (AM5). Instead, the interviewees indicated that the identity and roles do not have to be planned top-down but may emerge while further launching the use-case: *"They just have to start launching it: concept, idea and test it. However, they do not act and keep talking about what AMdEX should be"* (AM5). This might be a clash of two different types of business models that require different roles from the convener: top-down or bottom-up approach. AMdEX did not provide financial resources, rather it acted as a "*think tank"* (AM4) or *"sparring partner"* (AM6) according to the interviewees. However, one of the interviewees mentioned the social network that AMdEX provides as a valuable source for business development.

While the subcase members value a match making and networking role of AMdEX in the initial phases, they envision a mediating role in the future. First, AMdEX could provide technical and legal standardization of data exchange between parties on their platform. Interviewees foresee a role in ensuring technical interoperability (e.g. data types) and reference architecture. An example from one of the interviewees is *"secure storage in an EU location*" (AM6). Second, AMdEX could act as a trusted-third partner that guarantees a neutral governance of the platform, minimizing the risk of power imbalances or abuse. Last, interviewees mention that as a trusted third party, AMdEX could engage and manage an online community of developers that help to improve the data exchange on the platform (AM4; AM5).

The foregoing analysis summarized in Table 6 above shows that the role of the convener differs in the eyes of the subcase actors and the convener themselves. Overall, the subcases and convener overlap in their perception that the connecting role, mobilizing social capital for the subcases, is at the time of the interviews the primary role. The mediating role, standardizing and enacting a legal framework for parties on AMdEX, is seen as a critical future role by the convener and both uses cases. However, the convener and the subcases also deviate in the perceived or planned role of the convener. First, the subcases mention that the convener currently adds value to their subcases with a learning catalyst role, e.g., by sharing their knowledge about the organizational, legal and technical aspects of data platforms. This learning catalyst function was actually not envisioned in the convener's initial plans but seems valuable in the initial stages of their data markets. Second, while the convener strives for a role in the infrastructure provision of a joint data platform, the subcases do not assign this role to the convener. Last, the subcase providers would like to see a stimulating role of the convener in the innovation process, e.g., providing seed funding to new data markets or act as a living lab for new subcases. A role in the innovation process was not discussed in the convener's initial plans, which, according to the subcases, provides uncertainty about the convener's identity, role and approach in the innovation process of the subcases.

4.3. The case of Data Ecosystem of Statistics Netherlands

Data Ecosystem[5] is a data collaboration set up by Statistics Netherlands ('Centraal Bureau voor de Statistiek', CBS) in the Netherlands. Data Ecosystem can be considered a keystone-centric data collaboration since it strongly depends on the initiator CBS. The organization started the data collaboration by connecting 'triple helix' organizations: businesses, academia, and local governments. Furthermore, CBS is the driving force behind the activities of Data Ecosystem, organizing networking events and setting up a website. As such, CBS performs the typical roles of a keystone-centric convener: connecting, stimulating, and infrastructure provisioning.

CBS' role traditionally is to gather data and develop statistics to be used by government agencies. "*Until five years ago the organization was oriented towards creating statistics for the national government*" (DE1). In 2015 social welfare provisioning became the responsibility of municipalities, and to perform these tasks, municipalities also need data and statistics. Therefore, CBS started to develop activities specifically geared towards municipalities and provinces. "*Many municipalities, however, need support towards developing data-driven policy beyond merely providing datasets and statistics. They need advice on how to change their organization and on how to develop their infrastructure*" (DE1). To this end, CBS set up Data Ecosystem, involving organizations that CBS already collaborated with, and by setting up a steering board with representatives from 'triple helix' organizations. "*It supports the movement towards data-driven policy in collaboration with businesses, government, and academia*" (DE1).

The main activities of Data Ecosystem performed by CBS were the organization of four networking and matchmaking events in 2019, the development of a website listing best practices, and involving national government organizations that have an interest in data-driven policy making, such as the Ministry of the Interior and the Association of Netherlands Municipalities ('Vereniging Nederlandse Gemeenten', VNG). The collaboration grew quickly: "*the first meetings were attended by 10 to 12 people, later on we needed to rent out large meeting rooms*" (DE1). The website data-ecosysteem.nl lists 24 best practices.

An important goal of Data Ecosystem is to experiment with a collaborative setting for data exchange and identifying success factors for data-driven policy, including the governance of such a collaboration. Therefore, while being the initiator of Data Ecosystem, CBS considers itself a matchmaker, but not the driving force for future development. "*CBS is not responsible for data-driven policy making within municipalities*" (DE2). CBS envisioned Data Ecosystem to become a self-sufficient collaboration, and for other organizations to play a role in this. "*To make the transition to such a self-sufficient initiative, CBS considers itself important in developing the organization structure and governance, including financial matters and determining who will steer and make decisions concerning memberships of the ecosystem*" (DE1). The organization expects to remain involved as a member of the ecosystem bringing in datasets and expertise of data processing, integration, and interpretation. However, in 2020 the initiative was put on hold. Few activities were undertaken, and the main persons involved in setting up the initiative took on different roles. This was partly due to "*the Covid-19 pandemic which slowed down everything*" (DE2), but mainly because it remained unclear who will take on key roles and contribute financially, making it difficult to transition to such a self-sufficient ecosystem. At the end of 2020, a report was issued with the main conclusions of the experiment and recommendations for setting up a governmental data ecosystem.[6]

The first use is the Data lab Gelderland-Oost (GO) which is a collaboration between eight municipalities in the east of the Netherlands, in Gelderland province. The purpose of this collaboration is to facilitate data-driven policy in these municipalities by jointly setting up a data lab. Although the collaboration between the municipalities started before the CBS-led Data Ecosystem initiative emerged, Data lab GO is showcased as a best practice on the website. The data lab was involved early on in Data Ecosystem and gave several presentations at the matchmaking and networking events to share lessons learned and stimulate comparable initiatives. One of the initiators of Data lab GO is a member of the Data Ecosystem steering board. The Data lab considers the main role of CBS and Data Ecosystem to be stimulating. "*The Data lab presented itself a few times at Data Ecosystem meetings, testing the concept. This is the value added of Data Ecosystem*" (DE3–1). Furthermore, CBS as convener is seen as having an important role in providing its knowledge and expertise to different data collaborations.

[5] https://data-ecosysteem.nl

[6] ABDTOPconsult (2020) 'Tussen droom en data… Verkenning Ecosysteem voor een Datagedreven Overheid', 21 December, Tussen droom en data …; verkenning ecosysteem voor een datagedreven overheid | Publicatie | Algemene Bestuursdienst.

Table 7
Comparison of perceived convener roles in the Data Ecosystem of Statistics Netherlands case in the present and future (default font – present, *italic font* – future).

ROLES	Convener	Subcase actors 1	Subcase actors 2
Connecting role	Organizing matchmaker and networking events (C2) Connecting national stakeholders responsible for stimulating data-driven policy making in order to have them take over some of the roles and responsibilities (C1)	Organizing networking meetings (C2) *Scaling and connecting regional data collaborations (C1)*	Organizing networking meetings (C2)
Mediating role	Setting up a steering board with 'triple helix' representatives (M2) *Continuing to play a role in development of organization and governance structure to achieve a 'self-sufficient' collaboration in the future (M2)*		
Stimulating role	Funding all activities, such as the networking events and website development (S2)		Providing expertise on the available data (S2) *Focusing more on the specific demand of a regional data collaboration (S1)*
Learning catalyst	Developing and maintaining a website with best practices (L2) *Continuing to share expertise on data collection, processing, and publication (L2)*	Providing space for testing the concept of Data lab GO during the networking events (L3) Sharing relevant expertise for data collaborations (L2) *Focus should be more on creating a 'learning movement' for data-driven policy making. (L1)*	*Focusing more on demand of the data collaboration on the level of a project or deliverable (L2)*
Infrastructure provision	Acting as the main data provider for the collaboration (I1)		

CBS as data provider and provider of expertise is, therefore, likely to have a role within Data lab GO, but the data lab exists independently from Data Ecosystem. Data lab GO expects CBS and Data Ecosystem to have a role in scaling data collaborations and stimulating learning between data collaborations in different regions. "*In addition to this, a wider 'learning movement' is necessary to exchange knowledge on data-driven policy. To gain momentum and speed up*" (DE3–2). Next to providing and exchanging knowledge and providing (structural, lumpsum) funding, the stimulating and learning movement of Data Ecosystem is considered very important, for example by means of professional education.

The second subcase concerns Etil (short for Economic technological institute Limburg) which is a policy research institute for regional economy and innovation in Limburg province. The Etil "*Labour market in sight*" monitor focuses on monitoring the labour market in different regions. This monitor can be used by regional authorities to make policy decisions. "*Etil has transformed from a traditional research institute to a platform for scientific research. We used to do research and write reports to inform policy. Now we employ data scientists that use open and closed, primary and secondary data sources in dashboards, viewers and platforms that are 24/7 accessible*" (DE4). Most data in the Etil labour market monitor is CBS data. Therefore, CBS involved Etil early on in the Data Ecosystem initiative.

Data Ecosystem may provide Etil's *Labour market in sight* with new exposure and may lead to new collaborations. CBS and Data Ecosystem may thus function as connectors. "*CBS, however, should not have been initiator, but data provider. CBS has tried to connect existing data ecosystems. But this is not the right way to go forward. The right way would be to explore first what is already happening and how we can stimulate knowledge exchange. It has not created new business for CBS*" (DE4). "*The initiative seems to have sizzled, no updates have been made to the website. Yet bilateral collaboration around the Labour market in sight monitor is better than ever*" (DE4). According to Etil, it would have been better to further develop regional data collaborations around their demand, instead of organizing large scale networking meetings with a focus on the available data. "*CBS' vision has failed. The initiative has led to good discussions, but operationalization did not succeed*" (DE4).

The foregoing analysis summarized in Table 7 above shows that the perception of the convener of its role is different from that of the subcase actors. They agree that the main role currently performed by the convener CBS is the connecting role, and that in the future the role of the convener should shift towards the stimulating and learning catalyst roles. However, their perception of how this role should be performed by the convener in the future differs. CBS itself perceives its role as providing datasets and expertise on how to process and integrate data. Furthermore, it expects to bring in its expertise in setting up such data collaborations, including the organization and governance structure. While they do see CBS as a data provider who may also bring in expertise on data processing, both subcase actors, Data lab GO and Etil, perceive the critical roles of the convener to be the intermediary and stimulating role for regional data collaborations, focusing on scaling up and sharing best practices. Furthermore, Data lab GO foresees a role for the convener in setting up a 'learning movement' focusing on professional education and ensuring organizational and culture change. In conclusion, while CBS focuses on data provisioning on a national level, both subcases foresee a shift towards a focus on the regional demand or ecosystem. Infrastructure provisioning was not considered critical by the convener nor by the subcase actors.

4.4. The case of JobTech Dev

This case presents JobTech Dev,[7] a data ecosystem initiated and facilitated by the Swedish Public Employment Service (SPES) with the vision to improve the digital matchmaking and guidance services on the Swedish labor market through the sharing and collaboration on open data. One of the data sets shared in the ecosystem regards job ads relating to the Swedish labor market. For this data set, SPES functions as an intermediary by collecting the job ads from 10 large ad providers and publishing the ads through a public job ad API. The ecosystem can, hence, be characterized as having an intermediary-based structure with elements of a keystone-centric structure where SPES makes up the focal point as the convener for the ecosystem.

The ads are collected from the ad providers' websites (or API:s when available) using open source software that is customized, hosted, and maintained by SPES. Ads are then transformed to a common industry format, enriched with statistical job type identifiers and other metadata, filtered for duplicates, and then reduced to conceptually only include the job title with a link back to the original ad ("back-to-source"). The reduction of ad content and use of back-to-source links is a consequence of negotiations as ad providers consider the ads as having a differentiating value. The solution thereby becomes a value driver for the ad

[7] https://jobtechdev.se/en

Table 8
Comparison of perceived convener roles in the JobTech Dev case in the present and future (default font – present, *italic font* – future).

	Convener	Subcase actors 1	Subcase actors 2
ROLES			
Connecting role	Connecting competing ad providers to collaboration (C1) *Recruitment of further ad providers (C1)*	Driving force for the collaboration (C1)	Focal point in collaboration (C1) *Recruitment of further ad providers (C1)*
Mediating role	Mediating negotiations and discussions (M1) Facilitating and leading governance in collaboration (M2)	Leading negotiations and solving conflicts (M1) *Clarify and anchor vision for project among actors in the collaboration (M1)*	*Maintain open dialogue with ad providers (M1)* *Satisfy different business needs, and keep commitment to sharing ads (M1)*
Stimulating role	*Stimulate an open collaboration and co-development (S4)*		
Learning catalyst	Enable research and validation of technology used (L2) *Stimulate exploration of complementary data sets (L1)*		
Infrastructure provision	Development and maintenance of infrastructure (I1)	Development and maintenance of infrastructure (I1) *Continue to lead development and communicate progress on high level (I1)* *Provide data on insights on ads (I1)*	Development and maintenance of infrastructure (I1) Development of key technology (e.g., algorithms for processing of data) (I1)

providers as *"[job seekers] would hopefully get a more customized ad service"* while "*more -right- individuals see our ads and apply" (JT3)*. A second value driver that emerged with time was the improved statistics related to the labor market. By capturing a large and growing percentage of the available job ads, improved statistics could render in more accurate analysis, reporting, and prediction of the current state and direction of the labor market. As highlighted by Statistics Sweden, "[*the dataset] has the potential to become a complement to our existing vacancy statistics and enable us to reduce our own sampling"* (JT4). Below we will report further on these two value drivers as two subcases.

In the first subcase – improved matchmaking – the convener SPES had an active role in connecting and mediating among the actors within the collaboration from its initiation. The dialogue and negotiations leading up to the collaboration have taken about two years. A possible reason from SPES' viewpoint is that the ad providers have generally been skeptical of the collaboration as they considered themselves direct competitors and that job ads were something differentiating and sensitive, which was confirmed in an interview by an ad provider (JT3). This skepticism in the beginning was also directed towards SPES, as some actors viewed SPES as potentially turning into a direct competitor that would collect and publish all ads on their ad platform. In its mediating role, however, SPES managed with time to establish a level of trust and neutrality among the ad providers. SPES explains part of this success to its standing as a non-competing and not-for-profit actor with overall responsibility for enabling effective matchmaking on the Swedish labor market.

The two-year mediation process has, however, taken its toll among at least some of the ad providers. As expressed by an interviewee, "*we have come further away from the project… I believe everyone feels the same way, show us the product… Considering it has taken one and a half years"* (JT3). The interviewed ad provider acknowledges the complexity of the project but would have preferred a more continuous communication and high-level visualization of the project to maintain interest as well as the vision for the project. The communication has thus far mostly been a technical level while "*most of us just want to see what happens with our ads" (JT3)*. For SPES as the convener, this highlights a need for balance between mediation and communication to maintain commitment from the actors within the ecosystem.

From SPES' perspective, the focus will be on developing its capacities in the roles of mediation and infrastructure provisioning for the collaboration to be able to grow. They consider the initial phase of the collaboration as an experiment where the intention is to arrive at a solution and evaluate it with feedback from the included providers and identify hurdles (e.g., related to technical, process, or legal aspects) that could hinder further providers in joining the collaboration. "*It can't take a year to join the collaboration. It needs to be streamlined with standardized agreements and technical infrastructure*" (JT1). From the mediator's perspective, SPES aims to standardize the governance and processes for the collaboration, and from the infrastructure provisioning perspective, they aim to develop and mature their infrastructure for collecting, enriching, and publishing data.

Long-term, SPES anticipates that their focus as a convener will shift from a mediating to a connecting role as they aim to recruit more ad providers to the collaboration after the demo period. The more providers that agree to join the collaboration strengthens the main value driver for those participating in the collaboration, i.e., increased exposure and traffic related to the ads that they publish respectively. From the perspective of the interviewed ad provider, "*it is still unclear… if it will be an improvement once the solution is in place. I have no idea, will it become big, or will anyone find this? It is hard to tell" (JT3)*. The ad provider further adds that SPES should continue to "*develop the infrastructure and take input as they currently do… but to communicate and be more transparent in their work" (JT3),* again highlighting the need for a more open yet high-level communication within the collaboration.

The second long-term goal of SPES is that the collaboration will continue to mature and evolve into a space for joint knowledge creation and innovation that can be diffused and shared within the ecosystem. In this regard, SPES would take on the role of a learning catalyst, aiming to foster an open collaboration within the ecosystem, similar to what can be seen in open source software and open data communities. To nurture such development, SPES strives to release and develop as much as possible of their infrastructure as open source software. This ambition, however, needs to be balanced against the (as of now) more business-oriented focus from the ad providers. As put by the interviewed ad provider, they are "*less interested in talking code and more interested in how it is going to look? Where should I click?" (JT3)*.

The second subcase – improved statistics – is a collaboration with Statistics Sweden that highlights the importance for the convener to grow partnerships and thereby stimulate innovation and co-creation within the ecosystem. For the ad providers, the validation efforts from Statistics Sweden create further trust in the collected and enriched data. In this regard, SPES acts as a learning catalyst in the ecosystem by enabling research and validation efforts on the underlying technology. However, there is also an interest in complementary data sets as highlighted by the respondent from SPES: *"everyone is interested in the management of duplicates, algorithms, but also in the taxonomy and how they can get systematic in mapping competencies to job types" (JT2)*. According to the ad provider, *"it would enable us to analyze what makes one ad more*

popular than the other, know what soft skills to use, and optimize the ads" (JT3). According to Statistics Sweden, a further value can be created by *"making it easier for employers to proactively assign appropriate job type identifiers to their ads",* and by *"providing statistical insights and reporting on job market segments of importance to the ad providers" (JT4).*

Another aspect that drives ad providers to partake in the collaboration is the feeling of social responsibility as improved statistics of the labor market benefit society as a whole. For some providers, it is even considered as the main motivation for potentially entering the collaboration. SPES refers to an example where a provider "*said the employers who publish ads on [the provider's platform] don't have the muscles to manage the number of calls that would come, should the ad be published via the API and on our ad platform" (JT2).* The employers behind the ads hence only want the limited exposure through the chosen ad provider. However, the ad provider did agree to have its ads collected and used for statistical purposes, as long as it was not published on the public job ad API. This highlights the importance for the convener to be flexible in mediating and finding different ways in how actors can enter the ecosystem and provide value.

In terms of ability to influence, Statistics Sweden considers SPES as having *"…an open mind and interested in hearing our thoughts, we consider ourselves having good opportunities in influencing and providing our perspective" (JT4).* The interviewed ad provider confirms, *"We have been able to provide input, such as identification of duplicates… and I definitely think they have listened" (JT3).*

Taking a long-term perspective, the respondent from Statistics Sweden highlights the importance of a stable solution that can ensure high-quality and free data (JT4). They further believe that SPES' role as a convener going forward is two-fold *"in maintaining an open dialogue with the many actors in the collaboration, as well as packing a solution that cleans and packages the data the is collected, of which none is an easy task" (JT4).* Regarding the former, Statistics Sweden is specifically concerned about the ad providers being motivated to continue the collaboration and not adding fees to the data they share. In contrast to SPES' understanding as noted in the former subcase, mediation can therefore be considered as a key role that they may have to continue to fill even long-term as the ecosystem grows.

Both SPES and the two subcase actors highlight the importance of the connecting role to drive the collaboration forward. SPES and Statistics Sweden also emphasize the importance of growing the ecosystem, while the ad providers are more focused on the short-term results. All actors also consider the importance of SPES as a convener to mediate and lead the discussions with the collaboration to find solutions to business risks or technical issues. The interviewed ad provider especially highlights a need to improve communication, anchoring of the vision, and iterative development of the infrastructure – issues not reflected upon by the convener at the time of the interviews. Contrary to the two subcase actors, SPES anticipated less focus on mediation and more on the connecting role to grow the collaboration.

As a convener, SPES also provides the infrastructure and performs all product development related to the job ad API - a key enabler for the collaboration according to all three actors. SPES may in this sense be compared to a software vendor, but where the customers pay in terms of data and attention, rather than in monetary terms. SPES's long-term ambition, however, is to share and enable collaborative development of the common infrastructure and become the facilitator of a networked ecosystem, rather than the focal point of the ecosystem's current spoke-wheeled structure.

5. Discussion

In the previous section we reported on our case analysis from the perspective of three groups of actors (convener, actors in subcase 1 and actors in subcase 2). In this section we present a comparative analysis of which roles are perceived as critical across the analyzed cases. We propose to discuss the criticality of the roles using the following categories:

- *primary* role – role that was perceived as critical by all groups of actors,
- *secondary* role – role that was perceived as critical by more than one group of actors, and
- *emerging* role – role that was perceived to be critical *in the future* by more than one group of actors.

We further integrated the primary, secondary, and emerging roles into convener meta-roles (last row of Table 9).

Table 10 below relates the findings of the cross-case comparison to our theoretical framework. We found the theoretical framework comprehensive and made several additions to it based on the case

Table 9
Comparative analysis of convener roles perceived as critical across the cases (1 – primary role, 2 – secondary role, E – emerging role; * shows the roles not envisioned by the convener).

ROLES	AI innovation	AMdEX	JobTech Dev	Data Ecosystem
	Platform-based	*Marketplace-based*	*Intermediary-based*	*Keystone-centric*
Connecting role	1 – Central hub and networking partner E – Main national contact point for AI innovation	1 – Networking partner	1 – Neutral party connecting competitors E – Recruiter of new partners	1 – Organizer of networking events
Mediating role		E – Trusted third party and mediator	2 – Mediator in negotiations and conflict resolution E – Project leader responsible for project vision and partners' commitment	
Stimulating role	E – Provider of funding, subcase ambassador maximizing societal impact*	E – Provider of funding and business developer*		2 – Initiator, organizer and provider of funding
Learning catalyst	2 – Meeting place for knowledge exchange in the network	2 – Advisor on legal and technical issues*	2 – Focal point for data innovation	2 – Expertise and best practices of data processing and collaborations E – Tailormade data expertise on demand and initiator of a larger learning movement
Infrastructure provision		E – Provider of data facilities	1 – Value-added service provider	
CONVENER META-ROLES	**Central hub and knowledge center**	**Mediator and market master**	**Trusted third party service provider**	**Main data provider and expert**

Table 10
Critical convener roles in relation to the theoretical framework (*italics* indicate the additions to the framework we made based on the case analyses).

ROLE CATEGORIES	SUB-ROLES	AI Innovation	AmDEX	JobTech Dev	Data Ecosystem
Connecting role	C1 Connecting actors, mobilizing and committing actors *also to realize a sustainable collaboration*	1E	1	1E	1
	C2 Matching demand and supply		1		1
	C3 Connecting disparate actors and initiatives in a one-stop-shop	1E			
Mediating role	M1 Mitigating potential conflicts of interest, crafting a shared vision, managing expectations			2E	
	M2 Providing governance tools and coordinating structures, *including project management*		E	2	
Stimulating role	S1 Stimulating the innovation process and outcomes, *including tailormade demand-driven support*		E		
	S2 Identifying sources of advice, support, and funding for the innovation process	E	E		2
	S3 Creating the urgency for innovation and providing inspiration	E			
	S4 Ensuring openness and transparency in the data ecosystem				
Learning catalyst	L1 Providing structure or space for joint knowledge creation and mobilization of a collective body	2	2E		E
	L2 Providing research and expertise on the issue, providing best practices and resources for collaboration		2E		2E
	L3 Legitimizing and diffusing knowledge				2
Infrastructure provision	I1 Provide an initial set of resources (platform, tools, *datasets*, expertise, *office space*)		E	1	
	I2 Leverage capabilities across the network				
	I3 Making the creation of products/services by third parties more efficient				

analyses (highlighted in italics in Table 10). Our case analyses also prompted a conclusion that some sub-roles are more focused on supporting individual subcases (e.g. providing advice or expertise on certain issues), while other sub-roles may be directed towards supporting the whole data ecosystem (e.g. providing an infrastructure). This became apparent in the cases of Data Ecosystem and AMdEX where this focused demand-driven approach was suggested by the subcase actors.

In all four cases the convener had to connect public and private interests, and in the cases of AmDEX, AI Innovation and Data Ecosystem, to position themselves in a network of existing collaborations and relationships among subcase actors. This task involved earning trust and assuring their non-competing status to the subcase actors (JobTech Dev) and searching for a value proposition that would satisfy the various subcase actors' needs (AmDEX, AI Innovation, and Data Ecosystem). The Data Ecosystem case also showed that it did not (yet) succeed in setting up a 'self-sufficient' network that is able to operate independently from the convener, as was the objective. This illustrates the challenges discussed by van Hille, Bakker, Ferguson, and Groenewegen (2019) for the convener to balance between being a stakeholder and a neutral party and between leading and facilitating collaboration between parties.

Interestingly, all four cases showed that there was a perception gap between how the convener viewed themselves and which convener roles were found most valuable by the subcase actors. This is the contribution of our study as prior research paid little attention to the different perceptions of convener roles. In some cases, like in the case of AmDEX, this even manifested itself in an 'identity crisis' when the convener intended to move in one direction, while the expectations of the subcase actors were different. Namely, the advisory role was found to be the most critical by the subcase actors, whereas it was not envisioned by the convener. This may be partly explained by the convener lacking prior knowledge and certainty about all potential stakeholders for this initiative which poses additional challenges for conveners, as also explained by Agogué et al. (2017). Similarly, in some cases conveners project that certain roles will be critical in the future, while the subcase actors highlight completely different roles. For instance, in the JobTech Dev case the convener viewed the stimulating role as important in the future, while the subcase actors highlighted other roles. And, whereas the stimulating role in the AI Innovation case was not particularly highlighted as such by the convener, the subcase actors brought it to the forefront based on their needs. In the Data Ecosystem case this perception gap was linked to Statistics Netherlands focusing on its connecting role around the supply of data, while the subcase actors organized around specific data propositions and they would have liked Statistics Netherlands to take on stimulating and learning catalyst roles. This even led to one subcase actors considering the initiative unsuccessful. These findings elevate the importance of articulating value creation when engaging in strategic ecosystem thinking and initiating data ecosystem development (Harrison et al., 2012).

From this analysis we arrive at a conclusion that the convener roles should be constructed and emerge from the interaction with and based on the understanding of the needs of the participating actors. This is in line with the research of Taillard, Peters, Pels, and Mele (2016) and their interactionist view on how ecosystems are developed. For instance, in the AI Innovation and the JobTech Dev cases the actors found the connecting role to be important at present and in the future. And in the Data Ecosystem case the learning catalyst role was considered to be critical by the convener and the subcase actors at present and in the future. These examples point to the convener roles in which there is shared interest of both convener and subcase actors.

We also find relative consensus regarding which roles were considered primary and secondary. We thus offer the following three propositions based on our case studies. Our propositions are meant to summarize our key findings and propose statements for future research validation, similarly to e.g. Andersen, Medaglia, and Henriksen (2012).

Proposition 1. The default critical convener roles at the stage of initiation of the data ecosystem are the connecting role and the learning catalyst role.

In all cases we observe that the connecting role is the primary role, regardless of the scale and level of the initiative. This is in line with the research of Dedehayir et al. (2018). Yet, in the four cases the connecting role is realized by means of different sub-roles or a combination thereof (see Table 10). In the AmDEX case the convener acts as a matchmaker between data providers and data users, thereby supporting business development along the data value chain. In the JobTech Dev case, the convener acts as an intermediary connecting competing data providers. In the AI Innovation case, the convener acts as a central hub for AI-related expertise in the different regions in Sweden. And in the Data Ecosystem case the convener is responsible for organizing most or all activities related to the data ecosystem such as matchmaking events, as well as for data provisioning and sharing relevant expertise.

Furthermore, in all cases we observe that the role of the learning catalyst is the secondary role. Yet again, comparing the four cases we notice that the learning catalyst role is manifested in different convener sub-roles (see Table 10). In the AI Innovation case the learning catalyst role is about offering space for meetings and knowledge exchange; in the

AmDEX case it is realized through expert advisory to subcase actors; in the Data Ecosystem case through exchanging best practices; and in the JobTech Dev case it is envisioned as the infrastructure being the focal point of data innovation in the ecosystem. None of the cases prioritized the development of an "open innovation arena" (Elmquist et al., 2016), although at least in the AI Innovation case the sub-role of stimulating collective knowledge creation was articulated as a desired one by a subcase actor. The learning catalyst role may be particularly critical for data collaboratives, compared for instance with open data initiatives. Prior research on open data initiatives emphasized other functions, such as stimulating (Martin et al., 2017), governance tools (Bonina & Eaton, 2020), and more generally (open data) infrastructure provision (Zuiderwijk et al., 2014). This can be explained by the fact that in the case of data collaboratives there is a large societal demand for innovation and for new ways of seeing problems that require more than one organization to find a solution. According to Agogué et al. (2013), these aforementioned conditions create the need for conveners to play an active role in enabling joint knowledge creation in the data ecosystem. Furthermore, in contrast to open data initiatives, data collaboratives that rely on private partners sharing their proprietary data may require more incentives and legitimatization from a convener to share their knowledge, learn openly in the data collaborative, and to contribute to a join knowledge production (Gupta et al., 2020). Private partners in data collaboratives may require a clearer business case and proofs of concepts with data pilots to commit to knowledge sharing in the collaborative.

With regards to the future convener roles, we formulated the following proposition based on our results:

Proposition 2. As the data ecosystem develops, the stimulating and mediating convener roles emerge as the critical future roles, in addition to the connecting role.

In some cases (AI Innovation and JobTech Dev) the convener is expected and intends to continue with the connecting role and further attract and mobilize relevant actors. In other cases, however, namely in the Data Ecosystem case, the convener aimed to discontinue with the connecting role in the future and ensure the collaboration between actors is self-sufficient. This way, as became evident in the Data Ecosystem case, as well as in the AMdEX case, the convener can move towards playing more stimulating and mediating roles – towards acting as a trusted third party (AMdEX) or as a provider of tailormade expertise on demand (Data Ecosystem). An important observation we make is that for some cases, such as AMdEX, future convener roles can be seen as a search process, considering the 'identity crisis' of the convener described above (M2, S1, S2, L1, L2, I1 in Table 10). In the Data Ecosystem case this search process of the convener manifested itself in the convener's desire to hand over some of its connecting and mediating roles to a wider and 'self-sufficient' partnership of actors. On the other hand, in other cases, like JobTech Dev and AI Innovation, the future convener roles are fewer and more streamlined indicating a clear path forward.

Let us now dive into the nuances differentiating the cases further. The four cases included in our analysis were sampled to represent the four data ecosystem structures proposed by Oliveira et al. (2019) and Gelhaar et al. (2021). Thus, as explained in section 5, we assumed that the convener roles are part of and are shaped by data ecosystem structures. By data ecosystem structure, as explained in section 3 and following Oliveira et al. (2019), we mean how and to what degree actors are connected and depend on each other. We see differences in how the data ecosystem structure shaped the role and sub-roles of the conveners and vice versa. For instance, in the JobTech Dev case, SPES can be seen as a Public Open Innovation Intermediary (Bakici et al., 2013) and a keystone actor making up the focal point of collaboration among the ad providers. Furthermore, in the JobTech Dev case the convener had to orchestrate the collaboration 'from scratch' as no prior collaborations among the use actors existed. Therefore, earning trust and providing an initial infrastructure was of importance. In contrast, in the AI Innovation case, the convener is a consortium where public, private and research interests are represented. Furthermore, this initiative is on a national level posing an ambitious task of creating "a united front" for AI innovation in Sweden. This shaped the convener role as the national hub and knowledge and expertise centre. In the AMdEX case the convener is a non-profit player who is searching for a value proposition to play a convener role in the network of existing collaborations and relationships. Last but not least, in the Data Ecosystem case Statistics Netherlands as convener was the initiator and main organizing and funding partner, but the initiative aimed to become 'self-sufficient' and attract other partners to become jointly responsible for the connecting role, but this was not successful (yet). The meta-roles that we proposed in Table 9 capture these important differences across the cases. We thus offer the following proposition.

Proposition 3. Different convener meta-roles are associated with archetypical data ecosystem structures.

We urge future research to investigate Proposition 3 in more depth and to elaborate in greater detail on the dependency between how the data ecosystem is organized and the role of the convener in it. Future research can investigate how the meta-roles might evolve as the data ecosystem matures.

6. Conclusions

In our study we investigated the question what convener roles are perceived to be critical when developing data collaboratives from the perspective of data ecosystem thinking. We particularly focused on data collaboratives aimed at innovation for addressing societal challenges. Through the analysis of four cases from Sweden and The Netherlands, we identified several archetypical roles that can be played by conveners of such ecosystems.

As a result, we provided a rich and nuanced picture of *how the convener roles can vary* (dependent on the stage of ecosystem development, the underlying data ecosystem structure) and that there is in fact a perception gap between how conveners themselves define their roles and what is seen as critical by the ecosystem actors. Prior research did not provide substantial knowledge in this respect. We further stipulated (by means of propositions) which roles can be seen as primary and secondary and under which conditions. We established that there is a shift in the critical convener roles between the stages of initiation and further development of the data ecosystem. This finding is important to consider for conveners in order to understand how their roles should evolve in order to sustain and grow the data ecosystem. This might help tackle the issue that many data collaboratives have struggled with, namely scaling up towards continued development (European Commission, 2020). Furthermore, through our cases we also concluded that the convener roles are shaped by (and shape) the data ecosystem structure, hence it is important that conveners closely consider the existing (and evolving) structure of the data ecosystem when positioning themselves in it. Thereby, we provided an empirical validation of the hypotheses from the previous literature about the dependencies between convener roles and data ecosystem structure (Dedehayir et al., 2018; Lis & Otto, 2020).

Our research makes a theoretical contribution to research on data collaboratives and on data ecosystems more broadly in a number of ways. First, we reviewed and systematized the convener roles found in several literature streams, such as open data ecosystems, innovation ecosystems, partnerships research, that are relevant for convening data collaboratives. The framework presented in Table 1 thus can be used for further theory building (and testing) in the field. Based on our case analyses, we extended the framework with additional sub-roles as well (Table 10). This provided important nuances since our findings showed how convener roles can be realized through different sub-roles. As explained in the Discussion, the choice of the appropriate sub-role is important as the sub-roles can differ significantly. For example, the connecting role can be realized through such sub-roles, as matchmaker

of data providers and users or as intermediary connecting competing data providers. Without this nuance of the sub-roles, conveners could make false conclusions regarding how to fulfill a certain role. In the aforesaid example, by mixing the two sub-roles the convener could be unaware the competing interests that need to be balanced. Thus, our extended framework of roles and sub-roles can support conveners to identify the activities needed to perform a certain role.

The main limitation of our work regards its generalizability and comes as a natural consequence of the qualitative nature of our research design. We have based our analyses on a limited number of cases, although we included diverse cases in our study. We also focused on initiatives in the early phases of development, although we included a forward-looking perspective in our case analyses. Convener roles change and evolve over time, thus different roles might be critical at later stages of data ecosystem development. We also are aware that the data ecosystem structures of the initiatives may evolve over time. In our study we only focused on early stages of development of the initiatives; future research can investigate the evolution of data ecosystem structures in relation to convener roles.

Limitations also come from how the classification is performed of when an ecosystem passes between different stages in its life cycle, e.g., from an initiating, to a developing, and to a mature stage. This becomes a contextual factor where observers may draw conclusion by comparing and contrasting ecosystems by their respective maturity progress. The reader should consider the contextual factors of our studied cases when comparing with other cases in order to understand what roles and activities should be prioritized. Future research on maturity and life-cycle aspects of data ecosystems may hence be of value both to researchers and practitioners in order to position advice and practice better.

A further aspect in terms of limitations relates to the theoretical sampling which we based on the categorization different data ecosystem structures by Oliveira et al. (2019) and Gelhaar et al. (2021). Yet, as these four types are not mutually exclusive, the cases we have analyzed may fall into multiple categories, although each case was attributed to one overarching category. Readers should hence consider the characteristics of the underlying case when attempting to transfer results to other cases. Future research could help bring clarity and further context to the different data ecosystem structures to help conveners position themselves as to what roles and activities they should take on.

We call for future research to test our framework of convener roles in larger multiple case study designs. We also urge future research to further develop our ideas about the relationship between convener roles and data ecosystem structures. Additionally, there is also room for exploration regarding which competences are required from organizations in order to be able to pursue the convener roles that we identified.

CRediT authorship contribution statement

Iryna Susha: Conceptualization, Methodology, Investigation, Formal analysis, Validation, Writing – original draft, Project administration. **Tijs van den Broek:** Conceptualization, Methodology, Investigation, Formal analysis, Writing – review & editing. **Anne-Fleur van Veenstra:** Conceptualization, Investigation, Formal analysis, Writing – review & editing. **Johan Linåker:** Investigation, Formal analysis, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing ffnancial interests or personal relationships that could have appeared to infiuence the work reported in this paper.

Acknowledgements

The first author received funding for this research from the Swedish Research Council under the grant agreement 2015-06563 “Data collaboratives as a new form of innovation for addressing societal challenges in the age of data” during her previous employment at Örebro University.

Appendix A

Topic list for interviews with conveners

1. Background of the initiative
   a. How did the initiative came to be (short history)
   b. Who are the key players and driving forces (organizational structure)
   c. How do players relate to one another, what kind of shared history do they have
   d. What are the main objectives (goal of data innovation)
2. Main functions of convener
   a. How do you see your role in this initiative
   b. What key activities do you perform in relation to the data ecosystem
   c. How do you add value to the data innovation process
   d. How did or does this role evolve as the initiative progresses
   e. How do parties perceive you as an actor, in what role
   f. How successful have you been in this role
   g. What challenges have you faced in performing your role
   h. How does the future look like for this initiative

Topic list for interviews with subcase participants

1. Background of the subcase
   a. How did the subcase came to be (short history)
   b. Who are the key players and driving forces (org structure) and what are their roles in the subcase
   c. How are you related to convener
   d. What are the main objectives (goal of data innovation)
   e. What are your objectives in this subcase as an organization
   f. What are your challenges and the challenges of the collaboration
2. Perceived functions of convener
   a. What is the role of convener in this project (from research conveners play these and these roles, what do you observe in your subcase)
   b. What key resources do they make available to you
   c. How do (or would) they add value to your collaboration
   d. Do you see how their role would or has changed as the project progresses
   e. What needs could or does the convener meet in this partnership
   f. How successful have they been in their role
   g. How does the future look like for this initiative

References

Agogué, M., Berthet, E., Fredberg, T., Le Masson, P., Segrestin, B., Stoetzel, M., Wiener, M., & Yström, A. (2017). Explicating the role of innovation intermediaries in the “unknown”: A contingency approach. *Journal of Strategy and Management, 10*(1), 19–39. https://doi.org/10.1108/JSMA-01-2015-0005

Agogué, M., Yström, A., & Le Masson, P. (2013). Rethinking the role of intermediaries as an architect of collective exploration and creation of knowledge in open innovation. *International Journal of Innovation Management, 17*(02), 1350007. https://doi.org/10.1142/S1363919613500072

Andersen, K. N., Medaglia, R., & Henriksen, H. Z. (2012). Social media in public health care: Impact domain propositions. *Government Information Quarterly, 29*(4), 462–469.

Bakici, T., Almirall, E., & Wareham, J. (2013). The role of public open innovation intermediaries in local government and the public sector. *Technology Analysis & Strategic Management, 25*(3), 311–327. https://doi.org/10.1080/09537325.2013.764983

Bonina, C., & Eaton, B. (2020). Cultivating open government data platform ecosystems through governance: Lessons from Buenos Aires, Mexico City and Montevideo. *Government Information Quarterly, 37*(3), Article 101479.

van den Broek, T., & van Veenstra, A. F. (2018). Governance of big data collaborations: How to balance regulatory compliance and disruptive innovation. *Technological Forecasting and Social Change, 129*, 330–338.

Crosby, B. C., Hart, P., & Torfing, J. (2017). Public value creation through collaborative innovation. *Public Management Review, 19*(5), 655–669. https://doi.org/10.1080/14719037.2016.1192165

Curry, E., & Sheth, A. (2018). Next-generation smart environments: From system of systems to data ecosystems. *IEEE Intelligent Systems, 33*(3), 69–76.

Dawes, S. S., Vidiasova, L., & Parkhimovich, O. (2016). Planning and designing open government data programs: An ecosystem approach. *Government Information Quarterly, 33*(1), 15–27. https://doi.org/10.1016/j.giq.2016.01.003

Dedehayir, O., Mäkinen, S. J., & Ortt, J. R. (2018). Roles during innovation ecosystem genesis: A literature review. *Technological Forecasting and Social Change, 136*, 18–29.

Dorado, S., & Vaz, P. (2003). Conveners as champions of collaboration in the public sector: A case from South Africa. *Public Administration and Development, 23*(2), 141–150. https://doi.org/10.1002/pad.270

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

Elmquist, M., Ollila, S., & Yström, A. (2016). Beyond intermediation: The open innovation arena as an actor enabling joint knowledge creation. *International Journal of Technology Management, 72*(4), 273–295. https://doi.org/10.1504/IJTM.2016.081573

European Commission. (2020). Towards a European strategy on business-to-government data sharing for the public interest. Final report prepared by the high-level expert group on business-to-government data sharing Accessed 17 June 2022 from https://digital-strategy.ec.europa.eu/en/news/experts-say-privately-held-data-available-european-union-should-be-used-better-and-more.

Gelhaar, J., Groß, T., & Otto, B. (2021). A taxonomy for data ecosystems. In *Proceedings of the 54th Hawaii International Conference on System Sciences* (p. 6113). https://doi.org/10.24251/HICSS.2021.739

Gupta, A., Panagiotopoulos, P., & Bowen, F. (2020). An orchestration approach to smart city data ecosystems. *Technological Forecasting and Social Change, 153*, Article 119929. https://doi.org/10.1016/j.techfore.2020.119929

Gupta, R., Mejia, C., & Kajikawa, Y. (2019). Business, innovation and digital ecosystems landscape survey and knowledge cross sharing. *Technological Forecasting and Social Change, 147*, 100–109.

Harrison, T. M., Pardo, T. A., & Cook, M. (2012). Creating open government ecosystems: A Research and Development agenda. *Future Internet, 4*(4), 900–928. https://doi.org/10.3390/fi4040900

Heimstädt, M., Saunderson, F., & Heath, T. (2014). Conceptualizing open data ecosystems: A timeline analysis of open data development in the UK. In *In Conference for E-democracy and open governement* (p. 245).

van Hille, I., Bakker, F. G. A., Ferguson, J. E., & Groenewegen, P. (2019). Navigating tensions in a cross-sector social partnership: How a convener drives change for sustainability. *Corporate Social Responsibility and Environmental Management, 26*(2), 317–329. https://doi.org/10.1002/csr.1684

Holzmann, T., Sailer, K., & Katzy, B. R. (2014). Matchmaking as multi-sided market for open innovation. *Technology Analysis & Strategic Management, 26*(6), 601–615. https://doi.org/10.1080/09537325.2014.913786

Iansiti, M., & Levien, R. (2004). Strategy as ecology. *Harvard Business Review, 82(3), 68–78*, 126.

de Juana-Espinosa, S., & Luján-Mora, S. (2019). Open government data portals in the European Union: Considerations, development, and expectations. *Technological Forecasting and Social Change, 149*, Article 119769.

Kalegaonkar, A., & Brown, L. D. (2000). Intersectoral cooperation: Lessons for practice. *IDR Reports, 16*(2), 1–25.

Klein, H. K., & Myers, M. D. (1999). A set of principles for conducting and evaluating interpretive field studies in information systems. *MIS Quarterly*, 67–93.

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

Koster, V., & Suarez, G. (2016). Open data for development: Experience of Uruguay. In *Proceedings of the 9th international conference on theory and practice of electronic governance* (pp. 207–210). ACM.

Koznov, D., Andreeva, O., Nikula, U., Maglyas, A., Muromtsev, D., & Radchenko, I. (2016). A survey of open government data in Russian Federation. In *, vol 3. IC3K 2016—Proceedings of the 8th international joint conference on knowledge discovery, knowledge engineering and knowledge management* (pp. 173–180).

Lee, D. (2014). Building an open data ecosystem: An Irish experience. In *Proceedings of the 8th international conference on theory and practice of electronic governance* (pp. 351–360). ACM.

Lis, D., & Otto, B. (2020). Data governance in data ecosystems–insights from organizations. In *AMCIS 2020 Proceedings* (p. 12).

Lis, D., & Otto, B. (2021). Towards a taxonomy of ecosystem data governance. In *Proceedings of the 54th Hawaii International Conference on System Sciences*.

Lnenicka, M., & Komarkova, J. (2019). Developing a government enterprise architecture framework to support the requirements of big and open linked data with the use of cloud computing. *International Journal of Information Management, 46*, 124–141.

Martin, S., Turki, S., & Renault, S. (2017). Open Data Ecosystems. In A. Kő, & E. Francesconi (Eds.), *Vol. 10441. Electronic Government and the Information Systems Perspective. EGOVIS 2017. Lecture notes in computer science*. Cham: Springer. https://doi-org.db.ub.oru.se/10.1007/978-3-319-64248-2_5.

Mercado-Lara, E., & Gil-Garcia, J. R. (2014). Open government and data intermediaries: The case of AidData. In *Proceedings of the 15th Annual International Conference on Digital Government Research* (pp. 335–336). https://doi.org/10.1145/2612733.2612789

Mergel, I., Edelmann, N., & Haug, N. (2019). Defining digital transformation: Results from expert interviews. *Government Information Quarterly, 36*(4), Article 101385.

Modgil, S., Gupta, S., Sivarajah, U., & Bhushan, B. (2021). Big data-enabled large-scale group decision making for circular economy: An emerging market context. *Technological Forecasting and Social Change, 166*, Article 120607.

Oliveira, M. I. S., Barros Lima, G., & Farias Lóscio, B. (2019). Investigations into data ecosystems: A systematic mapping study. *Knowledge and Information Systems, 61*(2), 589–630. https://doi.org/10.1007/s10115-018-1323-6

Oliveira, M. I. S., & Lóscio, B. F. (2018). What is a data ecosystem?. In *Proceedings of the 19th annual international conference on digital government research: Governance in the data age* (pp. 1–9). https://doi.org/10.1145/3209281.3209335

Orlikowski, W. J., & Baroudi, J. J. (1991). Studying information technology in organizations: Research approaches and assumptions. *Information Systems Research, 2*(1), 1–28.

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

Stadtler, L., & Probst, G. (2012). How broker organizations can facilitate public–private partnerships for development. *European Management Journal, 30*(1), 32–46. https://doi.org/10.1016/j.emj.2011.10.002

Sun, T. Q., & Medaglia, R. (2019). Mapping the challenges of artificial intelligence in the public sector: Evidence from public healthcare. *Government Information Quarterly, 36*(2), 368–383.

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

Susha, I., & Gil-Garcia, J. R. (2019). A collaborative governance approach to partnerships addressing public problems with private data. In *Proceedings of the 52nd Hawaii international conference on system sciences*.

Susha, I., Janssen, M., & Verhulst, S. (2017a). Data collaboratives as a new frontier of cross-sector partnerships in the age of open data: Taxonomy development. In *Proceedings of the 50th Hawaii International Conference on System Sciences. Paper presented at 50th Hawaii International Conference on System Sciences (HICSS 2017), Waikoloa, USA, January 4–7, 2017* (pp. 2691–2700). IEEE.

Susha, I., Janssen, M., & Verhulst, S. (2017b). *Data collaboratives as "bazaars"? A review of coordination problems and mechanisms to match demand for data with supply*. Transforming Government: People, Process and Policy.

Susha, I., Rukanova, D., Ramon Gil-Garcia, J., Hua Tan, Y., & Gasco, M. (2019). Identifying mechanisms for achieving voluntary data sharing in cross-sector partnerships for public good. In *Proceedings of the 20th Annual International Conference on Digital Government Research* (pp. 227–236).

Taillard, M., Peters, L. D., Pels, J., & Mele, C. (2016). The role of shared intentions in the emergence of service ecosystems. *Journal of Business Research, 69*(8), 2972–2980.

Verhulst, S., & Sangokoya, D. (2015). *Data collaboratives: Exchanging data to improve people's lives, in medium* (vol. 2015).

Walsham, G. (1995). Interpretive case studies in IS research: Nature and method. *European Journal of Information Systems, 4*(2), 74–81.

World Economic Forum. (2019). Data Collaboration for the Common Good: Enabling Trust and Innovation Through Public-Private Partnerships Accessed 6 July 2020 from http://www3.weforum.org/docs/WEF_Data_Collaboration_for_the_Common_Good.pdf.

Yin, R. K. (2018). *Case study research and applications* (6th ed.). Sage.

Zeleti, F. A., & Ojo, A. (2016). Critical factors for dynamic capabilities in open government data enabled organizations. In *Proceedings of the 17th International Digital Government Research Conference on Digital Government Research* (pp. 86–96). https://doi.org/10.1145/2912160.2912164

Zuiderwijk, A., Janssen, M., & Davis, C. (2014). Innovation with open data: Essential elements of open data ecosystems. *Information Polity, 19*(1–2), 17–33. https://doi.org/10.3233/IP-140329

Iryna Susha is an Assistant Professor at the Copernicus Institute of Sustainable Development, Faculty of Geosciences, Utrecht University (The Netherlands) and previously a postdoctoral researcher at the Department of Informatics at Örebro University (Sweden). Iryna's current research focus is on data collaboratives and how they can be efficiently used to address societal challenges. Her earlier works on open government, open data, and citizen participation have appeared in various international journals and conferences.

Tijs van den Broek is an Assistant Professor at the Organization Sciences department of the Vrije Universiteit Amsterdam. His work addresses data-driven innovation, online social networks, corporate social responsibility, and prosocial campaigning. Academically, he obtained a cum laude PhD in Business Administration from the University of Twente.

Anne Fleur van Veenstra is Director of Science at the Strategic Analysis & Policy department of the Netherlands Institute of Applied Scientific Research TNO. She leads TNO's Policy Lab research activities on data-driven policy, conducting policy experiments using data analytics and AI, e.g. for the energy transition or social services. Anne Fleur widely researched and published on the use and impact of (big) data and artificial intelligence in the public sector, data-driven policy, and transformational government.

Johan Linåker is a Postdoctoral Researcher in the Software Engineering Research Group at Lund University in the wonderful south of Sweden. His focus and passion is helping software organizations move forward by opening up and using open source software, data and innovation as their tools. In his Ph.D., Johan focused on creating guidelines for how software-intensive organizations can develop contribution and community strategies in the context of Open Source Software (OSS) Requirements Engineering.