Data & Policy (2025), 7: e79
doi:10.1017/dap.2025.10018

TRANSLATIONAL ARTICLE

Government data ecosystems for addressing societal challenges: Lessons from two cases in the Netherlands

Annelieke van den Berg, Marissa Hoekstra and Anne Fleur van Veenstra

TNO Vector, Centre for Societal Innovation and Strategy, TNO, The Hague, The Netherlands
Corresponding author: Annelieke van den Berg; Email: annelieke.vandenberg@tno.nl

Received: 25 October 2024; **Revised:** 28 March 2025; **Accepted:** 30 May 2025

Keywords: data ecosystems; data-informed policymaking; societal challenges

Abstract

Societal challenges such as climate change and health inequalities require complex policy decisions, for which governmental organizations rely on a good information position. Having access to data from various domains is seen as a facilitator of making evidence-informed decisions that are more legitimate and less uncertain. To identify and make data available that is stored at various organizations, stakeholders participate in sociotechnical networks, also known as data ecosystems. Data ecosystems aimed at addressing societal challenges are characterized as complex because knowledge about societal issues is uncertain, information is scattered among (governmental) actors, collaboration extends beyond existing organizational networks, and values and interests of network actors can be conflicting. In this translational article, we examine how to successfully establish and maintain data ecosystems aimed at addressing societal challenges, given these complexities. We analyze two cases of successful data ecosystems in the Netherlands and present five narratives about how these data ecosystems navigated these complexities. We find that establishing collaboration among network actors, using bottom-up approaches, contributed to the success of both cases. The cases created structures in which participants were able prioritize the right questions, find common interests, and work together. The narratives present insights for government officials about collaboration in data ecosystems and add to the literature by highlighting the importance of organizational capabilities.

Policy Significance Statement

Policymakers are often faced with complex policy questions to address societal challenges. In the Netherlands, to inform policy with data, data ecosystems are established that may cater to multiple policy questions. This translational article draws lessons from practice about aspects that support the development of such data ecosystems. Most importantly, we find that organizational capabilities of data ecosystems are experienced as particularly important for the success of data ecosystems. We provide five narratives with lessons learned and advice for organizational practices that improve collaboration in data ecosystems to address societal challenges.

1. Introduction

Policymakers who address complex societal challenges of our time, such as climate change, health inequalities, and public safety, have to collaborate with other governmental partners to acquire sufficient information to make policy decisions. In the Netherlands, we see that data ecosystems are created to improve the information position of policymakers pertaining to specific societal challenges. Data ecosystems are sociotechnical networks in which actors collaborate to find, archive, publish, consume,

© The Author(s), 2025. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/dap.2025.10018 Published online by Cambridge University Press

or reuse data as well as to foster innovation and create value (Oliveira et al., 2019, p. 589). Data ecosystems connect different levels of government and cater to multiple policy questions. They are characterized by complexity because knowledge about societal issues is uncertain, information is scattered among (governmental) actors, collaboration extends beyond existing organizational networks and values, and interests of network actors can be conflicting (Klijn, 2002). Participants in such data ecosystems often have a pioneering role, because there are relatively few precedents of data ecosystems addressing societal challenges across government organizations. They encounter technical, organizational, and interpersonal barriers, for which there is not yet a recipe for how to overcome them.

Literature on data sharing among organizations to address societal challenges encompasses both data ecosystems and data collaborations. Data collaborations are cross sectoral, meaning that public, private, or societal partners engage in data sharing and have the purpose of a specific societal challenge (Susha et al., 2017; Ruijer, 2021), while data ecosystems are characterized by intentionality, value creation, and sustainability, meaning that they are "intentionally created, developed, and managed for the purpose of achieving a managerial and policy vision" (Susha et al., 2023, p. 3). Since we focus on exploring the intentional creation of sustained data sharing to develop policy to address societal challenges among public organizations, in this article, we conceptualize our object of study as intergovernmental data ecosystems. Because of their technological and collaborative nature, to better understand how data ecosystems for societal challenges develop and evolve, we need an "integrated technical and organizational perspective" to study them (Bartolomucci and Bartalucci, 2024). In this translational article, we therefore examine two successful data ecosystems in the Netherlands, to answer the question: Which aspects contribute to the successful functioning of intergovernmental data ecosystems aimed at addressing societal challenges?

By answering this question, we provide insights into what helps these data ecosystems thrive. In so doing, we provide insights for government officials about collaboration in data ecosystems, the complexities that they face, and how these can be addressed. At the same time, we hope to inspire current and future research with these rich case descriptions about data ecosystem practices. The article is structured as follows. In Section 2, we briefly touch on literature on data-informed policymaking and data ecosystems. Section 3 describes the methodology used in this research and the background of the examined cases. This is followed by the findings and lessons learned in Section 4. We conclude our article with a discussion and conclusion in Section 5.

2. Literature

2.1. Data for addressing societal challenges

Government organizations become increasingly acquainted with adding data and analytics to the evidence base on which decisions are made in different steps of policymaking (van Veenstra et al., 2020). Examples of such data use can for instance be found in the domains of public health (Danemayer et al., 2023; Urbano et al., 2024), mobility (Liu and Dijk, 2022; Martinazzo, 2023), and social policy (Kersing et al., 2022; Dietrich et al., 2024). It is argued that data can help to understand the issue at hand, estimate the effects of interventions, arrive at decisions that contain less uncertainty, and ultimately achieve better policy outcomes that can also be monitored and evaluated with the use of data (Giest, 2017; Head, 2022). In the execution of policy, it is believed that making decisions based on data and algorithms can improve efficiency and remedy favoritism that might inadvertently be introduced by prejudiced individual human judgments (Kersing et al., 2022; Rij et al., 2024). At the same time, the extent to which value can be derived from data is highly dependent on the capacity of decision makers to comprehend it and weigh it against other forms of information (Giest, 2017).

In this article, we zoom-in on the use of data in the public sector in relation to decision making about societal challenges. The principles about data informed policymaking as outlined before still hold in this context, but the dynamic becomes more complex because addressing a societal challenge is not contained to work being done in one governmental department or organization. Klijn (2002) outlines three characteristics of the complexity in decision making on societal problems.

https://doi.org/10.1017/dap.2025.10018 Published online by Cambridge University Press

The first concerns knowledge uncertainty, meaning that there is a lack of knowledge about both the nature of the problem as well as the right way to solve the problem (Klijn, 2002). Over the past 20 years, the data revolution has significantly increased opportunities for attaining knowledge about a problem through data, but even when these data or information are available, it is scattered among a large set of actors that each may hold relevant information (Head, 2022). When information is fragmented, it is desirable, or even necessary, to set up collaboration between actors to arrive at a proper information basis for decision making (Alford and Head, 2017).

This relates to the second characteristic, namely institutional uncertainty (Klijn, 2002). Large societal challenges do not fit well in existing organizational networks. Societal problems can span multiple policy arenas, and each arena in turn can contain actors of different networks. It takes time and effort to find out where certain data and information is present, and establish arrangements in which this is shared between actors (Alford and Head, 2017).

Setting up such arrangements can be complicated by the third characteristic of complexity, namely strategic uncertainty (Klijn, 2002). Different actors can have conflicting values and interests (Head, 2022), for example, because they have diverging mandates or responsibilities. Actors may even choose to strategically deploy control over to their data, in order to gain an advantage in protecting their interest (Alford and Head, 2017). Cumulatively, these three characteristics outline key complexities in decision making for addressing societal challenges, which have implications for the data ecosystems in which these actors interact and collaborate with each other.

2.2. Data ecosystems for policymaking

Oliveira et al. (2019) define data ecosystems as "socio-technical complex networks in which actors interact and collaborate with each other to find, archive, publish, consume, or reuse data as well as to foster innovation, create value, and support new businesses" (p. 589). Propensities of data ecosystems typically include intentionality, value creation, dynamism, demand–supply relationships, and sustainability as well as evolution (Harrison et al., 2012). According to Heimstädt et al. (2014) actors may play the following roles in a data ecosystem: data suppliers, data intermediaries, and data consumers. Following Harrison et al. (2012) that mention an additional role of a central actor "taking the initiative […] to achieve specific goals" (p. 907), this initiating and coordinating actor is termed convener of the data ecosystem by Susha et al. (2023).

As data ecosystems are sociotechnical systems (Bartolomucci and Bartalucci, 2024), both technical and social elements are required for their functioning. Technical elements that are needed to facilitate data sharing within the ecosystem include data sources, the presence of a (data) infrastructure, (digital) resources, services and applications, software, and standards (Shin and Choi, 2015; Oliveira et al., 2018, 2019). Social and organizational elements are the social, cultural, legal, and ethical context of a data ecosystem. Moreover, participants that share data and broader (societal) stakeholders play an important role within the data ecosystem. Additionally, the skills and expertise of participants and stakeholders are of importance (Zeleti and Ojo, 2017; Oliveira et al., 2019). Between all these elements, interactions and dependencies can be identified.

Data ecosystems are often embedded in a specific context (Harrison et al., 2012). Depending on the purpose, scope, and level of maturity of the data ecosystem, the way organizations work together can vary (Lis and Otto, 2021). It is becoming more common that data ecosystems are set up with the aim to create, store, and share data to support policy or decision making more structurally (Lanza, 2021). Data ecosystems are considered to be more sustainable than interorganizational information sharing (Susha and Gil-Garcia, 2019) and may evolve intentionally from data collaborations (Susha et al., 2023).

Literature has identified success factors and barriers for data ecosystems. These include technical barriers such as the lack of technical knowledge and resources to maintain the ecosystem (Dawes et al., 2016; Oliveira et al., 2019; Diran and van Veenstra, 2020). Moreover, there are also a number of organizational and network barriers that are of influence to the success of the data ecosystem, such as the high complexity of tasks, lack of actor participation and interaction, a lack of organizational structure,

https://doi.org/10.1017/dap.2025.10018 Published online by Cambridge University Press

unclear division of roles and responsibilities, data literacy of stakeholders, data ownership, fragmentation of data, and data privacy (Oliveira et al., 2019; Acar et al., 2021). Also, "lack of clarity in terms of roles and responsibilities and turf issues (non-cooperation between organizations with seemingly common interests) have been identified as important challenges" (Susha and Gil-Garcia, 2019, p. 2899). It is yet unknown whether primarily technical or organizational factors have a more significant role in the success of data ecosystems that address societal challenges, and how these play out in practice.

3. Method

3.1. Research design and setting

This article describes findings of a comparative case study of two data ecosystems based in the Netherlands. In general, the Netherlands is considered as one of the front runners in the Digital transformation. In 2022, the Netherlands scored third in the general Digital Economy and Society (DESI) index and fourth in the category *digital public services* in the EU (European Commission, 2022). This makes the Netherlands an appropriate setting for conducting a study on success factors of data ecosystems.

3.2. Cases

Cases were selected for this study on the basis that the data ecosystem (1) is aimed at addressing a societal challenge, (2) involves participants from various governmental organizations, and (3) has a relatively high level of maturity. We included two cases that address different challenges, namely sustainability and public safety. The first data ecosystem is a program titled "Improving Information Provision for Energy Transition" ("Verbetering Informatievoorziening Energietransitie" in Dutch; abbreviated as VIVET), which is focused on improving the availability of data for the energy transition at a national level. The second data ecosystem is "View of Subversive Crime" ("Zicht op Ondermijning" in Dutch; abbreviated as ZoO), which is focused on increasing the availability of data for the prevention subversive crime, that is, illegal activities that undermine legal structures and have disruptive effects on society.

3.2.1. Case A: National energy transition (VIVET)

VIVET originated in 2019 out of a demand for high-quality data to improve the information provision for the Dutch energy transition. These data would help to substantiate and implement the visions, plans, and strategies for the shift toward a carbon-free energy system. Participants in this data ecosystem can roughly be grouped into three categories. First, there is a user panel, which is made up of parties that experience the need for accessible and reliable data to carry out the task of the energy transition. This includes municipalities, represented by the Association of Netherlands Municipalities (VNG), provinces (IPO), and Regional Energy Strategy regions. Second, there are national actors that manage information that can be insightful for the energy transition. These can be labeled as the data management parties and include Statistics Netherlands (CBS), Netherlands Environmental Assessment Agency (PBL), Netherlands Cadastre, Land Registry and Mapping Agency (Kadaster), Netherlands Enterprise Agency (RVO), and the executive agency of the Ministry of Infrastructure and Water Management (Rijkswaterstaat). Third, there are the financing parties. VIVET is financed by the Dutch Ministry of Economic Affairs and Climate Policy (EZK) and the Ministry of the Interior and Kingdom Relations (BZK) as it is considered of national importance to optimize the provision of information.

VIVET addresses two main obstacles in the information provision. First, VIVET focuses on solving data gaps. For example, there is a need for low-regional information about the generation of solar or wind energy, and the infrastructure of, for example, electricity grids. The data management parties collect a lot of data, but the required data are fragmented, uncoordinated, and cannot always be found. Second, VIVET tackles technical obstacles. For example, the data come from different registers which were not originally set up with the aim of monitoring the energy transition. Therefore, there are bottlenecks that need to be solved before these registers can be linked.

https://doi.org/10.1017/dap.2025.10018 Published online by Cambridge University Press

3.2.2. Case B: Subversive crime (ZoO)

The initiative Zicht op Ondermijning (ZoO) started in 2017. The goal of ZoO is to make patterns visible and gather insights that contribute to tackling subversive crime and thereby to contribute to making society resilient for the undermining effects of organized crime. Different types of parties are involved such as several Dutch municipalities, national government actors such as the Ministry of Interior and Kingdom Relations, Ministry of Finance, Ministry of Justice and Security, Dutch Tax Authority, Netherlands Public Prosecution Service, Police, Statistics Netherlands (CBS), and research institution Leiden University. ZoO focuses on data accessibility aimed at preventing undermining criminal activities. The network has created a dashboard that gives insights into undermining activities within regions of the Netherlands. The core of the dashboard consists of data from CBS, which is complemented with data from government organizations.

In practice, there are a number of challenges that make it difficult to exchange data within the public safety domain. This not only counts for specific use cases, but also when a broader analysis of data sources takes place. Challenges are both technical, such as the availability and quality of data, and legal, such as that legal frameworks are not always easy to apply in large partnerships. As part of ZoO, a number of municipalities set up pilot projects to examine the possibilities and limitations of data analysis. An important output of ZoO is a publicly available dashboard.[1] This dashboard presents a number of indicators that provide insight into local criminal phenomena. Data used in the dashboard come from Statistics Netherlands (CBS), municipalities, Kadaster (The Netherlands' Cadastre, Land Registry and Mapping Agency), and the Financial Intelligence Unit (FIU).

3.3. Procedure, data sources, and instruments

The selected cases were analyzed by a combination of desk research and semistructured interviews with participants in the data ecosystem. For desk research, we analyzed program reports and other available documentation about the data ecosystems. This resulted in a document with a rich case description about both data ecosystems. Furthermore, we conducted interviews with three participants per case (six interviews total) who were each employed within a different (governmental) organization. The interviews with the participants were conducted between October and December 2022 and were 1-h long. The interviewees had different roles, such as program manager (2x), director of information provision, researcher, and policy advisor (2x).

During the interviews, an interview protocol was used that consisted of two sections. In the first section, the protocol included questions aimed at gaining a better understanding of the data ecosystem (e.g., its goals, status, and participants), and the respondent's role in this ecosystem. The second part focused on the aspects that played an important role in the success of the data ecosystem, aspects which were more challenging for the data ecosystem to navigate, and lessons learned for future data initiatives. Our interview guide was foremost set-up to give participants room to highlight the aspects that they personally deemed as the most important challenges or successes. Later on in the interview, we guided the conversation toward three factors that are known to impact data collaborations, namely technical aspects (data, infrastructure, etc.), stakeholder management, and collaboration practices. During the interviews, we made detailed notes, and the interviews were recorded so that we could transcribe passages of the interview for direct quotes.

3.4. Data analysis

We used an inductive approach for analyzing and coding the case descriptions and interview transcripts across the cases, following the principles of thematic analysis (Braun and Clarke, 2006). First, we started with open coding, where we created initial codes for all sections of the interview where respondents spoke about success factors, challenges, and lessons learned. Next, we analyzed these initial codes to search for

[1] https://www.zichtopondermijning.nl/

https://doi.org/10.1017/dap.2025.10018 Published online by Cambridge University Press

themes. In this process, we iteratively created lists of preliminary themes in one interview, which we then crosschecked across other interviews. After this review, we started to define and name the themes. When this was completed, we began writing out narratives, to present the story that emerged from the data in a compelling manner. During this step, further refinement of the themes took place and the analysis was completed.

4. Findings

We find that the main challenge for intergovernmental data ecosystems aimed at addressing societal issues lays in shaping successful collaboration among network actors. The success of the data ecosystem is to a large part dependent on the extent to which parties in the data ecosystem are able to find common interests, decide which actions get prioritized, and establish ways to work together. As one of the interviewees noted, the key challenge of sharing data within an ecosystem does not have so much to do with the technical aspects, but "it's about all the fuss surrounding it" [V1]. Although technical barriers are experienced, there is almost always a solution that can be found as long as the collaboration works well. In this section, we untangle the fuss surrounding data sharing and present five narratives that describe how this was successfully navigated in the two analyzed cases. These findings are also summarized in Table 1.

4.1. Use the policy question as guide to articulate data demands

> "Very good demand articulation is necessary. [Be clear about] what do you need and why, and what attribute or specifications you need. For example, people say 'I need information about heat.' What do you want to know about heat and for what purpose? Why do you need that data and how do you want to use it?"—V1.

Both cases experienced that it is important to have a central demand or policy question at the heart of the data ecosystem that is derived from the societal issue. This policy question can serve as the glue that holds the data ecosystem together and guides the next steps that need to be taken. As a respondent notes: "data must be at the service of the social task" [V3], and having a clear research goal helps to define what data (sources) are necessary to gain the required insights that can bring the social task forward. Working with this approach requires a reversal from a supply driven to a demand driven mindset. It was stated that when the demand is not clearly articulated, collaborating parties will commonly ask for as much data as they can possibly receive, in which case, the availability of the data becomes leading in the development of new knowledge. However, when you spend more time to figure out which question people wish to answer, it becomes evident that certain data are not needed. A demand driven approach serves as a filter for assessing which of the data sources are relevant to use and provides an incentive to go the extra mile to try and unlock data that are not readily available but would provide valuable observations.

The ability to articulate a clear question is a capability that both data ecosystems cultivated over time. In ZoO, it was first experienced that "it was not possible to conduct the data analysis in one or two weeks. […] As a result it became clear that a big question first needs to be narrowed down and further articulated before you know which data source you can use to answer the question" [Z1]. They found that there are often several questions that arise at the same time, which must be brought together. Because of this experience, ZoO established a way of working to do so, where they gather input from all partners to formulate initial research questions, and then hold thematic meetings to refine and prioritize the questions. Once this is completed, it is explored how feasible it is to do the data analysis and afterward data preparation is started. ZoO found that this process helps to make it easier to translate a question toward data analysis.

VIVET also found that it was a challenge to formulate good questions. One of the complicating factors was that, at the onset of VIVET, the energy transition field was very new and dynamic and many things were yet to be explored and defined. A tool that they experienced to be helpful is a vision document. Such

https://doi.org/10.1017/dap.2025.10018 Published online by Cambridge University Press

Table 1. *Summary of the five narratives*

Narrative	Description	Advice
1. Use the policy question as guide to articulate data demands	A central policy question derived from the societal issue is crucial for unifying and guiding a data ecosystem.	First, gather input from all partners to formulate initial research questions, and then hold thematic meetings to refine and prioritize the questions. Afterward, the feasibility of the data analysis can also be taken into account when deciding which questions to answer first.
2. Invest in a common language and shared definitions	Multidisciplinary experts from various organizations have to collaborate in data ecosystems that focus on societal challenges. Mutual understanding between these experts is essential and requires effort.	Ask detailed questions about what various experts mean when they use specific terms. Make definitions of important concepts explicit, if possible, with the help of an independent project leader.
3. Secure capacity through managerial support	Managerial support is an essential factor that impacts whether enough experts with the necessary skills are available for long-term collaboration in the data ecosystem.	All collaborating partners should keep up good relations with their higher management, to sustain the sense of urgency of both the societal issue and the collaboration in the data ecosystem at this level.
4. Foster collaboration through shared interests and expectation management	Collaborating partners need to respect each other and be willing to sometimes prioritize the interests of the collaboration and the societal issue over the interests of their own organization. Interpersonal relationships must be managed well.	Dedicate a work session to providing clarity about the collaborating partners' expectations, interests and ways of working, to signal and resolve any mismatches at an early stage in the collaboration. When necessary, come back to these aspects in interim reviews to keep improving interpersonal relationships.
5. Organically create structured ways of working	It provides clarity when ways of working within the data ecosystem are formalized. However, let these methods of working together arise naturally between the collaborating partners, rather than imposing a preexisting method.	Do not start by implementing a top-down structure on the collaboration. Rather, give experts room to jointly arrive at a way of working that bridges the gap between their various ways of working, and solidify this method afterward.

a document paints a picture of how societal partners, in this case decentralized governmental bodies such as municipalities, envision the future. From there, it is possible to distill data demands and work backward. In general, VIVET experienced that reasoning along the lines of a societal challenge helped to get things moving and made it easier to establish data sharing.

https://doi.org/10.1017/dap.2025.10018 Published online by Cambridge University Press

4.2. Invest in a common language and shared definitions

The complexity of data projects and societal challenges requires that policy and domain expertise, data expertise, and legal expertise—across several institutions—come together in the initiative. This indispensable collaboration between multidisciplinary actors in a data ecosystem makes it a necessity to invest time and energy in making sure that all participants understand each other well. Interviewees remark that people with data expertise approach things differently than policymakers. It is therefore not always easy to establish a common language: "Dedicate a lot of attention to what are we exactly talking about? How do we define crime? When you zoom in it quickly becomes more complicated" [Z2]. It is recommended to make room to ask and answer such detailed questions, because, as an interviewee warns, otherwise there is a chance that information is produced that people do not agree on and that everyone interprets in their own way.

Investing in shared definitions is not only necessary so that participants have the same understanding about what a concept entails, but it also impacts whether data can be connected and compared. One interviewee presented a detailed account of how the process for establishing shared definitions took place in the VIVET case, to harmonize data about solar energy and wind energy on land. Several parties publish relevant information about these concepts, but their data are not always the same. This causes a problem because it leads to discussions about the validity of the different numbers, rather than about the actions that should be taken on the basis of these numbers. VIVET chose to appoint an independent project manager to facilitate the harmonization of these data, because all other parties had a particular interest in the outcome. The process was started by taking stock of the various definitions that were used by the different parties and for what reason. In some cases, there was a good argument for using a particular definition, for example, because this was a requirement based on European legislation. This inventory was used for a kind of negotiation between the different parties. A proposal was made for which definitions would be best to use, based on how well the definitions matched with policy documents. Then there was a discussion about whether or not it would be possible to use these definitions and what advantages or disadvantages this choice entailed. From that, a conceptual framework was created, which was validated over the course of a year. This way of working was appreciated and they want to return to it in the future when they need to harmonize new concepts.

4.3. Secure capacity through managerial support

> "You need to be within means, with some money of course […] and the agreements at management level, to actually do something about it. […] It never, or almost never, depends on the willingness of people. But the moment you come to management and they say 'Yes, but this has to be done, but there is no money for it and there is no time for it' - then it becomes difficult."—V2.

To be able to bring together all the necessary expertise in a data ecosystem, sufficient capacity, that is, the availability of experts with required skills, must be secured. Managerial support is highlighted as a crucial factor that impacts whether this capacity is available. As an interviewee states: "It is important to manage expectations and secure managerial support, and therefore also the required resources. Freeing up people in the organization who can participate. [..] Without this managerial support, no extra capacity will be created. Without created capacity, you miss out on quality and expertise and you are in fact dependent on yourself, which is too weak for such a major development" [Z1]. When sufficient capacity and budget are available, it not only makes collaboration possible, but it also becomes easier for a data initiative to solve technical challenges, such as transforming data, improving data quality, and keeping data up-to-date.

Data ecosystems addressing societal challenges are vulnerable to fluctuations in managerial support, because whether societal issues remain high on administrative agendas is dependent on various factors, such as the political climate, notable events, and whether new urgent issues emerge. ZoO experienced that when the initiative was started, the topic of subversive crime was high on the agenda. However, as time passed, it slowly faded away and other issues were started to be seen as more important. This made it hard

https://doi.org/10.1017/dap.2025.10018 Published online by Cambridge University Press

to keep people involved who initially said yes or signed an agreement, and had a direct impact on capacity in the data ecosystem: "maintaining administrative commitment is a very difficult one and it has had an impact on project leaders, who were overshadowed by their department heads with more work. Many project leaders [indicate that they] would like to be able to continue contributing, but just don't have the time for it anymore, and do it more on the side. [… It is] very sad because most project leaders have good will, are really interested in this subject, […] but often do not have the time to seriously devote time to this" [Z1].

4.4. Foster collaboration through shared interests and expectation management

This narrative emphasizes the importance of managing the interpersonal relations within the collaboration. It is crucial for fruitful collaboration that people develop respect with and for each other. Especially when working together with different parties and different types of experts, it is necessary that people understand each other and respect each other's varying positions. It helps when it is made evident what people's expectations and interests are of participating in a data-sharing program, so that all participants know why certain parties are involved and what their priorities are.

The ZoO case illustrates how expectation management is part of this process. As soon as it becomes clear that parties are involved in the data ecosystem with expectations that deviate from what is previously been decided or what is in scope, this should be addressed. When the positions of the partners are clear, it can become evident that work needs to be done to align the interests. In ZoO, there was some frustration at the beginning of the collaboration because some parties had made inaccurate assumptions about the way of working of another party. An interim review and separate work session were organized to resolve this issue, with the goal of providing clarity about the interests of the various collaborating parties. Spending time on this subject created stability and mutual respect.

A successful way in which ZoO managed to create shared interests is that when a question was articulated and selected for further analysis, they took the approach of answering the question in such a way that more stakeholders could use the data. To illustrate this, if one municipality posed a question, the analysis was done for every municipality in the Netherlands. This led to new insights for the entire data ecosystem working on tackling subversive crime and not just for one specific stakeholder. Interviewees see this as one of the main success factors of the ZoO data ecosystem. Other factors that made collaboration a success and were briefly mentioned by interviewees are that working groups were kept small and that it was decided not to keep pushing when the collaboration met resistance.

Because the interest of the collaboration sometimes needs to be prioritized over the interest of the participants' institution, flexibility of participants is required. This is emphasized by V2: "An important lesson learned is the flexibility that's necessary […] How you can prioritize the interests of the collaboration over the priorities of your organization" [V2]. VIVET observed that parties were sometimes reluctant to share their data, but the urgency of the energy transition helped participants to take on a more flexible attitude than they might otherwise. Putting the societal challenge at the center helps people to see sufficient need for a particular activity so that they are willing to invest in it. "It is mainly about trust and that parties reason from the perspective of the social task and no longer from their organization. Then things start moving" [V1]. This sense of urgency can best be created in small work groups and then flow upward, as is elaborated further in Section 4.5. Moreover, an interviewee noted that flexibility is about allowing yourself space for learning together and keeping ego's at bay. "You have to allow yourself to say 'ok, that didn't go so well, but now we will go in this direction'" [V2].

4.5. Organically create structured ways of working

In order to create more structure, both data ecosystems established a system to formalize the ways of working and governance of the data ecosystem. For example, an interviewee from VIVET mentioned that "we do not share data on a physical platform. We work on the organizational conditions that make it possible to share data" [V1]. Moreover, interviewees underlined that clarity about organizational structure

https://doi.org/10.1017/dap.2025.10018 Published online by Cambridge University Press

is necessary. However, one interviewee also stated that "the organizational structure should emerge in an organic way that appeals to all participants" [V2]. This is how VIVET emerged from a bottom up manner. They realized that improving the information provision for the Energy Transition initiative would not start top-down. Therefore, the parties involved in VIVET saw the need to do this themselves.

One aspect that can contribute to achieving structure is that someone takes up a convener role and coordinates between different participants of the ecosystem. This role was present both within VIVET and ZoO. In VIVET, there was a program manager appointed that coordinated between different participants of the ecosystem. In ZoO, a ministry took on a facilitating role and there were two project leaders. Moreover, both data ecosystems established a formal meeting structure with different moments for meetings on different management levels, such as an interdepartmental meeting and an intergovernmental meeting. In ZoO, the core team worked in an agile manner.

Furthermore, as part of structuring processes in the data ecosystem of ZoO, a process description was developed. The process description consists of the 13 steps that detail in each step which type of roles should be involved. Important steps in this process are: (1) demand articulation, such as the prioritization and sharpening of questions and examining the feasibility of answering questions with required and available data; (2) data preparation and analysis; (3) validation of the analysis and output; (4) publication of the results and underlying data. This process description helped to create clarity on when and which stakeholders need to be involved in the process and made the established way of working explicit by outlining when certain steps need to be undertaken.

5. Discussion and conclusion

In this translational article, we examined two cases of data ecosystems in the Netherlands that are aimed at addressing societal issues and presented five narratives that cover aspects that lead to successfully establishing and maintaining such an ecosystem. For government officials, whether in a policy or technological role, these lessons from other initiatives can provide valuable insights about collaboration in data ecosystems, the complexities that are faced, and how these can be addressed. The narratives are summarized in Table 1.

In these five narratives, the importance of including diverse perspectives as well as working bottom-up is repeatedly emphasized. The cases illustrate that this increases the likelihood that data are unlocked that holds the right information to answer the policy questions at hand, thereby reducing knowledge uncertainty (Klijn, 2002). At the same time, the cases highlight how the ecosystem is both in need of a managerial structure (Oliveira et al., 2019) as well as that it is dependent on managerial support for the resources to keep functioning. What data ecosystems can do to sustain this vital managerial support, even when societal issues lose prominence on the agenda, is an open question.

Another overarching lesson is that it is vital to have extensive interaction between network partners that cover a broad range of expertise and operate at different organizational layers at the onset of a data ecosystem, in order to invest in good working and interpersonal relations. This creates an environment where work structures, network goals, and roles in the ecosystem (Heimstädt et al., 2014; Susha et al., 2023) can organically emerge. This also ensures that experts are able to identify and address technical or legal barriers for connecting and comparing data at an early stage. This is in line with the finding of Klievink et al. (2018) that "the reinforcement process between collaboration and trust […] resulted in institutionalisation of the collaboration itself" (p. 392), which may lead to reduced institutional uncertainty (Klijn, 2002).

We thus find that, in practice, organizational capabilities are experienced as very important for the success of data ecosystems. Moreover, our results illustrate that focusing on organizational capabilities can also advance technological data governance aspects, showing that organizational and technical capabilities are sometimes closely related. This is line with Bartolomucci and Bartalucci (2024) who point out "the interconnection and interdependence between collaborative governance dimensions and technological ones" (p. 19). For example, it was stressed that collectively formulating a good policy question aids in data collection, as it outlines the data demands more clearly and provides an incentive to

https://doi.org/10.1017/dap.2025.10018 Published online by Cambridge University Press

work harder to overcome barriers to unlock specific data. This is in line with what van Veenstra and Kotterink (2017) found in relation to data-driven policymaking. Additionally, data quality can be enhanced through the process of creating shared definitions. This helps to avoid that data are combined that does not completely match and puts forward more precisely what criteria the data need to meet to be useful. Finally, establishing shared interests between ecosystem partners also contributes to data collection, as it improves the willingness of parties to share their data (Susha, 2020).

The narratives point to several directions for further research. First, it is worthwhile to expand on how data ecosystems combine the varying dynamics of bottom-up formation on the one hand, and sustainable embedding in existing organizational structures on the other. This could be done by establishing a connection between literature streams, including organization studies and organization behavior. Second, it is necessary to further explore the apparent contradiction of data ecosystems needing to focus on a joint specific question, while there is also the need for generic data ecosystems that may be used to address several adjacent or domain focused societal challenges. The development of structured ways of working should be explored further to find common and generalizable elements that may help establish such ecosystems. Third, it is valuable to expand the relationship between the technical and organizational capabilities of ecosystems. From a sociotechnical perspective, it is relevant to not only identify technical and organizational capabilities, but explore whether new capabilities emerge, for example, building on expertise from data science as well as policymaking. Similarly, it would be worthwhile to build on the interrelation between information systems and political science, to investigate questions of power in relation to storage and control over data.

Given that this article is only focused on two cases in a specific country setting, we want to end by reflecting on the impact this might have had on our findings. The Netherlands has a specific administrative climate that allows for room for discussion, is not very authoritative, and relatively cautious in sharing data. It is possible that in the context of another country, more aspects could be managed top-down. For example, in a more authoritative context, narrative three concerning managerial support might be of greater importance, whereas narratives four and five about expectation management and organically creating structured ways of working may resonate a little less. Second, while it was a research goal to focus on intergovernmental ecosystems, we recognize that the studied cases are indeed heavily oriented toward the public sector and did not include partners from private organizations. Generally, when looking at cross-sectoral data sharing, additional concerns that come into play focus on data "ownership" and a stronger incentive to "monetize" data (Susha et al., 2023), for example, compared to value creation in the form of better policy. This may reduce generalizability of these findings in relation to data collaboratives in which private organizations are involved. Third, both examined initiatives set out to make data available at the level of statistics and had access to a safe computing environment of Statistics Netherlands (CBS). More technical lessons learned might have come up in the context of sharing privacy sensitive microdata, or when a different environment needs to be set-up for computations. Fourth, the studied cases were both examples of successful data ecosystems. It could also have been interesting to study data ecosystems that failed to highlight mistakes to be avoided in the future.

In conclusion, data ecosystems that address societal issues operate in a complex environment that is characterized by uncertainty about knowledge, collaboration that extends over existing organizational networks, and values and interests of network actors that may be in conflict. In this translational article, we outlined how two cases of successful data ecosystems in the Netherlands navigated these complexities. We found that both cases were successful in a large part because of how they established collaboration among network actors, using bottom-up approaches. The cases created structures in which participants were able to prioritize the right questions, find common interests, and work together. This stresses the importance of investing in organizational capabilities of data ecosystems, if they are to address the societal challenges of our time.

Data availability statement. We are not able to share data collected from the interviews.

Acknowledgements. This article derives from an abstract accepted into Data for Policy 2024. We thank all interviewees for sharing their insights and we thank the reviewers for their feedback that helped to improve our work.

https://doi.org/10.1017/dap.2025.10018 Published online by Cambridge University Press

Author contribution. Conceptualization: A.B., M.H. (equal), A.F.V. (support). Data curation: A.B., M.H. (equal). Formal analysis: A.B., M.H. (equal). Funding acquisition: A.B., A.F.V. (equal). Investigation: A.B., M.H. (equal), A.F.V. (support). Methodology: A.B., M.H. (equal), A.F.V. (support). Validation: A.B., M.H. (equal). Writing—original draft: A.B. (lead), M.H., A.F.V. (support). Writing—revision: A.B., M.H., A.F.V. (equal). All authors approved the final submitted draft.

Funding statement. This work was funded by the Dutch Ministry of the Interior and Kingdom Relations.

Competing interests. The authors have no competing interests to declare.

Ethical standard. The research meets all ethical guidelines, including adherence to the legal requirements of the study country.

References

Acar F, **Raes L**, **Rosseau B and Satta M** (2021) *Making Policies with Data: The Legacy of the PoliVisu Project*. Cham, Switzerland: Springer International, pp. 105–122.

Alford J and Head BW (2017) Wicked and less wicked problems: A typology and a contingency framework. *Policy and Society 36*, 397–413.

Bartolomucci F and Bartalucci S (2024) Designing data collaboratives' governance dimensions for long-term stability: An empirical analysis. *Data & Policy 6*, 1–23.

Braun V and Clarke V (2006) Using thematic analysis in psychology. *Qualitative Research in Psychology 3*, 77–101.

Danemayer J, **Young A**, **Green S**, **Ezenwa L and Klein M** (2023) Responding to the coronavirus disease-2019 pandemic with innovative data use: The role of data challenges. *Data & Policy 5*, 1–15.

Dawes SS, **Vidiasova L and Parkhimovich O** (2016) Planning and designing open government data programs: An ecosystem approach. *Government Information Quarterly 33*, 15–27.

Dietrich S, **Malerba D and Gassmann F** (2024) Predicting social assistance beneficiaries: On the social welfare damage of data biases. *Data & Policy 6*, 1–28.

Diran D and van Veenstra AF (2020) Towards data-driven policymaking for the urban heat transition in the Netherlands: Barriers to the collection and use of data. *In 19th International Conference on Electronic Government (EGOV)* pages, 361–373.

European Commission (2022) *Digital Economy and Society Index (DESI) 2022*.

Giest S (2017) Big data for policymaking: Fad or fasttrack? *Policy Sciences 50*, 367–382.

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

Head BW (2022) *Wicked Problems in Public Policy: Understanding and Responding to Complex Challenges*. Palgrave Macmillan.

Heimstädt M, **Saunderson F and Heath T** (2014) From toddler to teen: Growth of an open data ecosystem. *JeDEM 6*, 123–135.

Kersing M, **van Zoonen L**, **Putters K and Oldenhof L** (2022) The changing roles of frontline bureaucrats in the digital welfare state: The case of a data dashboard in Rotterdam's work and income department. *Data & Policy 4*, 1–18.

Klievink B, **Voort HVD and Veeneman W** (2018) Creating value through data collaboratives: Balancing innovation and control. *Information Polity 23*, 379–397.

Klijn EH (2002) Governing networks in the hollow state: Contracting out, process management or a combination of the two? *Public Management Review 4*, 149–165.

Lanza G (2021) *Data-Related Ecosystems in Policy Making: The PoliVisu Contexts*. Cham, Switzerland: Springer International.

Lis D and Otto B (2021) Towards a taxonomy of ecosystem data governance. *Proceedings of the 54th Hawaii International Conference on System Sciences*, 6067–6076.

Liu X and Dijk M (2022) The role of data in sustainability assessment of urban mobility policies. *Data & Policy 4*, 1–20.

Martinazzo L (2023) A comparative analysis of public transport accessibility to hospitals in Córdoba (2019-2023): Where are we now? *Data & Policy 5*, 1–22.

Oliveira MIS, **Oliveira LERA**, **Batista MG and Lóscio BF** (2018) Towards a meta-model for data ecosystems. *dg.o '18: Proceedings of the 19th Annual International Conference on Digital Government Research: Governance in the Data 72*, 1–10.

Oliveira MIS, **de Fátima Barros Lima G and Farias Lóscio B** (2019) Investigations into data ecosystems: A systematic mapping study. *Knowledge and Information Systems 61*, 589–630.

Rij WV, **Dekker R and Meijer A** (2024) Integrating data, knowledge, and expertise for policy advice: Understanding the practices of Dutch organized crime control professionals. *Data & Policy 6*, 1–14.

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

Shin DH and Choi MJ (2015) Ecological views of big data: Perspectives and issues. *Telematics and Informatics 32*, 311–320.

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

Susha I and Gil-Garcia JR (2019) A collaborative governance approach to partnerships addressing public problems with private data. *Proceedings of the 52nd Hawaii International Conference on System Sciences*, 2892–2901.

Susha I, **Janssen M and Verhulst S** (2017) Data collaboratives as a new frontier of cross-sector partnerships in the age of open data: Taxonomy development. *Proceedings of the 50th Hawaii International Conference on System Sciences*, 2691–2700.

https://doi.org/10.1017/dap.2025.10018 Published online by Cambridge University Press

Susha I, **van den Broek T**, **van Veenstra AF and Linåker J** (2023) An ecosystem perspective on developing data collaboratives for addressing societal issues: The role of conveners. *Government Information Quarterly 40*, 1–18.

Urbano VM, **Bartolomucci F and Azzone G** (2024) Determinants for university students' location data sharing with public institutions during COVID-19: The Italian case. *Data & Policy 6*, 1–20.

van Veenstra AF and Kotterink B (2017) Data-driven policy making: The policy lab approach. *9th International Conference on Electronic Participation (ePart)*, 100–111.

van Veenstra AF, **Grommé F and Djafari S** (2020) The use of public sector data analytics in the Netherlands. *Transforming Government: People, Process and Policy 15*, 396–419.

Zeleti FA and Ojo A (2017) Open data value capability architecture. *Information Systems Frontiers 19*, 337–360.

Cite this article: van den Berg A, Hoekstra M and van Veenstra AF (2025). Government data ecosystems for addressing societal challenges: Lessons from two cases in the Netherlands. *Data & Policy*, 7: e79. doi:10.1017/dap.2025.10018

https://doi.org/10.1017/dap.2025.10018 Published online by Cambridge University Press