Data & Policy (2024), 6: e37
doi:10.1017/dap.2024.29

RESEARCH ARTICLE

Designing data collaboratives' governance dimensions for long-term stability: an empirical analysis

Federico Bartolomucci and Simone Bartalucci

Department of Management Economics and Industrial Engineering at Politecnico di Milano, Milan, Italy
Corresponding author: Federico Bartolomucci; Email: federico.bartolomucci@polimi.it

Received: 06 June 2023; **Revised:** 15 December 2023; **Accepted:** 13 May 2024

Keywords: collaboration; cross-sectoral partnership; data collaboratives; governance; ecosystems

Abstract

The momentum surrounding the use of data for the public good has grown over the past few years, resulting in several initiatives, and rising interest from public bodies, intergovernmental organizations, and private organizations. The potential benefits of data collaboratives (DCs) have been proved in several contexts, including health, migration, pandemics, and public transport. However, these cross-sectoral partnerships have frequently not progressed beyond the pilot level, a condition hindering their ability to generate long-term societal benefits and scale their impact. Governance models play an important role in ensuring DCs' stability over time; however, existing models do not address this issue. Our research investigates DCs' governance settings to determine governance dimensions' design settings enhancing DCs' long-term stability. The research identifies through the literature on collaborative governance and DCs seven key governance dimensions for the long-term stability of DCs. Then, through the analysis of 16 heterogeneous case studies, it outlines the optimal design configurations for each dimension. Findings make a significant contribution to academic discourse by shedding light on the governance aspects that bolster the long-term stability of DCs. Additionally, this research offers practical insights and evidence-based guidelines for practitioners, aiding in the creation and maintenance of enduring DCs.

Policy Significance Statement

The article advances the discussion on the governance of data collaborations at a theoretical level while offering practitioners concrete design principles. It demonstrates the role that public entities play in initiating and nurturing data collaboratives (DCs) over time and advocates for their engagement. The article can serve as a point of reference for designing DCs by public, private, and not-for-profit entities since it provides specific design guidelines and potential errors to avoid.

1. Introduction

In recent years, the growing awareness of the potential benefits of socially oriented data use (Verhulst and Young, 2016) has prompted practitioners, institutions, and high-level organizations to engage in numerous efforts to promote the use and reuse of data for the common good. Implementing data-driven initiatives for the social good not only requires data but also frequently necessitates a combination of complex assets and skill sets, including technical and socially oriented expertise that a single actor may rarely possess (Susha et al., 2019a). Two main perspectives have emerged: those advocating for the necessity of conceiving data as a common good, opening data freely for any use (Zuiderwijk and Janssen, 2014; Varshney and Mojsilovic, 2019; Nikander et al., 2020), and those advocating for the creation of cross-sectoral partnerships using data for the social good, embracing a "data as a club good" perspective (Savona, 2020; Global Partnership for Sustainable Development Data, 2023). In the second perspective, many have converged on the necessity of creating cross-sectoral partnerships that, by leveraging the experiences, knowledge, and assets of multiple actors, enhance the design and implementation of context-specific (Chignard and Glatron, 2023) and innovative solutions that would otherwise be unachievable by a single actor (Susha et al., 2022). These types of partnerships have assumed different forms and have been alternatively called data collaboratives (DCs; Verhulst and Sangokoya, 2015), data-driven social partnerships (Susha et al., 2019a), or local data ecosystems (Liva et al., 2023). Regardless of their names, these forms of partnerships, hereinafter referred to by the umbrella term DCs, share a few characteristics: a clear social purpose, a cross-sectoral nature, data analysis as the main value-creation activity, and the sharing of data among a restricted number of partners.

Despite DCs' increasing diffusion, proved by the growing number of such initiatives around the world (see the repository datacollaboratives.org) and the European Union (EU) regulatory effort included in the recent Data Governance Act,[1] their impact remains limited (Flanagan Anne and Sheila, 2022). A major limitation of DCs' impact-generation capacity is their struggle to progress beyond the pilot stage and to the production stage (Lapucci and Cattuto, 2021). Predominantly being in the form of small and one-off initiatives (EU Commission, 2020; Susha et al., 2022) hinders DCs' ability to perform impact-generation activities over time (GSMA, 2018; Flanagan Anne and Sheila, 2022), to demonstrate their impact, and to retain or entice the commitment of existing or new actors in the partnership. Creating stable DCs could allow them to develop standard practices of data sharing (e.g., agreements and data-sharing protocols) (The New Hanse Project, 2023), create empirical evidence of their results (Lapucci and Cattuto, 2021), and revitalize partners' motivation, thereby scaling their activities and the consequent impact, through either the expansion of their activities or the replication of their model in different contexts. DCs' inability to sustain themselves beyond the pilot stage is attributable to a number of factors, encompassing technological (Otto and Hompel, 2022), economical (Nikander and Elo, 2019; Charles and Tonetti, 2020; Nikander et al., 2020; Savona, 2020), ethical (Lepri et al., 2018; Noriega-Campero et al., 2019; Suresh et al., 2022), and organizational ones (van den Broek and van Veenstra, 2018; Bughin et al., 2019; Stalla-Bourdillon et al., 2019).

To prioritize key research questions for the field's development, Susha et al. (2018) established a research agenda. However, many of those questions remain unanswered, and a structured body of empirically tested knowledge supporting practitioners in their efforts remains lacking (Lapucci and Cattuto, 2021). Among the research questions put forth by Susha et al. (2018), few concern the governance aspects of these partnerships, which are essentially a set of coordinating and monitoring elements that ensure the survival of a partnership (Bryson et al., 2006) and are therefore directly connected to DCs' capacity to progress beyond the pilot stage. Since then, a few studies have addressed this issue from an organizational perspective (Susha and Gil-Garcia, 2019, Ruijer, 2021), while others have focused on data governance aspects (Bharosa and Janssen, 2015; Groves and Neufeld, 2015; Otto and Hompel, 2022). From an organizational perspective, DCs' characteristics, which include social purpose and data centrality, require adapting existing cross-sectoral partnership governance frameworks (Agranoff, 2006; Ansell and Gash, 2008; Provan and Kenis, 2008; Emerson et al., 2012; Bryson et al., 2015) to their specific logics and characteristics. Therefore, previous studies (Susha and Gil-Garcia, 2019; Ruijer, 2021) have focused on determining which alternative governance dimensions need to be considered when designing DCs' governance. Building on their findings and acknowledging the key role different governance dimensions' settings play in ensuring a partnership's long-term stability (Provan and Kenis, 2008; Susha et al., 2019a), we pose the following research question: *How do different governance dimensions' design settings favor DCs' long-term stability?*

[1] https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32022R0868.

https://doi.org/10.1017/dap.2024.29 Published online by Cambridge University Press

In answering this research question, this article contributes to identifying, through empirical evidence, which design settings of the analyzed governance dimensions favor the long-term stability of a partnership. This article thus expands the existing literature on DCs (Gil-Garcia and Sayogo, 2016; Stalla-Bourdillon et al., 2019; Susha and Gil-Garcia, 2019; Ruijer, 2021), moving from the identification of the governance dimensions to be analyzed to the identification of their best design settings that enhance a DC's long-term stability. By analyzing 16 case studies, this article also answers the call for more empirical evidence in the field (Lapucci and Cattuto, 2021), providing evidence that sustains our findings and previous ones. Moreover, this article has significant implications for practitioners, providing them with applicable design principles to strengthen the stability of their initiatives.

This article is structured as follows: Section 2 presents the literature systematization that led to the selection of the seven dimensions of analysis. Section 3 presents the applied methodology and briefly describes the analyzed cases. Section 4 illustrates the findings. Sections 5 and 6 discuss the results, draw conclusions, and suggest possible future research.

2. Theoretical background

2.1 DCs' governance

The literature on collaborative governance and that on DC has broadly addressed the need for governance schemas to make collaborations function and succeed (Provan and Kenis, 2008; Susha et al., 2019a). Few researchers (Susha et al., 2019a; Ruijer, 2021) have challenged existing collaborative governance models (Bryson et al., 2006; Thomson and Perry, 2006; Provan and Kenis, 2008; Emerson et al., 2012) in terms of DCs' uniqueness. Interestingly, Ruijer (2021) adopts Bryson et al.'s framework (2015), as a theoretical lens to develop her analysis. According to Bryson et al. (2015), whose work builds on other collaborative governance frameworks (Ansell and Gash, 2008; Emerson et al., 2012; Vangen et al., 2015), governance lies at the intersection of processes and structures. This interpretation permits transcending the formal division of processes and structures and recognizing the role of these elements at their intersection, such as leadership, technology, collaborative capacity, and competencies (Bryson et al., 2015), and the funding model (Bharosa, 2022). Ruijer (2021) defines a set of governance dimensions that must be considered when designing DC governance schemas. She identifies seven components (institutional environment, initial conditions and shared motivation, collaborative structure, collaborative processes and activities, leadership, tensions, and outcome) and three layers (organizational, political and policy, data and technical). The combination of the seven components and the three layers helps identify 40 different dimensions that must be considered when designing DCs' governance.

2.2 Governance and stability

Governance elements, such as collaborative processes and structures, determine a partnership's capacity to operate and make the governance arrangement an element determining its ability to operate in the long term (Agranoff, 2006). Long-term stability is recognized as a critical success factor concerning DCs (Susha, 2020) and a crucial area of focus for private–public data partnerships (Flanagan Anne and Sheila, 2022). It facilitates the creation of a collective identity, fostering both internal and external legitimacy (Provan and Kenis, 2008; Koschmann et al., 2012). Furthermore, it enables greater managerial efficiency (Provan and Kenis, 2008), mitigates the impact of start-up-phase costs (GSMA, 2018), and ensures continuity in social innovation and solution refinement (Kjaer and Vestergaard, 2002; Agranoff, 2006; Manning and Roessler, 2014; Sadabadi and Rahimi Rad, 2022). Stability also facilitates long-term and interorganizational planning (Agranoff, 2006; GSMA, 2018) and enables scaling impact through the expansion of activities or the replication of the model (GSMA, 2018; van Tulder and Keen, 2018). It fosters knowledge exchanges among partners and the establishment of a common knowledge repository (Agranoff, 2006; Manning and Roessler, 2014; GSMA, 2018). Finally, it aids partnerships in attracting external funding (GSMA, 2018).

By cross-analyzing literatures on collaborative governance (e.g., Thomson and Perry, 2006; Ansell and Gash, 2008; Emerson et al., 2012; Bryson et al., 2015), DCs' governance (e.g., Gil-Garcia and

https://doi.org/10.1017/dap.2024.29 Published online by Cambridge University Press

Sayogo, 2016; Klievink et al., 2018; Susha, 2019), and partnerships' stability (Kjaer and Vestergaard, 2003; Manning and Roessler, 2014; Sadabadi and Rahimi Rad, 2022), we have identified, among those cited by Ruijer (2021), seven governance macro dimensions that are consistently considered in the literature to be related to DCs' long-term stability. These are elements that, despite influencing partnerships' long-term stability, may also refer to a partnership's initiation phase (e.g., leadership or trust), implementation (e.g., formal structures or the intermediation schema), or long-term operations (e.g., impact measurement). However, the boundaries among the different phases are indistinct, and as further discussed, one element may need to be addressed and managed differently according to the development phase of a partnership.

The first elements to consider are the *initial conditions* (Bryson et al., 2006; Ansell and Gash, 2008) constituted by a complex articulation of conditions defining the social context in which a collaboration begins or the immediate preconditions that influence the formation of a collaboration (Bryson et al., 2015). Among the elements that define the initial conditions (e.g., authoritative texts and recognized interdependence) (Bryson et al., 2015), we have focused on the concepts of interdependence and leadership. The first refers to the involved actors' realization that addressing complex societal challenges requires collective action (Bryson et al., 2006). Interdependence in DCs is exemplified by how the various resources and competencies that drive a collaboration are often spread across multiple actors and industries. Leadership refers to the presence of a formal or informal leader who can initiate and help secure resources and support for a collaboration (Emerson et al., 2012). The leader acts as a boundary-spanning agent (Agranoff, 2006; Thomson and Perry, 2006) and engages and facilitates the interaction of other parties (Susha et al., 2022). The two previously mentioned dimensions are connected to the concept of *trust*, which affects the different phases of initiation, development, and long-term stability of a collaboration (Agranoff, 2006; Bryson et al., 2006; Thomson and Perry, 2006; Provan and Kenis, 2008). Trust is not a static concept, nor is it always reciprocal; rather, it is frequently built during the collaborative timeframe through dedicated trust-building activities (Bryson et al., 2015). The mutual reinforcement effect between trust and collaborative activities suggests the possibility of a catalyst effect (Klievink et al., 2018) and its influence on the governance structure (Provan and Kenis, 2008). In addition to the standard trust logics of collaborative partnerships, the concept of trust in DCs depends on data-related activities (Stalla-Bourdillon et al., 2019), as well as continuous communication and engagement cycles (Farmer et al., 2023).

Incentives comprise another important dimension that is relevant both at the beginning and during the implementation phase of a partnership. These are contingent on actors' expectations in relation to the time and effort required for collaboration (Ansell and Gash, 2008) and the risk associated with data exploitation (Klein and Verhulst, 2017). How to motivate private actors to share their data is one of the primary incentive challenges in the context of DCs (Susha et al., 2019a). While public actors are motivated by an inherent mission to promote the common good and by public pressure (Gil-Garcia and Sayogo, 2016), private actors should be given distinct incentives. Klein and Verhulst (2017), supplemented by Moretti et al. (2022), identify three primary categories of incentives: knowledge and insight incentives, brand equity, and license to operate.

Formal structures, which refer to the formal rules that partners have agreed upon, are a relevant factor related to the governance structures of a partnership. This is the only identified dimension that reflects the regulatory function of collaborative governance (van der Voort, 2017). Formal structures frequently take the form of formal agreements, making it possible to influence efforts to attract the required resources, formalize the participation of various parties (Koschmann et al., 2012), and forge connections between actors (Bryson et al., 2006). Formal agreements are especially relevant with regard to sharing data and data infrastructures, as they define the connected responsibilities of each actor (Stalla-Bourdillon et al., 2021) as well as data auditability and accountability (van Donge et al., 2022). In the long term, agreements can affect the adaptability and flexibility of a governance structure by restricting, regulating, or incentivizing the entry of new actors (Ruijer, 2021).

Formal agreements may also define the *intermediation* model; however, given its influence on collaborative stability, this has been considered a separate dimension of analysis. The intermediation model may adopt different grades of complexity. In its less structured forms, intermediation may be

https://doi.org/10.1017/dap.2024.29 Published online by Cambridge University Press

employed by one of the partners and entails facilitating the exchange of information between the actors involved; at its most complex level, it may imply decision-making, including on behalf of the partners (Flanagan Anne and Sheila, 2022), and revenue generation (Susha et al., 2020). In most formal DCs, facilitation may be performed by ad hoc organizations acting as intermediaries (Digital Civil Society LAB, 2017; Perkmann and Schildt, 2015; Stalla-Bourdillon et al., 2021). In the case of DCs, these organizations may also serve as data stewards (Verhulst et al., 2020).

Although the literature agrees on the potential benefits that the presence of an intermediary may bring to a partnership, there is no consensus on how an intermediary should sustain its operations (Stalla-Bourdillon et al., 2021; Susha et al., n.d.). However, several researchers (e.g., Martin et al., 2021; Susha et al., 2022) recognize DCs' capacity to economically support their activities as a crucial factor in ensuring a partnership's long-term stability (Smichowski, 2019; Flanagan Anne and Sheila, 2022; Micheli et al., 2023). In our research, we refer to this capability as the *business model*, using the term in a comprehensive sense to encompass all the potential income sources a collaborative can generate, ranging from traditional research grants to revenue generation. The GSMA (2018) outlines four major benefits a well-structured business model could provide: proven solutions can be improved and scaled up; knowledge sharing and mutual learning can be increased; funding can be secured; and the value generated in terms of social, economic, and reputational value can be increased.

Finally, the literature agrees on the importance of governance configuration embracing a certain level of *flexibility* over time. Flexibility refers to partnerships' adaptability to internal and external pressures for change. Some of these tensions may be caused by the structural elements of a partnership, such as the diverse interests driving partners to join a collaboration, or power imbalances, which are exacerbated when partners are highly heterogeneous in terms of size, funding, or reputation (Bryson et al., 2006). An important governance decision in this regard is how to manage the trade-off between inclusiveness in collaborative decision-making and administrative efficiency (Bryson et al., 2006). On the one hand, inclusiveness is necessary for the legitimacy and effectiveness of a collaboration; however, on the other hand, "the more organizational participants are involved in the network decision process, the more time-consuming and resource-intensive that process will tend to be" (Provan and Kenis, 2008).

Table 1 summarizes the aforementioned seven dimensions of analysis and their subdimensions, which were used as a foundation to construct the interview protocol. It also reports the literature sources from which the dimensions were derived.

The seven dimensions of analysis explained above and summarized in Table 1 were used as a basis to develop the case studies further described in Section 3.

2. Methodology

To answer the research question, an interpretive approach was utilized (Walsham, 1995; Klein and Myers, 1999). The research was designed as an explanatory holistic multiple-case analysis, complying with the framework proposed by Yin (2009). To explore connections between conceptual frameworks and practical instances, an explanatory research methodology was implemented. The specific objective was to elucidate the ways in which governance dimensions impact DCs' long-term stability. The rationale for selecting a multiple case study approach, as opposed to a single case study approach, was grounded in the acknowledged benefit of generating more reliable findings as a consequence of the enhanced quantity of the evidence collected (Yin, 2009). The heterogeneity inherent in the DC phenomenon was another factor influencing this choice (Bartolomucci and Bresolin, 2022). To identify governance design factors applicable to all varieties of DCs, it was critical to conduct this research on a substantial number of diverse cases. Moreover, considering the unprecedented nature of the phenomenon being studied, we sought to address the current scarcity of empirical evidence in the field of research (Lapucci and Cattuto, 2021). The decision to utilize holistic case studies, as opposed to embedded case studies, was motivated by the problem of identifying subunits framed in the same structure, the absence of which would have rendered the analysis impossible to replicate. Consequently, the DCs' global governance framework was utilized as a unit of analysis.

https://doi.org/10.1017/dap.2024.29 Published online by Cambridge University Press

Table 1. *Dimensions of analysis and literature review*

Dimension	Subdimension	Cross-sector social partnership publications	Data collaborative publications
Initial conditions	• Leadership • Interdependence	Bryson et al. (2006) Thomson and Perry (2006) Ansell and Gash (2008) Agranoff (2006) Emerson et al. (2012)	Susha et al. (2022)
Trust	• Type of trust • Commitment	Bryson et al. (2006) Thomson and Perry (2006) Ansell and Gash (2008) Provan and Kenis (2008) Emerson et al. (2012)	Stalla–Bourdillon et al. (2021) Klievink et al. (2018) Clarke and Crane (2018)
Incentive system	• Incentive model	Ansell and Gash (2008) Provan and Kenis (2008) Emerson et al. (2012)	Susha et al. (2019b) Susha (2020) Gil Garcia and Sayogo (2016) Moretti et al. (2022) Susha et al. (2017)
Formal Structures	• Formal agreements	Bryson et al. (2006) Koschmann et al. (2012) Emerson et al. (2012)	Stalla–Bourdillon et al. (2021) van Donge et al. (2022)
Intermediation	• Facilitative leadership • Intermediation model	Ansell and Gash (2008) Provan and Kenis (2008) Bryson et al. (2006) Thomson and Perry (2006)	Gil–Garcia and Sayogo (2016) Perkmann and Schildt (2015) Susha et al. (2022)
Business model	• Value proposition • Revenue model		Carballa Smichowski (2019) Susha (2020) Robin et al. (2016) GSMA (2018)
Flexibility	• Inclusivity versus efficiency • Stability versus flexibility	Bryson et al. (2006) Thomson and Perry (2006) Agranoff (2006) Ansell and Gash (2008) Provan and Kenis (2008) Emerson et al. (2012) Lockwood (2009)	Stalla–Bourdillon et al. (2021) Ruijer (2021)

Our data collection was mainly based on the testimonies of individual members of the collaborations. Only in three cases were we able to analyze responses from members of two different organizations within the same DC. Furthermore, our group of respondents consisted mainly of professionals serving in data intermediary organizations. To counterbalance potential biases, we cross-referenced the information collected via the interviews with publicly available secondary sources or documents provided by the respondents. However, the high representation of data intermediaries in our interview sample may distort the importance of some elements in our resulting model, such as the importance of intermediaries. In addition, most of our interviewees from data intermediary organizations held top positions, such as

https://doi.org/10.1017/dap.2024.29 Published online by Cambridge University Press

directors or co-founders. While this lends credibility to our results because of their global perspective, it may detract from the influence of certain governance choices on DC operations.

3.1 Case selection

Cases were selected based on the theoretical sampling technique (Eisenhardt, 1989), starting from the DC taxonomy developed by Bartolomucci and Bresolin (2022). In their classification, through the analysis of multiple variables (e.g., actors involved, value proposition, geography, and data used) and building on the taxonomy developed by Verhulst and Sangokoya (2015),[2] the authors identify and describe five DC clusters facing different development challenges. Notably, only the collaboratives belonging to three clusters out of the five are designed to last; the other two, created to answer emergency conditions or respond to short-term challenges, are not. Therefore, the latter were excluded a priori since they are structurally time-limited and thus not in line with our research objective of investigating DCs' governance settings that foster their long-term stability. Among the remaining 80 DCs pertaining to the three clusters considered, we selected a subsample of 40 DCs based on heterogeneity. All sample cases were contacted via email or social media. The representatives of 14 collaboratives ascribable to the three clusters selected (e.g., Clusters 1, 3, and 5 of Bartolomucci and Bresolin, 2022) agreed to be interviewed.

To strengthen the research findings, empirical heterogeneity was pursued in terms of collaborative capacity to achieve long-term stability. Thus, examples with a proven track record of at least 3 years of activity were compared to either cases that were still active but openly reporting their struggles to continue or cases that had to suspend their operations. Although comparing success and failure cases, often adopted in innovation design (Rhaiem and Amara, 2021) and learning practices (Kapur and Bielaczyc, 2012), introduces a high degree of heterogeneity, it offers a better learning experience. Information about failed cases is often qualitatively superior and more detailed than that about successful cases, especially in cases in which the key to success is avoiding mistakes (Eskreis-Winkler and Fishbach, 2022). Analyzing failure allows researchers to uncover key concepts and induces thoughtfulness in problem-solving (Jackson et al., 2022). Moreover, comparing the assumptions derived from analyzing successful cases with those derived from analyzing failed ones, and vice versa, creates the opportunity to make a preliminary validation of the results and strengthen the research findings (Rhaiem and Amara, 2021).

We conducted 16 interviews with the representatives of 14 DCs. We also performed a review of the academic and grey literature on the cases under analysis. In total, 30 documents were analyzed. In the case of two collaboratives (C10 and C11), which had to interrupt their activities, it was not possible to interview their representatives directly; however, owing to the abundance of available written material, we included them in the analysis to test our findings and hypothesis against cases that were forced to stop.

The selected cases are highly heterogeneous in terms of (i) the macro sector in which they operate (e.g., economic development and infrastructure, health, education, and public safety); (ii) the geographical location, with cases distributed across five continents, although the EU and the US are predominant; (iii) the actors involved, with private, public, and third-sector actors well represented; and (iv) the data used. This heterogeneity makes the results more robust and generalizable, although it introduces some rumor into the research. Notwithstanding the differences observed among the cases, the external validity of the case study results was confirmed by the principle of replication (Yin, 2009). The analysis was limited to factors that could be replicated in a minimum of two cases.

Table 2 provides descriptive information about the cases analyzed.

3.2 Data collection and analysis

The exploratory case studies were developed primarily through semi-structured interviews (Dearnley, 2005; Whiting, 2008) that explored the design and management of the seven previously discussed theoretical governance dimensions. The data collected were analyzed through inductive coding

[2] https://www.unglobalpulse.org/2014/09/mapping-the-next-frontier-of-open-data-corporate-data-sharing/.

https://doi.org/10.1017/dap.2024.29 Published online by Cambridge University Press

Table 2. *Cases analyzed*

ID	DC name	Continuity	Main sector	Region	DC cluster	Data used	Data purpose	Actors involved
C1	Estonia: Mobile Positioning Data for Tourism Statistics (MoPoTuSa)	Ongoing	Economic development	Europe and Central Asia	3. Continuous effort to improve structural responses	Observed personal data	Tertiary	Private Public
C2	California Data Collaborative	Ongoing	Infrastructure	North America	3. Continuous effort to improve structural responses	Disclosed non–personal data	Secondary	Private Public No profit
C3	Civity	Ongoing	Infrastructure	Europe and Central Asia	3. Continuous effort to improve structural responses	Disclosed non–personal data	Primary	Private Public
C4	Salus Coop	Ongoing	Health	Europe and Central Asia	1. Collaborative effort to support wide–scale research projects	Disclosed personal data Observed personal data	Secondary	No profit Civil society
C5	Impact Deal	Ongoing	Economic development	Europe and Central Asia	5. External responses to structural problems	Disclosed non–personal data	Tertiary	Private No profit
C6	Consumer Data Research Centre	Ongoing	Economic development	Europe and Central Asia	3. Continuous effort to improve structural responses	Disclosed non–personal data Observed personal data Observed non–personal data	Tertiary	Private Public
C7	T1D Index	Ongoing	Health	Worldwide	1. Collaborative effort to support wide–scale research projects	Disclosed personal data	Secondary	Private No profit
C8	Needs Map	Ongoing	Social inclusion	Europe and Central Asia	5. External responses to structural problems	Disclosed personal data Disclosed non–personal data	Primary	Private No profit Civil society
C9	Act Now Coalition	Ongoing	Security/public safety	North America	3. Continuous effort to improve structural responses	Observed non–personal data	Secondary	No profit Public
C10	Sidewalk Toronto	Terminated	Infrastructure	North America	3. Continuous effort to improve structural responses	Observed non–personal data	Secondary	Private Public

C11	InBloom	Terminated	Education	North America	3. Continuous effort to improve structural responses	Disclosed personal data Observed personal data	Secondary	Private Public
C12	Bendigo Data Coop	Suspended	Urban development	Australia	3. Continuous effort to improve structural responses	Disclosed non–personal data	Tertiary	Private Public Civil society
C13	Civic Data Design Lab	Ongoing	Urban development	Worldwide	1. Collaborative effort to support wide–scale research projects	Observed personal data Observed non–personal data	Primary Secondary	Public Civil society
C14	Green City Force	Ongoing	Education and training	United States of America	3. Continuous effort to improve structural responses	Observed personal data Observed non–personal data	Primary	Public Civil society
C15	SciExpeM	Ongoing	Economic development	Worldwide	1. Collaborative effort to support wide–scale research projects	Disclosed non–personal data	Secondary	Civil society Private
C16	MATSim	Ongoing	Urban development	Europe	1. Collaborative effort to support wide–scale research projects	Disclosed personal data	Primary	Civil society Private No profit Public

(Yin, 2009) of the interview transcripts and documentary material. The two authors independently extrapolated the codes and classified them based on epistemic similarity (Yin, 2009). Subsequently, the codes were aggregated to construct governance-related categories named first-order codes (Gioia et al., 2010). The process was iterative and abductive, as the logical scheme of the categories was reviewed and checked with reference to the theory several times before the final framework was formulated. Supplementary Annex 1 provides details about the duration of each interview and the role of the interviewee, and Supplementary Annex 2 presents the interview guide.

To enhance the results' interpretability, the first-order codes were grouped according to epistemic affinity in the second-order codes, with the last one divided into (i) processes, (ii) structures, and (iii) elements referring to actors' agency (Bryson et al., 2006; Ruijer, 2021). Supplementary Annex 3 depicts in detail, using a graphical representation, the relations between the first- and second-order codes.

The next section comprehensively presents the findings based on the first-order codes with reference to the elements that emerged in the different case studies. In the discussion, we generalize the findings and discuss them according to the second-order codes. To increase the generalizability of our findings, only design settings that emerged from multiple cases have been considered good design practices.

3. Findings

Table 3 details the 18 first-order codes (i.e., the identified design configurations), their relationships to the seven dimensions of analysis, and the cases from which they emerged. In accordance with our research objective, we only report and remark on findings related to DCs' ability to sustain their activities over time. Each dimension's design setting is described by examples derived from the case studies. Bryson et al. (2006) served as the interpretive lens for the transition from the first-order codes to the second-order ones, which form the discussion's foundation.

3.1 Initial conditions

Four major critical factors have been determined to be fundamental in the initial phase of a partnership. Although these factors pertain to the early stages of a partnership's development, they have considerable implications for its stability over time. An enabling factor for starting a partnership is the *pressure and facilitation exercised by public institutions.* The urgency to act experienced by the public sector, which is frequently accompanied by a corresponding allocation of public resources, has positive effects on the formation of partnerships and facilitates their stability over time. Even formal (i.e., regulatory) or informal pressure from the public sector, without direct participation, can motivate other actors to take action. Regarding this, the California Data Collaborative reports, "There was also a lot of pressure on water suppliers from the State to save water to report data on how much water they're using." The pressure on the public sector to act is often triggered by the *presence of a clear need.* The nature of this need may differ; it can be of exogenous origin, such as an economic or environmental crisis, or of endogenous origin, such as process inefficiencies. The presence of a clear need is directly proportional to how fast a partnership can start and how well it will evolve over time. The MoPoTuSa project states, "They had a need; we had a solution. I must say that even with other kind of projects that we've done and collaborations, […] there must be some sort of pain that needs to be solved." Having a clear need helps in swiftly identifying the best possible contribution a data-intensive solution can bring, thus aligning actors' actions and objectives. Conversely, projects characterized by an unclear value proposition frequently encounter challenges in starting and maintaining their operations due to the increased complexity of comprehending the developed project's impact. The Bendigo Data Coop, which is struggling to sustain its activities, reports, "So we did struggle to find topics, or you know questions to answer that everyone felt were collectively good questions."

Other important factors in the initial phases of partnership development, which also influence a partnership's stability over time, are the presence of preexisting *relationships among partners*, the existence of *conjunct previous experience*, and a strong *individual leadership.* The presence of preexisting

https://doi.org/10.1017/dap.2024.29 Published online by Cambridge University Press

Table 3. *First-order codes and the relation with cases*

Critical factors (first-order codes) / Dimensions	Presence of a clear need and interdependence	Pressure and facilitation by public institutions	Conjunct previous experience	Competence and data expertise	Innovation, technology, and data infrastructure	Vision and mission alignment	Communicating privacy compliance	Operational agreements with partners	Non-disclosure and data-limitation agreements and the presence of an external committee	Presence of a dedicated and neutral intermediary	Partner and stakeholder engagement	Interest in data outcome	Non-profit legal entity	Social impact measurement	Revenue generation	Communication and transparency	Lean development	Openness toward new partners and partner-selection process
Initial conditions	C1 C2 C5 C7 C9 C14 C15	C1 C2 C6 C15 C16	C1 C2 C4 C5 C6 C7 C8 C9 C11 C15 C16		C1 C8 C14						C12						C3 C4 C11	
Trust		C1 C8 C16 C13	C1 C2 C3 C4 C5 C8 C12 C14 C16	C1 C6 C9 C12 C13 C16		C2 C4 C6 C7 C8 C11	C1 C3 C4				C6 C7 C10 C11 C12 C13		C2 C4 C8 C11 C14	C4 C6 C8		C1 C2 C8 C9 C10 C11 C16	C12C13	
Formal structures					C14		C16	C4 C5 C6 C7 C8 C14 C16	C2 C3 C6 C9 C14 C16									
Intermediation				C2 C3 C4 C6 C8 C12 C15	C3 C5 C6		C5 C9 C10 C11		C12	C1 C2 C3 C4 C5 C6 C7 C8, C9 C12	C3 C8 C12		C4 C9 C14					
Incentive system		C3 C8 C14			C1 C3 C14	C12					C6 C8 C15	C1 C2 C3 C6 C7 C8 C9 C12 C14 C15	C4 C6	C1 C4 C5 C6 C8 C12				C12
Business model															C1 C2 C3 C4 C8 C9 C12			
Flexibility																C1 C6 C8	C2 C1 C13	C1 C2 C3 C4 C5 C6 C7 C8 C9 C13 C15

relationships facilitates trust building and maintenance, thereby compressing the start-up time and aiding in overcoming critical moments of a collaboration during its entire lifetime. Personal—more than organizational—preexisting relations play a crucial role in the start-up phase of a collaboration and in keeping high levels of engagement even without immediate positive results. The California Data Collaborative reports, “The initial startup trust was just trust in [person name]. The other water suppliers trusted what she was saying, they trusted her vision, and so they went along with it.”

Finally, evidence suggests that it is advantageous for DCs to implement a *lean development* process that begins with the production of small achievements and evolves through continuous improvement. Regarding this, Civity reports, “So we made it as a showcase, but the showcase was so successful that more municipalities came to us, and they wanted to use this application.” Adopting this strategy not only increases partners’ motivation and commitment but also tests the project for potential flaws (particularly in relation to data use) and refines its activities and outputs over time. When a project begins on a grand scale, it may encounter ethical and privacy-related problems that result in its suspension or termination. In an interview, a manager of the InBloom project, which was forced to stop its activities, states, “The project has to scale fast. It’s a big, heavy, large-scale project, with a large initial investment. The belief is that it can’t be designed small, so we have to go big.”

3.2 Trust

Trust is a complex construct that depends on multiple elements; however, through the analysis of the case studies, a few of these have been isolated. The first element contributing to a partnership’s stability is complete *vision and mission alignment* among the partners with regard to its objectives; this is crucial to its long-term stability. In line with this, the Consumer Data Research Centre reports, “It was all about having senior academics who all had the same vision for setting up something which was responding to a call to do something in a research space where there was a need.” In contrast, when a partnership begins with a poor alignment, problems and breakups are common. This was observed in the MoPoTuSa project: “[…] some of the agencies departed early. The fundamental underlying issue was a difference in vision.” A second key factor in the case of DCs is the capacity to ensure *full privacy compliance.* Being compliant with privacy regulations is essential when managing data, particularly the sensitive type. To establish legitimacy, collaborations must not only be capable of defining and implementing secure data processes but must also effectively communicate these processes to stakeholders. Salus Coop reports, “To build trust, we have to go to street […] and explain that one of our values is keeping privacy at the very high level, because we know that this is a very sensitive thing for citizens.” In contrast, Sidewalk Toronto, which was forced to suspend its operations, encountered significant opposition due to ambiguous communication regarding data-use terms. A newspaper article on the project reports, “Who will own the data streaming from sensors in every park bench, lamppost and dumpster? No one at [company name], nor in local government, has given a straight answer to that question yet.” In addition to guaranteeing *communication and transparency* regarding data privacy, every aspect of the partnership must be communicated transparently to both partners and other stakeholders, including data use and other aspects of a collaboration’s management. The California Data Collaborative reports, “We try to be very communicative about what we’re working on at any given time. In terms of trust, we’re very transparent about our financials, how we’re spending money, what projects we’re working on.” On the same issue, Needs Map reports, “We should be transparent with our partners, including volunteers, about organizational, administrative, and financial information.” Lack of communication creates mistrust and may engender internal and external tensions, as in the case of InBloom, whose stakeholder declares, “Trust between states and districts, between districts and teachers, between parents and states, and teachers and technology was not adequately addressed by [the company’s] communication strategy.” Adopting proactive communication is the first step in promoting active *partner and stakeholder engagement.* Actively involving partners and stakeholders in decision-making processes from the outset and throughout the development of activities aids in building trust and legitimacy and in preventing significant setbacks that threaten a partnership’s survival. A representative of the Civic Data Design Lab reports, “At

https://doi.org/10.1017/dap.2024.29 Published online by Cambridge University Press

the beginning they didn't trust me, they didn't know who I was. I have built trust by being there, coming back frequently, continuously including them, involving them in workshops, continuing to help them even when the project was over." On a different level with regard to the first three elements, another critical point is the capacity to demonstrate the social value generated through a partnership's activities by setting a *social impact measurement* infrastructure. Measuring the social impact generated is essential for renewing stakeholders' motivation and gaining legitimacy for an organization's long-term operations. Few of the interviewed collaboratives have been able to do so, and many have struggled to assess and convey the value of the social impact generated despite recognizing its significance. On the matter of how to create trust, Needs Map reports, "We are sharing with the people and organizations our projects, our efforts in the field, and we are saying that we create this impact." In contrast, the Bendigo Data Coop clearly refers to the difficulty of demonstrating their impact: "We got valuable data out of it, where it gets a little bit tricky, and the rubber hits the road is: What is the data going to influence? What change is it going to influence?" Finally, the analysis of the case studies suggests that, being the convener, a *not-for-profit legal entity* positively influences trust creation and diminishes resistance. The California Data Collaborative reports, "By being a non-profit, it really does build trust, especially in the early days when no one really knew who we were. It's like a badge of honor that builds trust." On the other end of the spectrum, there is some evidence of the risk associated with a for-profit structure managing the collaboration. For example, in the InBloom case, reports indicate that "[there was a] deep worry that the data infrastructures underpinning schooling systems would potentially be controlled by private for-profit interests."

3.3 Formal structures

Regarding formal structures, three main elements appear to be critical for the long-term sustainability of a partnership. The first one is implementing a clear *operational agreement with partners* that guarantees partnership development and sets roles, responsibilities, and the resources to be allocated. These agreements often take the form of a memorandum of understanding (MOU). Regarding this, Salus Coop reports, "The agreements are not technical at all. They are very simple." However, many recognize the agreements' role in maintaining the organizations engaged in the project beyond the single data champions that may have started it. The T1D Index reports, "The MOU definitely helped keep people committed because there's change of leadership in all the organizations. Some partners have new presidents every year. So it was good to have a MOU to keep continuity." Therefore, operational agreements are needed to ensure organizational commitment transcending the will of individual reference people in the long term. Such simple operational agreements should, however, be coupled with detailed *non-disclosure and data-limitation agreements*, which are often complex in nature and may imply the creation of ad hoc data-use assessment processes. C16 reports, "So there is funding agreement, basically, there is the ethics agreement, there is the user agreement, and there is the agreement with the partner providing the app. The user's data is always a central aspect in all these agreements." The adherence to data-use terms and conditions is often guaranteed through the *presence of external committees*, which assess data-use requests and grant permission to proceed with data sharing in compliance with privacy regulations and agreements. The Consumer Data Research Centre reports, "We have some expert reviewers who look at all of those proposals, and then we take it to the data provider who has ultimate sign off on that proposal."

3.4 Intermediation

The intermediation modality has significant implications for DCs' long-term stability. One of the most influential elements in this regard is the *presence of a dedicated and neutral intermediary.* A collaborative may benefit from the presence of an intermediary organization in terms of power dynamics' balance, resource management, conflict resolution, and funding. The California Data Collaborative comments on this as follows: "Being able to move things to a third party made it so that the members and the participants could be more equal." The presence of an organization representing a DC facilitates the management of support activities for the initiative, such as the management of human and financial resources and the supervision of processes. Furthermore, having a dedicated entity aids in the management of tensions and the prevention of dropouts through the establishment of one-on-one relationships with various partners. Regarding this, the T1D Index states, “So the fact that [the organization] invested in a team which was only focused on this project, that was really important to the ongoing sustainability of the work.”

Evidence suggests that to foster the long-term stability of a partnership, intermediaries must incorporate advanced *data-related competence and skills.* The development of data-sharing infrastructure and standards and the execution of data cleansing and analysis are among the most frequently reported responsibilities of intermediaries. The California Data Collaborative reports, “You have to show your technical capabilities, and that you actually know what you are doing.” However, a DC must also possess other competencies, such as stakeholder engagement and interaction with specific beneficiaries in need. Needs Map expresses this as follows: “For big companies, it is not possible to interact with beneficiaries without [our] intermediation; they don’t have expertise. They can implement the projects, but it’s about knowing what disadvantaged and disabled people need.” To expedite the operation of a collaborative, intermediaries should take the form of knowledge centers in which various kinds of competencies are centralized.

Although the *technological and data infrastructure* is not the focus of this research, it has emerged as a critical factor for the success of a partnership, indicating the necessity to design it in coordination with the organizational aspects concerning the intermediation model and vice versa (Bharosa and Janssen, 2015). The Impact Deal representatives refer to this as follows: “Also the infrastructure theme, how can the data club evolve? One way may be to evolve it as a data space [...]. If the data space is a neutral subject, then the governance must also be neutral. It cannot be a single subject, then there could be the need for an independent third party.” However, the relationship between the two dimensions is not streamlined and must be further explored.

3.5 Incentive system

Excellence in the technological infrastructure is also required to entice and retain partners in a collaboration. Organizations are attracted by the possibility of utilizing innovative technologies, whether for understanding their potential or for reputational reasons. Regarding this, Civity reports that they “have a technology that is way more advanced with respect to what could be developed within municipalities. We also comply with the highest ISO standards and with the GDPR standards, which are really difficult sometimes. So we have all the certificates and competencies to manipulate the data.” Possessing not only the technological infrastructure but also data management and analysis capabilities is, for intermediaries and DCs in general, an important way to incentivize partners to participate in the collaboration.

Organizations are frequently motivated to participate in DCs by an *interest in data outcomes* (i.e., the intentional or unintended benefits they may obtain from the analysis of their data). Referring to this and the previous dimension, the MoPoTuSa project reports, “So for everybody, [the] incentive [is] just to do new things, and that is kind of the goal. For the operators, it might as well be that they want to learn more about their data because these kinds of initiatives could reveal for them something that they didn’t know before.” Other elements already discussed in this article that have significant implications for enticing partners to join a collaborative are (i) stakeholder engagement starting from the design phase of a partnership, (ii) the capacity to measure the impact generated, and (iii) the not-for-profit nature of an intermediary/project.

3.6 Business model

Developing robust funding models that include *revenue-generating activities* allows collaboratives to reduce their dependence on external funding to sustain themselves. Revenues may normally be generated through two broad categories of activities: (i) “internal services,” which are essential services offered as part of the collaboration, and (ii) “external services,” which are directed toward actors outside the

https://doi.org/10.1017/dap.2024.29 Published online by Cambridge University Press

collaboration and typically involve the exploitation of knowledge and resources already developed within the collaboration. Depending on the type of service or product being offered, the business model may be based on freemium models, one-time license payments, recurring payments, or transaction fees. The Act Now Coalition, for instance, reports, "We actually have been able to kind of have a funding model where we can license access to the API and the data to these businesses." Nearly all the collaboratives, particularly in the initial phase, received a grant that assisted them in establishing their partnership and acquiring the necessary assets to construct the first successful use case. However, organizations that rely primarily on donations struggle to maintain their partners' commitment over time and risk having to suspend their operations after a certain point. Therefore, the critical element is the capacity to design, from the beginning or during the start-up phase, valuable business models that may sustain a collaborative's activity once the initial grant used to establish it is depleted.

3.7 Flexibility

The analysis of the case studies also shows that it is important for a collaborative to embrace a certain degree of *openness toward new partners* and to *formalize a partner-selection process.* Openness to new partners favors the renewal and longevity of partnerships. New partners bring new resources, meaning, and commitment. For more market-oriented collaboratives, new partners are almost treated as new clients, such as in the case of the Act Now Coalition: "Partner for us at this point would be like new potential opportunities for sustainability like a new, we don't want to say customer, but that's really what they are." This implies that a collaborative develops initiatives or projects with a high degree of adaptability. New collaborators frequently entail new data, thereby increasing the value of the generated analysis. Impact Deal expresses it clearly: "There is a need for new partners. First of all, on the data side." Not-for-profit forms of collaboratives are often more cautious about accepting new partners because they want to preserve their value proposition, avoid deviating from it, and maintain control over their activities. This is the case of Salus Coop: "We have been very conservative on partnership. […] Projects that we host [need to] go through an ethical committee, and when [they are] approved, [they] can collaborate with us. It's a way in which we can filter quality and critical mass." Others have established formal or informal assessment processes to deliberate on the inclusion or exclusion of new partners. However, being able to balance flexibility and openness with new partners' mission alignment appears to be a key challenge that should be addressed. Another challenge in terms of flexibility is the ability to adequately handle internal and external tensions. In this regard, having a *fully transparent communication* strategy, supported by clear accountability procedures, is a safeguard against various forms of criticism and hazards. For example, the Consumer Data Research Centre reports, "There are sometimes tensions with the users of the data or people who want to collaborate, because quite often the understanding of what consumer data are is not that great […]. Trying to manage the expectations of collaborators is really important in that sphere, and again it comes down to the communication."

While the factors examined did not encompass every aspect of DC governance (as in the case of Ruijer, 2021), they have emerged as crucial determinants of DCs' stability. These factors aid in the reduction of risks and occurrences that typically culminate in the dissolution of a DC following its initial pilot phase. These may include partners' diminished interest in investing in a partnership, insufficient evidence regarding the effectiveness of a DC in delivering on its initial value proposition, inadequate incentives in comparison to the risk–benefit calculus, insufficient funds to maintain operations, and controversies surrounding data privacy.

4. Discussion

The distinction between elements pertaining to processes, structures, and actors' agency (Bryson et al., 2006) served as the interpretive lens for the transition from the first-order codes to the second-order ones. The aggregation in the second-order codes has facilitated the discussion of the results and their comparison with existing collaborative governance models (Agranoff, 2006; Bryson et al., 2006; Ansell

https://doi.org/10.1017/dap.2024.29 Published online by Cambridge University Press

and Gash, 2008; Provan and Kenis, 2008; Emerson et al., 2012). Aggregating the second-order codes under the tripartite lens of process–structures–actors' agency highlights the presence and interrelationship, in both governance processes and structures, of elements pertaining to data management and the enabling technological infrastructure as well as more organizational aspects. This evidence, obtained by starting from a predominantly organizational perspective, confirms and reinforces the need to adapt existing governance models, which often focus exclusively on organizational aspects, to the peculiarities of DCs (Ruijer, 2021). Furthermore, these findings support a few researchers' (Bharosa and Janssen, 2015; Susha et al., 2017; Bharosa, 2022; Otto and Hompel, 2022) call to consider the two dimensions of organizational governance and data and technical infrastructures' governance as factors that mutually influence each other. This perspective, although widely recognized, often suffers from the existence of research silos between information systems research and more organization-oriented research. The interrelationships among these two elements are evident in the governance factors' design settings reported in Table 4 and further discussed below.

The first factor emerging as a critical element in guaranteeing DCs' long-term stability is their reliance on robust, well-defined, and schematized *data-related processes* along the entire data value chain. This requires precise data collection, storage, access, and sharing procedures. Data-management processes are also determined by the organizational configuration a DC adopts. Notably, the data value chain is frequently administered by a single actor who possesses the required skills and knowledge. Nonetheless, the concentration of data capabilities in the hands of a single actor may result—more than in other forms of partnerships—in power imbalances and the emergence of endogenous and exogenous crises. Therefore, the cases analyzed indicate a necessity to activate governance processes (e.g., multi-stakeholder decision-making mechanisms) aimed at mitigating these risks. Compliance with privacy and data security regulations, as well as the adoption of a fully transparent and proactive communication strategy regarding the entire data value chain and addressing all stakeholders, may contribute to this objective. The combination of compliance and communication significantly reduces risks and encourages the participation of stakeholders in a partnership.

A transparent communication strategy also helps increase stakeholder *engagement.* In this regard, the importance of a strong value proposition alignment emerges as a key element for a partnership's continuity. This dimension, often cited in the collaborative governance literature (Bryson et al., 2006; Ansell and Gash, 2008; Provan and Kenis, 2008; Emerson et al., 2012), is overshadowed by the DC one, which often focuses more on the presence of a clear social need (Susha et al., 2020; Ruijer, 2021) as a factor triggering actors to participate in a partnership. While a clear social need is acknowledged as an important component in the initial development of a partnership, evidence shows a strong value proposition alignment to have a significant impact on partnership stability over time. In this regard, the more stable collaboratives among those analyzed often developed value proposition alignment activities both during the partnership's founding phase and throughout the course of its development via the implementation of *participatory design processes.* Having a value proposition that transcends a single need helps in modifying a partnership's commitment in the long term and avoids the loss of direction once the initial purpose is achieved or in case data do not fit the immediate need as it has been framed. Compared to other forms of partnerships, there is a greater need for the value generated for each stakeholder to be made clear, including the value for indirect beneficiaries and data owners that may or may not be included in a partnership. Including stakeholders in the partnership development and decision-making process may be operationalized through the adoption of a *lean development process* (Klievink et al., 2018; Global Partnership for Sustainable Development Data, 2023), which is based on the achievement of small and gradual goals (Ruijer, 2021). Following a strategy of gradual evolution helps in *demonstrating the social value* generated and contributes to generating buy-ins from existing and new partners. The same is true for the design and implementation of a structured impact management and measurement model. However, although both practitioners and academics recognize its value, efforts focused on this remain limited (Hlabano and Van Belle, 2019), and further research is required.

More than other kinds of collaborations, durable DCs consistently demonstrate a *high degree of adaptability* to external demands and a willingness to continuously recruit new partners. Engaging new

Table 4. *Governance design settings favoring DCs' stability*

Interpretative lens	Second-order dimensions	Governance design settings favoring DCs' stability
Processes	Data–related processes	– Establish explicit processes for data collection, storage, access, and sharing – Balance potential power imbalances caused by the concentration of data–related capabilities via a distributed governance model – Ensure processes' compliance with data regulations – Adopt a completely transparent and proactive data–usage communication strategy
	Engagement and participatory decision–making	– Ensure a strong value proposition alignment from the outset of the partnership and keep working on it throughout its development – Clarify the value proposition for all stakeholders, including data owners. Follow a lean development strategy through small and gradual achievements – Manage and measure the impacts generated – Include new partners continuously by assessing them beforehand – Be flexible in meeting market and partners' demands
Actors' Agency (In between)	Leadership	– Leverage informal and individual leadership, along with previous relations, networks, and reputations in the start–up phase of the partnership – Gradually transfer individual leadership to the organizational level after the start–up phase
	Collective intelligence	– Combine competencies and knowledge from different sectors and backgrounds to improve data interpretation and operationalization – Utilize anticipated and unanticipated data outcomes to incentivize and engage data providers
	Funding model	– Generate a mix of income sources to gain decisional independence and avoid being overdependent on funding partners – Provide data analysis services to generate revenue and incentivize partners to join the collaborative
Structures	Technological structures	– Ensure technological excellence and reliability, especially when dealing with sensible data – Demonstrate and communicate technological excellence to incentivize partners to join
	Organizational structures	– Constitute an independent intermediary, ideally in the form of a not–for–profit organization, to build trust and manage stakeholders' interests – Constitute intermediaries in the form of knowledge centers, integrating technical and social skills – Couple simple operative agreements with detailed and accurate data–usage agreements

https://doi.org/10.1017/dap.2024.29 Published online by Cambridge University Press

collaborators allows a collaborative to obtain access to new data and resources, thereby enhancing the analysis' value. Additionally, evidence demonstrates that stable DCs rely on *structured assessment processes* to evaluate partners' alignment with their values and activities prior to their entry into a partnership. Adaptability should not only be a value proposition characteristic but also be intentionally embraced in organizational and technological structures. Particularly important in this regard is the capacity to manage the trade-off between flexibility and the presence of a clear value proposition and agreements, especially regarding data use.

With regard to those governance dimensions in between processes and structures that refer to the agency of the actors involved, in accordance with previous literature (Bryson et al., 2006; Ansell and Gash, 2008; Emerson et al., 2012), our research provides evidence of the significance of preexisting collaboration experiences among actors, the relationship network, and reputation. Our research also emphasizes the significance of informal and individual *leadership*, particularly during the early phases of a collaboration. However, the findings discourage excessive reliance on personal relationships and single individuals. Relying on a single person's reputation and establishing data champions may engender success for a partnership at the beginning (Hoffman et al., 2019), but these expose it to a high degree of reliance on the individual's continuous involvement. To enhance their long-term stability, collaboratives must progressively gain independence from individual leaders and transfer leadership and trust capital at the organizational level. Transferring leadership and trust from the personal level to the organizational level is not an automatic process; rather, it needs to be planned and pursued through the creation of ad hoc organizational structures and processes.

Our research also emphasizes the significance of the activation of *collective intelligence* (Daly et al., 2019; Verhulst et al., 2019) as a crucial element for the long-term sustainability of partnerships and the transition from the descriptive use of data to the capacity to operationally use data. While the availability of data-related skills is a prerequisite for collaboratives' activation, more social-oriented skills allow collaboration deployment over time. Possessing in-depth knowledge of the beneficiaries or area of operations enables the creation of a link between high-level data analysis and empirical field data. Ground-truthing assumptions (Williams, 2020) and facilitating engagement and dialogue have emerged as key factors in this regard. Including actors working on the field of analysis helps in translating findings and evidence into actions and facilitates the creation of evidence on the impact generated. The activation of collective intelligence and the cross-contamination of diverse knowledge and backgrounds enable the adoption of novel analytical and interpretative lenses when analyzing data, which frequently results in the acquisition of unexpected data outcomes (i.e., shared knowledge derived from data analysis). *Expected and unexpected data outcomes* appear to be the clearest incentives for partners to participate in a collaboration. Furthermore, offering on-demand data analysis services incentivizes partners to join a partnership while also serving as the primary source of revenue for a DC's intermediaries, in addition to grant funding and membership fees (GSMA, 2018; Global Partnership for Sustainable Development Data, 2023). Our research confirms the difficulties that DCs face in establishing robust *funding models* and demonstrates the resilience of DCs that have succeeded in establishing such models (GSMA, 2018; Susha et al., 2022). It also demonstrates that the optimal funding model is frequently achieved as a result of multiple iterations in relation to the partners involved, the use of data, the available competencies, and the extant market demand for the services and products offered. Hybrid forms of funding, combining revenues, private grants, and research grants, appear to be the most resilient ones. Cross-subsidizing activities directed primarily at not-for-profit initiatives, with services offered to private and public actors, is one of the most robust options DCs can develop. However, establishing cross-subsidized models may create structural inefficiencies and trade-offs, which increase the management complexity (Battilana et al., 2015).

Concerning governance structures, a clear distinction between *technological and organizational structures* has emerged. However, the overall picture generated by the research provides clear evidence that the two are strongly related and should therefore be designed in relation to one another. The two dimensions can either exacerbate or mitigate their respective limitations. For instance, the centralization of technical competencies, which frequently generates power imbalances, can be counterbalanced by a participatory management and decision-making system. The trade-off between centralization and decentralization existing in both dimensions can thus be managed as a four-dimensional problem, with multiple configurations. However, further research on this is needed to understand the optimal combinations.

Technological infrastructures' excellence is a necessity for DCs, especially in those cases in which sensitive data sharing is involved. However, while safe data-sharing environments and data-sharing protocols are easily accessible at present, reference models on the managerial and organizational levels are lacking. A few elements can be derived from the research in this regard: In accordance with previous literature (Hernández-Chea et al., 2021; Reischauer et al., 2021; Susha et al., 2022), the presence of an *independent data intermediary* appears to be beneficial for the long-term stability of a partnership, as it pushes organizations to commit to sustain the partnership in the long term, manages one-to-one relationships with partners, and simplifies and hastens decision-making processes. Evidence from this study also suggests that intermediaries should take the form of *knowledge centers* and integrate more technical and data-related skills with socially oriented ones. Finally, it appears beneficial for trust creation and partner involvement to establish intermediaries under a *not-for-profit* legal form, which fosters trust generation and is generally less vulnerable to criticism.

5. Conclusions

This research advances the existing literature on DCs through three main contributions. First, it progresses from determining which aspects must be included when designing DC governance schemas (Susha, 2019; Ruijer, 2021) to determining the ideal design settings of those related to partnership stability. Second, it answers the call (Lapucci and Cattuto, 2021) to provide more empirical evidence on the field that can serve as a reference for others and challenge previous explorative studies (Smichowski, 2019; Susha, 2019; Ruijer, 2021). Third, it helps clarify which factors favor DCs' long-term stability, which is unanimously recognized as one of the major limitations affecting DCs' capacity to generate systemic impact (GSMA, 2018; European Commission, 2020; Farmer et al., 2022; Flanagan Anne and Sheila, 2022; Susha et al., 2022). Such a focus on the long-term stability of a partnership constitutes a new perspective in DC literature, which, despite recognizing long-term stability as an important element for DCs' development, has only partially focused on how to foster it. By adopting this research perspective, this research highlights the importance of a few elements that were overlooked by previous literature. For example, this is demonstrated by the importance of moving trust from the personal level to the organizational level, adopting mixed funding schemas, internalizing data-analysis competence at the intermediary level, and adopting a fully transparent and proactive communication strategy. The particular focus adopted has also allowed us to highlight elements that evolve during a partnership's development phases. This is illustrated by what motivates partners to engage in a DC. While the initial motivation to join a partnership is often driven by the presence of a clear social need, in the long term, motivation is primarily driven by (i) a strong value proposition alignment, (ii) the creation of small and simple successful cases (i.e., the adoption of a lean development strategy), (iii) the assessment and reporting of the impacts generated, and (iv) access to data-related services and data-analysis outcomes. While we have focused on individual dimensions' design settings' influence on partnership stability, we acknowledge the existence of multiple trade-offs and reciprocal influences among them, which require further analysis. These trade-offs are typical of social entrepreneurship and concern, for example, the balance between decisional centralization and stakeholder engagement, resource allocation among more socially oriented activities with regard to revenue-generating ones, and individual leadership and trust versus its institutionalization. Therefore, we recommend further research addressing these trade-offs to generate social and economic value (Battilana et al., 2015).

This research also highlights the interconnection and interdependence between collaborative governance dimensions and technological ones, calling for more transdisciplinary research on the topic. DCs are complex socio-technical structures (Liva et al., 2023); therefore, they must be studied with an integrated technical and organizational perspective. DCs require the creation of specific processes and structures for collecting, sharing, storing, and analyzing data in a safe and privacy-compliant manner. At the same time, data-related processes and structures are influenced by and influence organizational governance elements. As it is not possible to create a one-size-fits-all model (Klievink et al., 2018; Stalla-Bourdillon et al., 2019), the governance configuration should be adaptive with regard to the context and the different development phases (Bharosa and Janssen, 2015); thus, organizational and technological aspects need to be continuously adapted in synchrony. Therefore, we encourage future research to adopt a holistic approach toward DC governance, considering and further exploring various relationships between organizational and technological elements.

From a practitioner's perspective, our research provides empirical support and guidelines to design an effective DC governance setting. Despite only partially addressing the general lack of reference models for practitioners, our research represents a significant contribution in this direction. When designing the governance of their collaboratives, practitioners can refer to these research findings as a catalog of elements to consider and fallacies to avoid. In this regard, this study's reliance on 16 distinct cases makes it a significant and robust source of knowledge. Nevertheless, we encourage future research as well as other endeavors on this topic to aid practitioners in their efforts by creating a shared library of effective models to which they can refer when structuring their partnerships.

Finally, our research contributes to the broader discussion on the governance of new commodities (e.g., internet connection, artificial intelligence, and data) that, because of their actual or potential impact on communities, require experimentation involving different schemes of multi-stakeholder governance that would enable a commons-like (Avermaete, 2021) approach, ensuring the representativeness of all stakes in their management. These new approaches may govern the production, protection, enjoyment, and reproduction of such resources (Avermaet, 2021), ensuring social licensing over the asset (Verhulst and Saxena, 2022) and the pursuit of general interest objectives. Thus, our research may constitute one specific use case, showing which governance dimensions' design configurations may facilitate the management of these new forms of commodities. Therefore, we recommend further research in this direction and urge other researchers to explore the governance design settings of other types of assets in this category.

Abbreviations

DC data collaborative
EU European Union

Supplementary material. The supplementary material for this article can be found at http://doi.org/10.1017/dap.2024.29.

Acknowledgments. The authors are grateful to all the interviewees for their time and contribution to the research.

Competing interest. The authors declare none.

Author contribution. Conceptualization: F.B.; Data curation: F.B., S.B.; Data visualization: F.B., S.B.; Methodology: F.B., S.B.; Writing—original draft: F.B., S.B.; Writing—review and editing: F.B., S.B. All authors approved the final submitted draft.

Data availability statement. Data collected from the interviews are not shareable.

References

Agranoff R (2006) Inside collaborative networks: Ten lessons for public managers. *Public Administration Review 66*(Suppl. 1), 56–65. http://doi.org/10.1111/j.1540-6210.2006.00666.x.

Ansell C and Gash A (2008) Collaborative governance in theory and practice. *Journal of Public Administration Research and Theory 18*(4), 543–571. http://doi.org/10.1093/jopart/mum032.

Avermaete T (2021) Constructing the commons: Towards another architectural theory of the city? In *The New Urban Condition: Criticism and Theory from Architecture and Urbanism*. London: Routledge, pp. 54–72.

Bartolomucci F and Bresolin G (2022) Fostering Data Collaboratives' systematisation through models' definition and research priorities setting. *ACM International Conference Proceeding Series*, pp. 35–40. http://doi.org/10.1145/3543434.3543442.

Battilana J, **Sengul M**, **Pache AC and Model J** (2015) Harnessing productive tensions in hybrid organizations: The case of work integration social enterprises. *Academy of Management Journal 58*(6), 1658–1685. http://doi.org/10.5465/amj.2013.0903.

Bharosa N (2022) The rise of GovTech: Trojan horse or blessing in disguise? A research agenda. *Government Information Quarterly 39*(3), 101692. http://doi.org/10.1016/j.giq.2022.101692.

Bharosa N and Janssen M (2015) *Challenging the Chain: Governing the Automated Exchange and Processing of Business Information*. Amsterdam: IOS Press, pp. 1–420. http://doi.org/10.3233/978-1-61499-497-8-i.

Bryson JM, Crosby BC and Stone MM (2006) The design and implementation of cross-sector collaborations: Propositions from the literature. *Public Administration Review 66*(SUPPL. 1), 44–55. http://doi.org/10.1111/j.1540-6210.2006.00665.x.

Bryson JM, Crosby BC and Stone MM (2015) Designing and implementing cross-sector collaborations: Needed and challenging. *Public Administration Review 75*, 647–663. http://doi.org/10.1111/puar.12432.Designing.

Bughin J, Hazan E, Allas T, Hjartar K, Manyika J, Sjatil PE and Shigina I (2019) *Tech for Good Smoothing Disruption, Improving Well-Being. McKinsey Global Institute.* Available at https://www.mckinsey.com/~/media/mckinsey/featured insights/future of organizations/tech for good using technology to smooth disruption and improve well being/tech-for-good-mgi-discussion-paper.ashx (accessed June 2023).

Charles J and Tonetti C (2020) Nonrivalry and the economics of data. *ISSN 2502-3632 (Online) ISSN 2356-0304 (Paper) Jurnal Online Internasional & Nasional Vol. 7 No.1, Januari – Juni 2019 Universitas 17 Agustus 1945 Jakarta 53*(9), 1689–1699. Available at https://www.nber.org/papers/w26260 (accessed June 2023).

Chignard S and Glatron M (2023) Data collaborations at a local scale: Lessons learnt in Rennes (2010–2021). *Data and Policy 5*, e20. http://doi.org/10.1017/dap.2023.16.

Clarke A and Crane A (2018) Cross-sector partnerships for systemic change: Systematized literature review and agenda for further research. *Journal of Business Ethics 150*, 303–313. http://doi.org/10.1007/s10551-018-3922-2.

Daly A, Devitt KS and Mann M (2019) *Good Data Good Data*, Vol. 29. Available at http://networkcultures.org/blog/publication/tod-29-good-data/ (accessed June 2023).

Dearnley C (2005) A reflection on the use of semi-structured interviews. *Nurse Res. 13*(1), 19–28. http://doi.org/10.7748/nr2005.07.13.1.19.c5997. PMID: 16220838.

Digital Civil Society LAB (2017) *Workshop Summary: Trusted Data Intermediaries. Medium.* Available at https://medium.com/the-digital-civil-society-lab/trusted-data-intermediaries-c58426c2c994 (accessed June 2023).

Eisenhardt KM (1989) Building theories from case study research published by: Academy of management stable. *The Academy of Management Review 14*(4), 532–550.

Emerson K, Nabatchi T and Balogh S (2012) An integrative framework for collaborative governance. *Journal of Public Administration Research and Theory 22*(1), 1–29. http://doi.org/10.1093/jopart/mur011.

Eskreis-Winkler L and Fishbach A (2022) You think failure is hard? So is learning from it. *Perspectives on Psychological Science 17*(6), 1511–1524. http://doi.org/10.1177/17456916211059817.

EU Commission (2020) *Towards a European Strategy on Data Sharing for the Public Interest.* High-Level Expert Group on Business-to-Government Data. http://doi.org/10.2759/406717.

European Commission (2020) EU Data Strategy.

Farmer J, Aryani A, Tucker J and Woo J (2022) *City og Greater Bendigo Data Co-Op.*

Farmer J, Mccosker A, Albury K and Aryani A (2023) *Data for Social Good Non-Profit Sector Data Projects.*

Flanagan Anne J and Sheila W (2022) *Advancing Digital Agency : The Power of Data Intermediaries.*

Gil-Garcia JR and Sayogo DS (2016) Government inter-organizational information sharing initiatives: Understanding the main determinants of success. *Government Information Quarterly 33*(3), 572–582. http://doi.org/10.1016/j.giq.2016.01.006.

Gioia DA, Price KN, Hamilton AL and Thomas JB (2010) Forging an identity: An insider-outsider study of processes involved in the formation of organizational identity. *Administrative Science Quarterly 55*(1), 1–46. http://doi.org/10.2189/asqu.2010.55.1.1.

Global Partnership for Sustainable Development Data (2023) Effective and Ethical Data Sharing at Scale.

Groves RM and Neufeld A (2015) Accelerating the Sharing of Data Across Sectors to Advance the Common Good.

GSMA (2018) Scaling Big Data for Social Good: The Need for Sustainable Business Models, September. Available at https://aiforimpacttoolkit.gsma.com/resources/BD4SG_The_need_for_sustainable_business_models.pdf (accessed June 2023).

Hernández-Chea R, Mahdad M, Minh TT and Hjortsø CN (2021) Moving beyond intermediation: How intermediary organizations shape collaboration dynamics in entrepreneurial ecosystems. *Technovation 108*, 102332. http://doi.org/10.1016/j.technovation.2021.102332.

Hlabano M and Van Belle JP (2019) *Tracing the Impact of the City of Cape Town's Open Data Initiative on Communities and Development* IFIP Advances in Information and Communication Technology, Vol. *551*. Cham: Springer International Publishing. http://doi.org/10.1007/978-3-030-18400-1_23.

Hoffman W, Bick R, Boral A, Henke N, Olukoya D, Rifai K, Roth M and Youldon T (2019) Collaborating for the common good : Navigating public-private data partnerships, May.

Jackson A, Godwin A, Bartholomew S and Mentzer N (2022) Learning from failure: A systematized review. *International Journal of Technology and Design Education 32*(3), 1853–1873. http://doi.org/10.1007/s10798-021-09661-x.

Kapur M and Bielaczyc K (2012) Designing for productive failure. *Journal of the Learning Sciences 21*(1), 45–83. http://doi.org/10.1080/10508406.2011.591717.

Kjaer L and Vestergaard M (2002) *Local Partnerships in Europe. An Action Research Project Second Workshop Report.*

Kjaer L and Vestergaard M (2003) *Local Partnerships in Europe. An Action Research Project Second Workshop Report*.

Klein HK and Myers MD (1999) A set of principles for conducting and evaluating interpretive field studies in information systems. *MIS Quarterly: Management Information Systems 23*(1), 67–94. http://doi.org/10.2307/249410.

Klein T and Verhulst S (2017) *Access to New Data Sources for Statistics: Business Models and Incentives for the Corporate Sector.*

Klievink B, **Van Der Voort H and Veeneman W** (2018) Creating value through data collaboratives: Balancing innovation and control. *Information Polity 23*(4), 379–397. http://doi.org/10.3233/IP-180070.

Koschmann MA, **Kuhn TR and Pfarrer MD** (2012) A communicative framework of value in cross-sector partnerships. *Academy of Management Review 37*(3), 332–354.

Lapucci M and Cattuto C (2021) *Data Science for Social Good Philanthropy and Social Impact in a Complex World* (eds. Lapucci M and Cattuto C). Cham: Springer.

Lepri B, **Oliver N and Patrick P** (2018) Fair, transparent and accountable algorithmic decision-making processes. *Philosophy & Technology 31*(4), 611–627. https://doi.org/10.1007/s13347-017-0279-x.

Liva G, **Micheli M**, **Schade S**, **Kotsev A**, **Gori M and Codagnone C** (2023) City data ecosystems between theory and practice: A qualitative exploratory study in seven European cities. *Data and Policy 5*, e17. http://doi.org/10.1017/dap.2023.13.

Lockwood M (2009). Good governance for terrestrial protected areas: A framework, principles and performance outcomes. *Journal of Environmental Management 91*, 754–766.

Manning S and Roessler D (2014) The formation of cross-sector development partnerships: How bridging agents shape project agendas and longer-term alliances. *Journal of Business Ethics 123*(3), 527–547. http://doi.org/10.1007/s10551-013-1853-5.

Martin S, **Turki S and Kotsev A** (2021) *Establishment of Sustainable Data Ecosystems: Recommendations for the Evolution of Spatial Data Infrastructures.* http://doi.org/10.2760/04462.

Micheli M, **Farrell E**, **Carballa Smichowski B**, **Posada-Sànchez M**, **Signorelli S and Vespe M** (2023) *Mapping the Landscape of Data Intermediaries.*

Moretti LT, **Zahuranec A and Verhulst SG** (2022) The 9Rs Framework.

Nikander P and Elo T (2019) Will the Data Markets Necessarily Fail?

Nikander P, **Eloranta V**, **Karhu K and Hiekkanen K** (2020) Digitalisation, anti-rival compensation and governance: Need for experiments. In *Abstract from Nordic Workshop on Digital Foundations of Business, Operations, and Strategy*. Espoo, Finland: Aalto University, pp. 1–6.

Noriega-Campero A, **Garcia-Bulle B and Pentland AS** (2019) Active Fairness in Algorithmic Decision Making, pp. 77–83.

Otto B and Hompel M (2022) *Designing Data Spaces Designing Data Spaces.* http://doi.org/10.1007/978-3-030-93975-5.

Perkmann M and Schildt H (2015) Open data partnerships between firms and universities: The role of boundary organizations. *Research Policy 44*(5), 1133–1143. http://doi.org/10.1016/j.respol.2014.12.006.

Provan KG and Kenis P (2008) Modes of network governance: Structure, management, and effectiveness. *Journal of Public Administration Research and Theory 18*(2), 229–252. http://doi.org/10.1093/jopart/mum015.

Reischauer G, **Güttel WH and Schüssler E** (2021) Aligning the design of intermediary organisations with the ecosystem. *Industry and Innovation 28*(5), 594–619. http://doi.org/10.1080/13662716.2021.1879737.

Rhaiem K and Amara N (2021) *Learning from Innovation Failures: A Systematic Review of the Literature and Research Agenda Review of Managerial Science*, Vol. *15*. Berlin: Springer. http://doi.org/10.1007/s11846-019-00339-2.

Robin N, **Klein T and Jutting J** (2016) Public-private partnerships for statistics: lessons learned, future steps: A focus on the use of non-official data sources for national statistics. OECD Development Co-Operation, February.

Ruijer E (2021) Designing and implementing data collaboratives : A governance perspective. *Government Information Quarterly 38*, 101612. http://doi.org/10.1016/j.giq.2021.101612.

Sadabadi AA and Rahimi Rad Z (2022) How can cross-sector partnership promote social innovation? *Systemic Practice and Action Research 35*(4), 471–490. http://doi.org/10.1007/s11213-021-09581-w.

Savona M (2020) Governance of data value. *Policy Brief, Policy@Sussex 1*, 14.

Smichowski BC (2019) Alternative data governance models : Moving beyond one-size-fits-all solutions. *Intereconomics 11*(1), 222–227. http://doi.org/10.1007/s10272-019-0828-x.

Stalla-Bourdillon S, **Carmichael L and Wintour A** (2021) Fostering trustworthy data sharing: Establishing data foundations in practice. *Data & Policy 3*, e4. http://doi.org/10.1017/dap.2020.24.

Stalla-Bourdillon S, **Wintour A and Carmichael L** (2019) Building Trust Through Data Foundations a Call for a Data Governance Model to Support Trustworthy Data Sharing, 33. Available at https://cdn.southampton.ac.uk/assets/imported/transforms/content-block/UsefulDownloads_Download/E2360AAB5D274223BFDB863BAFC20F34/White Paper 2.pdf.

Suresh H, **Movva R**, **Dogan AL**, **Bhargava R**, **Cruxen I**, **Martinez Cuba A**, **Taurino G**, **So W and D'Ignazio C** (2022) Towards intersectional feminist and participatory ML: A case study in supporting feminicide counterdata collection. In *ACM International Conference Proceeding Series*. New York: ACM, pp. 667–678. http://doi.org/10.1145/3531146.3533132.

Susha I (2019) Establishing and Implementing Data Collaborations for Public Good: A Critical Factor Analysis to Scale Up the Practice.

Susha I, **Flipsen M**, **Agahari W and de Reuve GA** (n.d.) Towards Generic Business Models of Intermediaries in Data Collaboratives.

Susha I, **Flipsen M**, **Agahari W and de Reuve GA** (2020) Towards Generic Business Models of Intermediaries in Data Collaboratives. *IFIP Internationl Federation for Information Processing 2020.*

Susha I and Gil-Garcia JR (2019) A collaborative governance approach to partnerships addressing public problems with private data. In *Proceedings of the 52nd Hawaii International Conference on System Sciences*. Maui, Hawaii: HICSS, pp. 2892–2901. http://doi.org/10.24251/hicss.2019.350.

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. In *Proceedings of the 50th Hawaii International Conference on System Sciences (2017)*. Hawaii: HICSS, pp. 2691–2700. http://doi.org/10.24251/hicss.2017.325.

Susha I, **Pardo TA**, **Janssen M**, **Adler N**, **Verhulst SG and Harbour T** (2018) A research roadmap to advance data collaboratives practice as a novel research direction. *International Journal of Electronic Government Research 14*(3), 1–11. http://doi.org/10.4018/IJEGR.2018070101.

Susha I, **Grönlund Å and Van Tulder R** (2019a) Data driven social partnerships: Exploring an emergent trend in search of research challenges and questions. *Government Information Quarterly 36*(1), 112–128. http://doi.org/10.1016/j.giq.2018.11.002.

Susha I, **Rukanova B**, **Ramon Gil-Garcia TJ**, **Tan Y-H and Gasco Hernandez M** (2019b) Identifying Mechanisms for Achieving Voluntary Data Sharing in Cross-Sector Partnerships for Public Good. http://doi.org/10.1145/3325112.3325265.

Susha I, **van den Broek T**, **van Veenstra A-F and Linåker J** (2022) An ecosystem perspective on developing data collaboratives for addressing societal issues: The role of conveners. *Government Information Quarterly 40*(1), 101763. http://doi.org/10.1016/j.giq.2022.101763.

The New Hanse Project (2023) Governing Urban Data for The Public Interest, October, pp. 1–97.

Thomson AM and Perry JL (2006) Collaboration Processes: Inside the Black Box, December.

van den Broek T and van Veenstra AF (2018) Governance of big data collaborations: How to balance regulatory compliance and disruptive innovation. *Technological Forecasting and Social Change 129*, 330–338. http://doi.org/10.1016/j.techfore.2017.09.040.

van der Voort H (2017) Trust and cooperation over the public-private divide: An empirical study on trust evolving in co-regulation. In *Trust in Regulatory Regimes*. Cheltenham: Edward Elgar, pp. 181–204. http://doi.org/10.4337/9781785365577.00011.

van Donge W, **Bharosa N and Janssen MFWHA** (2022) Data-driven government: Cross-case comparison of data stewardship in data ecosystems. *Government Information Quarterly 39*(2), 101642. http://doi.org/10.1016/j.giq.2021.101642.

van Tulder R and Keen N (2018) Capturing collaborative challenges: Designing complexity-sensitive theories of change for cross-sector partnerships. *Journal of Business Ethics 150*(2), 315–332. http://doi.org/10.1007/s10551-018-3857-7.

Vangen S, **Hayes JP and Cornforth C** (2015) Governing cross-sector, inter-organizational collaborations. *Public Management Review 17*(9), 1237–1260. http://doi.org/10.1080/14719037.2014.903658.

Varshney KR and Mojsilović AM (2019) Open Platforms for Artificial Intelligence for Social Good: Common Patterns as a Pathway to True Impact. *International Conference on Machine Learning AI for Social Good Workshop.*

Verhulst S and Sangokoya D (2015) Mapping the Next Frontier of Open Data: Corporate Data Sharing | by Stefaan G. Verhulst | Internet Monitor 2014: Data and Privacy | Medium. Available at https://medium.com/internet-monitor-2014-data-and-privacy/mapping-the-next-frontier-of-open-data-corporate-data-sharing-73b2143878d2 (accessed 12 March 2021).

Verhulst SG and Young, A (2016) Open Data Impact Key Findings of the Open Data Impact Case Studies, March.

Verhulst SG and Saxena S (2022) The Need for New Methods to Establish the Social License for Data Reuse. Available at https://medium.com/data-policy/the-need-for-new-methods-to-establish-the-social-license-for-data-reuse-e7c67bdc4aff (accessed June 2023).

Verhulst SG, **Young A**, **Winowatan M and Zahuranec AJ** (2019) Leveraging Private Data for Public Good: A Descriptive Analysis and Typology of Existing Practices, October. Available at http://thegovlab.org/new-report-leveraging-private-data-for-public-good/ (accessed June 2023).

Verhulst SG, **Zahuranec AJ**, **Young A and Winowatan M** (2020) Wanted: Data Stewards, March.

Walsham G (1995) Interpretive case studies in IS research: Nature and method. *European Journal of Information Systems 4*(2), 74–81. http://doi.org/10.1057/ejis.1995.9.

Whiting L (2008) Semi-structured interviews for novice researhers. *Nursing Standard 22*(23), 35–40. Available at http://cat.inist.fr/?aModele=afficheN&cpsidt=20079971.

Williams S (2020) Data action using data for public good paper knowledge. In *Toward a Media History of Documents*, Vol. *3*. Cambridge, MA: MIT Press.

Yin RK (2009) *Case Study Research: Design and Methods*, 5th Edn. Thousand Oaks, CA: Sage.

Zuiderwijk A and Janssen M (2014) Open data policies, their implementation and impact: A framework for comparison. *Government Information Quarterly 31*(1), 17–29. http://doi.org/10.1016/j.giq.2013.04.003.

Cite this article: Bartolomucci F and Bartalucci S (2024). Designing data collaboratives' governance dimensions for long-term stability: an empirical analysis. *Data & Policy*, 6: e37. doi:10.1017/dap.2024.29