ORIGINAL PAPER

Data Governance in a Trilemma: A Qualitative Analysis of Rights, Values, and Goals in Building Data Commons

Jan Oleszczuk Zygmuntowski[1]

Abstract

Most of the discussions on data governance stress legal and technological aspects, while avoiding the discussion of power, stakeholder interests, and value. In order to move from formal debates to operational institution-building, we ask about the goals of data governance for different stakeholders and why data commons can be useful in achieving them. We discuss the evolution of data governance over time and cast light on the antagonism between three main goals of data governance: protecting fundamental rights, generating economic value and serving public interest. Given the necessity to navigate this conflicted political economy of data, we introduce the *data governance trilemma* (DGT) model. We turn to data commons as an institution capable of solving the collective action problem and negotiating acceptable configuration of the DGT goals. In order to operationalize data commons using DGT, we conducted workshops combining the critical success factors (CSFs) method with a deliberative Delphi technique. We find that the systemic configuration of data governance should be restructured towards data sovereignty, founded on institutional trust, protection of rights, and obligatory data sharing with the public interest in mind. Based on our results, we claim that unless we build data commons to steward data as a democratic medium, a lack of legitimization will riddle attempts to govern data better, and societal benefits will fail to manifest.

Keywords Data governance · Data commons · Political economy · Collective rights · Public interest · Deliberative Delphi · Critical success factors

1 Introduction

At a recent conference, a member of the audience asked a panel of grassroots data activists about the role of the public sector in data governance, only to hear that it should be "none." It was a baffling answer, given the involvement of public

✉ Jan Oleszczuk Zygmuntowski
jzygmuntowski@kozminski.edu.pl

[1] Kozminski University, Warsaw, Poland

welfare systems in data processing, but also a telling one that highlights the current state of discussions on data governance. The answer reflects the tendency to overlook the role of other key stakeholders, in this case the public sector, and more importantly, the implications of the goals they pursue in governing data. It seems that most of the discussions on data governance stress legal and technological aspects—the rather formal, objective obstacles to greater sharing—while avoiding discussions on political economy and their implications for power, stakeholder interests and goals, or generation and distribution of value (Sadowski et al., 2021).

A relational approach has emerged lately, thus shifting our attention from strictly individual to social harms—but also from the seemingly apolitical, market-focused "free data flow" to collective institutional designs (Viljoen, 2021). Most successful platforms enabling greater health data sharing are also marked by this evolution towards collective governance forms from previously siloed and later on managed individually (Kariotis et al., 2020). A growing body of scholarship attempts to systematically analyze data as commons, unpacking the social, economic, and political ramifications of data governance models and the implicit goals they serve (Fia, 2021; Prainsack, 2019; Purtova, 2017; Wong et al., 2022; Zygmuntowski et al., 2021).

However, an overview of existing scholarship suggests there is a need to expand the research agenda from ideal regulations into the study of social relations and communicative practices around common ownership of data (Hicks, 2022). Data commons literature remains very conceptual, with very few translations into operational frameworks, or further downstream—into practical research on building data commons. Hicks (2022) notes that "the bulk of the existing work on data ownership comes from legal scholars and technologists who, by the nature of their work, prefer formal regularity and consistency" while recognizing that the missing perspective approaches it "as an empirical, descriptive, and analytical task."

That task is not to eschew from the variety of actors involved in data governance, their goals and practices, but understand the relations between them and negotiate between their requirements and dealbreakers. It is exactly this gap that this article aims to fill in by answering two research questions:

RQ 1: What are the goals of data governance for different stakeholders?
RQ 2: How to build data commons that improve data governance?

The paper is structured as follows: in Sect. 2, we discuss the evolution of the notion of data governance over time and cast light on the antagonism between three main goals of data governance: protecting fundamental rights, generating economic value, and serving public interest. We also introduce the data governance trilemma (DGT) to navigate this conflicted political economy of data. In Sect. 3, we discuss why data commons can be useful as an institutional mechanism to solve collective action problem and negotiate acceptable DGT configurations. Finally, in Sect. 4, we describe the results of an empirical study consisting of a series of workshops on building data commons, combining the critical success

factors (CSFs) method with a deliberative Delphi technique. We tested not only the importance of resource, organizational and trust factors, but most importantly the usefulness of DGT model to choose governance goals. Section 5 provides further discussion of the results and explores possible avenues for future research.

Using the DGT for the analysis, we find that there is a sentiment for the restructuring of data governance towards greater protection of fundamental rights and increased recognition of the public interest in data. Apart from strict resource factors, institutional trust and stakeholder participation are also important, leading to the conclusion that data commons may be the right institution to embed collective data rights in governance. Unless data are democratized, societal benefits will fail to manifest.

The findings of this study should be accepted with understanding of its limitations. The paper makes a rather exploratory case, both theoretically and empirically. It is a preliminary assessment of the political economy of data governance and an operationalization of the challenge to bring the stakeholders together in order to build data commons.

2 The Antagonism in Data Governance

2.1 Evolution of Data Governance: from Internal Resource to Systemic Power

Not long ago, data governance was understood as "the exercise of authority and control over the management of data assets," conducted by organizations to "increase the value they get from their data assets" (DAMA International, 2017). Guided by business and technical drivers, the aim of data governance was to guide proper management of data as an internal company resource. Employee records, sales data from CRM systems, and business insights had to be properly managed within the organization not only for archival purposes, but for efficient use as an asset. Data governance referred to identifying "decision domains" and choosing managers accountable for the actual decision-making over data (Khatri & Brown, 2010). The perception of data governance was therefore limited to supervision within single organizational silo. Such a limited understanding did not take into account the realities of the expansion of data economy.

As data "colonized" an ever greater array of socio-economic life (Couldry & Mejias, 2019), and economic or political actors increasingly became dependent on the value captured from data flows, data governance became increasingly a matter of external arrangements between organizations. Thus, *data governance* aligned with the understanding of *governance* in social science. Informed by political scholarship, governance stresses the departure from hierarchical decision-making towards interdependencies between actors (Kooiman, 2003), driven by increasing institutional flexibility, participatory policymaking, and the power of networks (Levi-Faur, 2012). Governance is not restricted to state regulation, but involves "social interactions, cooperation and negotiations between stakeholders at the horizontal level" (Colebatch, 2014). In the end, stakeholders produce collectively binding decisions through both amical practices and competitive struggles. It is precisely the possibility to "govern by data" (Johns, 2021)—the non-administrative power to observe

and command populations by analyzing patterns and influencing behaviors—which posits data governance on a systemic level, as a cross-functional framework containing rights, obligations and procedures (Abraham et al., 2019). Data governance is thus not limited to a seemingly "neutral" technical ground, but is involved in social, economic and political conflicts. This is because data are created in heterogeneous data assemblages, such as ecosystems of technology, law, economic incentives, and socio-cultural conditions (Kitchin & Lauriault, 2014). Micheli et al. (2020) thus define data governance as "power relations between all the actors" and "various socio-technical arrangements set in place to generate value from data, and how such value is redistributed."

Because the concept of governance embraces the decentralized model of organizing society, it promotes participation and accountability in a normative way—as proper solutions to governance problems. Such "soft" regimes often benefit market actors more than they do civil society, which is what we observe in the data economy (Srnicek, 2017). Absent of strict regulation and strong institutions, the data economy has been shaped largely in order to privatize information and allow profit maximization (Bauwens et al., 2019; Fumagalli et al., 2019). The functioning of data as capital (Sadowski, 2019) mirrored the revolution of financialization: unrestricted, transnational data flows abound, and asymmetry of access to data further drives inequality. Not only privacy, but other fundamental rights are under threat from algorithmic profiling (Nemitz, 2018). This in turn results in conflicts such as aggressive and unfair competition, litigation, and disputes with regulators, as well as so-called *techlash* (technological backlash, a form of grassroots resistance to and withdrawal from digital platforms; see Syvertsen, 2020). With the deployment of artificial intelligence (e.g., large language models, stable diffusion) and the rising stakes of job automation and value capture, the conflicts over copyrights, compensations, and effectively decision-making between data producers (e.g., creators, artists) and data extractors (Big Tech companies) has not ceased but gained on intensity.[1]

Consequently, the agenda of data governance is no longer about "disciplining against forms of interpersonal violation" but rather seeks to (re)structure "the rules of economic production (and social reproduction) in the information economy" (Viljoen, 2021). Both the scholarship on data governance and policy debates are still looking for satisfactory models. To find it, one should first examine the antagonism between the goals of data governance stakeholders.

2.2 Three Goals of Data Governance

Governance by data continues the traditional governance's deployment of statistical techniques to manage populations, but the aim to "ensure [population] security and productivity" (Johns, 2021) is no longer universal. The discontinuation arises

---

[1] Consider the class action lawsuit filed against Stability AI Ltd.; Stability AI, Inc.; Midjourney Inc.; and DeviantArt, Inc., which "have created products that infringe the rights of artists and other creative individuals under the guise of alleged «artificial intelligence»" as plaintiffs claim. More information: https://stablediffusionlitigation.com/

because the stakeholders participating in data governance seek divergent goals, resulting in varying preferences for the socio-technical arrangement of governance. In our investigation of the data governance problem, we thus start by establishing the main goals of governance stakeholders. We follow the tripartite division into the state, market, and civil society. These macrosocial institutions have their origins in the work of Hegel, further developed by Marx, Gramsci, and more contemporary authors in the context of governance as the new paradigm of societal coordination (Jessop, 1998; Offe, 2000; Pelczynski, 1984).

In this framework, "ideal actors" of the state, market, and civil society are driven by an internal motivation, a rationale for their activity and a measure of success. Data governance goals are "meanings data represents for the interested actors" (Micheli et al., 2020). With regards to data, the state aims to collect them for administrative purposes, that is—to order according to larger public interest. For the Market, processing data is prerequisite to produce and transact better, thus increasing value output.[2] Whereas civil society strives to monitor misuses of data and act in defense of fundamental rights. We thus identify three main goals: protecting fundamental rights; generating economic value; and serving public interest.

Let us consider the limitation of such division to avoid reductionism. To begin with, in reality we face specific states, markets and civil societies, not abstracted "ideal actors." Real stakeholders are not limited to pursuing single goals. Actors participating in the data economy advance concepts which are idiosyncratic overlaps of motivations, reflecting their unclear priorities, political coalitions, or changing states of knowledge. Therefore, we briefly discuss each goal in the context of public debate and policy developments to ground them in empirically observed narratives and concepts.

The goal of protecting fundamental rights is often recognized through the right to privacy, although definitely is not limited to it. Various legislative efforts to protect privacy were started over a decade ago, and since the disclosure of the massive scale of surveillance in the digital world by Edward Snowden, broad coalitions have succeeded in passing regulations in Europe (Laurer & Seidl, 2021; Rossi, 2018), in US states, and some countries worldwide (Chander et al., 2021). The post-Snowden era fully recognizes data protection and data rights as goals of data governance, but given the asymmetry of power between data subjects and data controllers, the focus is rather on execution of rights and privacy-enhancing technologies (embracing privacy by design). Because there is a growing understanding of social harms in data, emerging on a population level, concepts like data justice seek to restructure data governance towards active prevention from discrimination, lack of representation or invisibility of people affected by decisions made by algorithms (Taylor, 2017). Ethical guidelines have also been proposed for AI systems in the context of data-derived products as detailed guidelines for algorithm designers, in order for systems to respect human autonomy, prevent harm, ensure fairness, and explainability (Niklas & Dencik, 2020).

---

[2] Whether that value is use-value beneficial for society or merely exchange-value to trade for profit is a separate discussion.

The goal of generating economic value is clearly linked to data-driven business operations, such as market insights, client profiling, offering artificial intelligence products, and services. However, these are still riddled with their own persisting problems. Monopolized data silos, gatekeeping, or fragmented systems are barriers to aggregation of data into larger sets or re-use of the existing ones, leading to less value being produced than possible with a well-governed, infrastructural approach to data. Various models are proposed for cross-sectoral partnership and exchange of data via data collaboratives (Klievink et al., 2018; Susha et al., 2017) or data trusts (Hardinges et al., 2019). Yet another view on value of data pertains to the uneven distribution of gains, with visions of personal data management or data ownership being presented as possible routes of emancipation of individuals from social structures (Micheli et al., 2020). A more pragmatic idea of data unions leverages the concept of collective bargaining to suggest wage negotiations between data subjects (cognitive workers) and actors profiting from data (Arrieta-Ibarra et al., 2018).

The goal of serving public interest stems from the administrative functions of the public sector and the knowledge-intensive needs of the welfare state. Increasingly, public and municipal managers take interest in data stewardship—the ethical and sustainable governance of data throughout its lifecycle (Verhulst et al., 2020b). Public sector often faces difficulties in accessing datasets crucial for operations related to collective wellbeing of the society, best exemplified by healthcare crisis during the COVID-19 pandemic. Data sharing mandates—requirements to share private data for public purposes—are gaining popularity both among experts (Alemanno, 2018) and regulators (i.a. EU's Data Act), while others recommend establishment of public data commons as institutions reshaping the entire digital economy (Zygmuntowski et al., 2021). So far, the tensions between state regulators and technology companies provoked greater interest in digital sovereignty at the local (such as Barcelona's government; see Monge et al., 2022; Morozov & Bria, 2018), national (French aspirations to *souveraineté numérique*), and supranational level (European ambitions expressed by the von der Leyen Commission).

2.3 Data Governance Trilemma

The problem of data governance can therefore be characterized as "the challenge of balance and enforcement" of "values, human rights, and public and private interests" (Zygmuntowski et al., 2021). The antagonism lays in whose values, rights and interests are prioritized. While stakeholders pursue their discrete combinations of goals, the outcome is a certain regime of data governance with all its institutions, legal rules, norms, and business models. Data governance is indeed not a collection of technical choices, but rather political economic ones.

A heuristic model of a trilemma can be created to make sense of these goals and how they are associated with or staked against one another, taking stock of the distinctive values and ideas that data governance brings into play. Figure 1 presents the data governance trilemma (DGT). The DGT triangle is set on three vertices: protecting fundamental rights, generating economic value, and serving the public interest. The three goals are interdependent to an extent that pursuing one of them leads

Fig. 1 Data governance trilemma
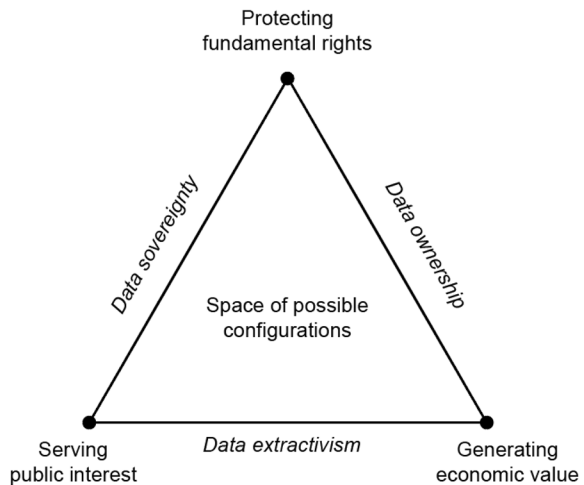
to departure from another one (Biga et al., 2022). Therefore, the edges connecting them are concepts that maximize two out of three goals but are antagonistic to the third goal. They are: data sovereignty (opposite to economic value), data ownership (opposite to public interest), and data extractivism (opposite to fundamental rights).

We perceive data sovereignty and data ownership as distinctive, separate concepts, which are advanced for different reasons and by different actors. Sovereignty stresses societal control and power over data (Hummel et al., 2021a, b), democratic legitimacy for such a control (Floridi, 2020; Roberts et al., 2021), or serves as a motivation for the public sector to develop its own infrastructure, enforce rights, and protect citizens (Calzada, 2019). On the other hand, ownership is mainly connected to fully individualized control over data (Lehtiniemi & Haapoja, 2019) and promises financial or symbolic gains from data governance (Hummel et al., 2021a, b). Data extractivism is perceived as a regime where firms and governments collaborate (either directly through market actions or indirectly through regulation) to use data in a subjugating, depleting, and non-reciprocal way (Hagolani-Albov et al., 2022).

As a result, inside the DGT triangle we observe a space of possible configurations of how society governs data. Inside it are both ideas and policy proposals attempting to improve data governance by suggesting a reconfiguration of systemic arrangements, countering or at the very least mitigating the perceived failings of the data governance currently in force. There is both an overlap and a conflict of data governance goals, stemming from various intersectional factors, from class interests, through logic of organizations, to temporary coalitions against overreach of power. Some governments bet on data-hungry companies to drive growth, while others expand the data rights of their citizens. Protecting freedom to conduct a business or the right to intellectual property[3] is not necessarily at odds with the goal of generating economic value.

---

[3] Both rights can be found e.g. in the Charter of Fundamental Rights of the EU as respectively: article 16 and 17.

We argue that all of these goals are in principle legitimate. Protecting fundamental rights secures a humane future, one which upends the instrumentarian alliance of the state and technology companies to extract data and rule society (Zuboff, 2019). Economic value in data is its power to increase resource efficiency and innovate, creating both tangible and intangible wealth outside the bureaucratic straightjacket. Whereas the public interest of societal welfare and provisioning of universal services is the myopically omitted piece in the world of fully-monetized data ownership.

However, acknowledging a valid interest in each goal does not imply that the pursuit of every goal is equally advantageous within a particular data governance system. The regime of surveillance capitalism is a result of widespread ideological maximalism denouncing rights and regulations in Cyberspace.[4] The currently existing data governance is not outside the DGT, nor in the center of it. Rather, we argue that it is skewed towards economic value, and contestability of such regime was diminished for many years because concerns over the other goals have been dismissed. The resulting social conflict over data rights and data sovereignty is a Polanyian "double movement" (Kenney et al., 2020) attempting to combat the excessive influence of the market logic on data governance and improve recognition of public interest and fundamental rights over data. Hence, new concepts emerge and coalitions form to shape data governance across various dimensions: regulations, technology, and culture.

It remains an open question to what extent the goals are antagonistic or not so much. Some theorists assert that political issues are by nature agonistic, meaning that they produce conflict over values and interests (Mouffe, 2005). Mouffe argues that "the political" is a constitutive part of human societies and cannot be eliminated or resolved completely. Instead, she emphasizes the importance of constructing a democratic "politics" that can accommodate and channel these inherent conflicts in a productive and inclusive manner. Agonism can be productive if there is an arena for conflict resolution. Then, the initial question can be reformulated as to whether DGT is a positive-sum game. And if so, how to build the right institutions for the antagonism in data governance?

3 Data Commons for Community Conflict Resolution

The ultimate decision on how to mediate the tensions between governance goals finds a strong legitimization in the sovereign, self-determined decision of the collective, because it is on the population level that social harms arise (McMahon et al., 2020). As Viljoen (2021) claims, "the relevant task of data governance is not to reassert individual control (…) but instead to develop the institutional responses necessary to represent the relevant population-level interests at stake in data production." Ensuring sustainable data governance can be perceived as a collective action problem, where

---

[4] For an example of such ideological maximalism, see John Perry Barlow's "A Declaration of the Independence of Cyberspace".

finding the right polycentric practices, such as boundary regulations or accountability, constitutes the commons (Benfeldt et al., 2020; Mindel et al., 2018).

One may understand commons-based approaches in the broadest sense, encompassing all frameworks that challenge the role of individual property as dominant means of organizing social relationships (Broumas, 2020; Marella, 2016), or, following the Ostromian tradition, treat the commons foremost as a social institution for managing common-pool resources by communities (Coyle et al., 2020; Ostrom, 1990; Prainsack, 2019). Although the study of the commons is widely associated with natural ecosystems, "a well-grounded domain of research exists focusing on shared knowledge, information, and data as objects and subjects of institutional governance" (Madison, 2020), namely the domain of knowledge commons. It is replete with findings on social groups, forms and flows which contribute to beneficial governance of intangible resources.

Reflecting on modern socio-technical assemblages, Frischmann (2012) redefines the commons as a strategy to govern infrastructural resources ("partially (non)rival goods"), a type of goods which are potentially sharable depending on capacity and flow of users. Like other infrastructures, data are "means for many ends," and thus should to be treated like shared resources to administer access of varying degree of control. There are cases when data functions as a public good—this is true for example for statistical data, where open access applies. But openness is only a particular form of access to data, which can fall prey to corporate, extractive practices (Bauwens & Niaros, 2017; Bodó, 2020). It is not a "tragedy of data commons"—to play on the unfortunate theory debunked on various occasions (Feeny et al., 1990)—but rather a failure to establish commons with specific governance rules. Openness lies on a spectrum of possible decisions on data governance, together with varied forms of permissioned access which are more suitable when data rights come into play (Taylor, 2017).

Increasingly, studies indicate an alignment between the commons and fundamental rights, as data commoning and protection of rights can be mutually reinforcing in the confrontation with surveillance and extraction for profit (Fia, 2021; Wong et al., 2022). Rights may empower generative and valuable use of data, while making sure the benefits are not appropriated by third party. At the same time, paying too much attention to strictly individual data rights slows down the development of socially valuable forms of data use. A classic example of this are barriers to medical research due to overarching privacy concerns, which commons-based and collective governance approaches try to solve (Kariotis et al., 2020).

Data commons can be therefore thought of as an institutional mechanism for negotiating and enforcing these choices, an arena of conflict resolution over access to data and infrastructure for utilizing data rights. Viljoen (2021) gives the example of "Waterorg"—a municipal public authority collecting household data to improve access to water—whose "basic governance structure allows for broader, democratic representation in the determination of societal goals." "Waterorg" is accountable to the local community, both in terms of execution of data governance strategy and the taking of responsibility for overreach; it also maintains the necessary infrastructure. Here, data commons go hand in hand with urban commons, as observed in various community wealth building projects (Webster et al., 2021) and especially in the case of Barcelona, which combined

its smart city strategy with digital sovereignty thanks to a data commons (Monge et al., 2022). We also observe attempts to establish data cooperatives and pool data as commons for collective access and value creation, with notable examples in industries such as financial services, healthcare, agriculture, energy provisioning, construction, and transportation (Bühler et al., 2023).

To what extent a given data commons will allow for collective data rights and direct control over access to data, whether it will provide valuable asset for (re)use by innovators and scientists and what will be the business model underpinning infrastructure, maintenance, cybersecurity: these are all possible configurations within the DGT. Instead of allowing for externalities, such as loss of trust, social unrest, or a decrease in data-driven innovation, the conflict may be resolved through internal procedures of debate, representative voting, or utilizing data rights in other ways directly (Zygmuntowski & Tarkowski, 2022). Data commons could serve as a collective negotiations tool to balance the three goals for a given data type, given community and governance needs of the moment. It is governance by data (Johns, 2021), yet on terms set by the population.

As stated in the introduction, scholarship faces the challenge of focusing less on formal discussions regarding legal definitions of data and instead grappling with the demands of empirical and operational analysis (Hicks, 2022). The commons are not one universal blueprint like hegemonic data governance models are (Carballa Smichowski, 2019), but need to be studied deeply in context, much like Ostrom studied natural common-pool resources. Given the wide array of different data governance models put forward (Micheli et al., 2020), we stress that data commons are not another model, but rather a set encompassing models of various legal and conceptual standing, associated by collective governance over data as a shared resource. In this view, data cooperatives are data commons, but data trusts only if provisions for collective governance are embedded.

Building data commons translates to designing mechanisms allowing communities to "collectively curate, inform, and protect each other through data sharing and the collective exercise of data protection rights" (Wong et al., 2022). Just as new institutions and democratic methods were established for all levels of societal ordering, including the right to nations' self-determination, there is a growing need now to create novel institutions embedding collective data rights in the participatory decision-making over data. Whether DGT model navigates through this challenge accurately and aids in operationalizing data commons, requires empirical validation.

4 Building Data Commons: a Critical Success Factors Analysis

The concept of data commons requires translation into operational, often sector-specific recommendations and pilot projects. Various industry and think-tank reports tackle this problem (Ctrl-Shift, 2018; Verhulst et al., 2020a), but only rarely fully acknowledging the political dimensions of data governance. Leveraging existing know-how of data sharing efforts and expanding it with the conscious negotiation of DGT goals is at the moment a strategic planning feat. Therefore, in this research, we ask what are the critical success factors (CSFs) for building data commons. Most

of the CSFs are elements related to the technological infrastructure, human staffing needs, or background for successfully establishing a new institution; yet, operating towards a specific goal can be treated as a factor as well.

4.1 Methodology

CSFs analysis can adopt various methods, such as case studies, structured interviewing, action research, and others. For the purpose of this research, we identified CSFs in the data governance scholarship and followed with workshops combining ranked poll with a deliberative Delphi method—a heuristic technique where the output of the poll was treated as an input for genuine deliberation and another round of inquiry (Glass et al., 2022). Our method follows Susha (2020), who conducted CSF research with data collaboratives. We adopted a three-stage approach: state of the art, identification of CSFs and relevance of CSFs.

4.1.1 Stage 1. State of the Art

This stage of research was described in the previous sections and resulted in adopting the DGT model.

4.1.2 Stage 2. Identification of CSFs

As a starting point, we adopted the top 15 critical success factors resulting from Susha's study on data collaboratives (2020). They were cross-checked with other data governance literature. Finally, we added additional factors based on scholarship and the DGT model that the previous study on data collaboratives, a particular model of data governance, did not take into account. Table 1 presents the final set of 21 factors, categorized into four types: organizational factors, resource factors, trust factors, and DGT goal factors.

4.1.3 Stage 3. Relevance of CSFs

Once we completed preparing the framework, we conducted two rounds of 2-day workshops with mixed groups of stakeholders from Polish public institutions (central and local government, executive agencies), think-tanks, data-driven companies (startups), and scholars. The workshops took place in April 2022 in the Chancellery of the Prime Minister of Poland and were a part of a larger research project led by the Instrat Foundation. A total of 31 experts participated in both rounds of the 2-day workshops. All the datasets generated and analyzed during the study are available from the corresponding author on request.

On the first day, the workshop consisted of presentations on data governance, regulatory changes and various models, as well as group analysis of the barriers for data sharing. On the second day, we run a ranked poll on CSFs of building data commons. Each expert was asked to answer the question: "which critical success

Table 1 List of critical success factors for data commons from the literature

No	Critical success factor	Select literature
Organizational factors		
1	Alignment of incentives of participants	(Jarke, 2020; Susha, 2020)
2	Appropriate business model	(Susha, 2020)
3	Collective (collaborative) governance	(Benfeldt et al., 2020; Jemielniak & Przegalinska, 2020; Kariotis et al., 2020)
4	Common terminology enabling integration	(Abraham et al., 2019; Susha, 2020)
5	Matching data supply with problem demand	(Hardinges et al., 2019; Susha, 2020)
6	Monitoring & sanctioning misuse	(Abraham et al., 2019; Prainsack, 2019)
Resource factors		
7	Availability of financial and human resources	(Hardinges et al., 2019; Susha, 2020)
8	Compatibility of technical infrastructure with interoperability standards	(Abraham et al., 2019; Susha, 2020)
9	Effective project leadership	(Abraham et al., 2019; Susha, 2020)
10	High quality of shared data	(Susha, 2020)
11	Regulatory data sharing mandate	(Alemanno, 2018; Zoboli, 2020)
12	Technical/analytical skillset of the team	(Benfeldt et al., 2020; Susha, 2020)
Trust factors		
13	Building trust in the institution	(O'Hara, 2019; Susha, 2020)
14	Clear responsibilities	(Susha, 2020)
15	Education and data stewardship ethics	(O'Hara, 2019; Verhulst et al., 2020b)
16	Participation of all stakeholders	(Ruhaak, 2020; Susha, 2020)
17	Public pressure, community and/or political support	(Benfeldt et al., 2020; Susha, 2020)
18	Shared understanding of objectives, values and outcomes	(Susha, 2020)
DGT goal factors[a]		
19	Economic value proposition	(Benfeldt et al., 2020; Jarke, 2020; Susha, 2020)
20	Designing methods to protect data & rights	(Stalla-Bourdillon et al., 2020; Wong et al., 2022)
21	Acting in public interest	(Rosenbaum, 2010; Zygmuntowski et al., 2021)

[a]The wording of DGT goal factors has been slightly redacted for the CSF. This reflects our choice to make the factors more operational, keeping them stylistically more uniform to reduce bias during workshops
Table of all 21 critical success factors divided in four categories (organizational, trust, resource, and DGT goal factors) and referenced to scholarly literature

factors are the most important for building data commons?" by individually ranking the CSFs from the most to the least important via a mobile web application. Participants essentially gave each option points from 1 (least important) to 21 (most important). We did not instruct participants on how to interpret "success," leaving it up for discussion. We then deployed the deliberative Delphi method: an automatically generated summary of the poll was displayed to everyone and a discussion commenced. We asked each expert both for explanations of individual ranking and for comment

on the total results. Learning what other experts think created a self-correcting feed-back loop. At the end of the workshop, we ran a second round of ranked poll on the exact same framework.

4.2 Findings

Figure 2 shows the results of the second ranked poll. Overall, resource factors are the most important as a category, receiving 31% of all points compared to 28% received by trust factors and 24% by organizational factors. The quality of data, their availability (including through regulatory intervention) and interoperability of data infrastructure is a prerequisite for governance, hence, those CSFs were ranked very high. However, along with the resource factors, we find that the relationship with community is the key to success, as indicated by the relative importance of institutional trust-building and stakeholder participation. Out of all organizational factors, the highest ranked is the one most associated with the function of commons as monitoring, access and exclusion management strategy.

Bar plot shows critical success factors in descending order of importance. The ranked score represents the average point allocation by workshop participants. Organizational factors are indicated in black, resource factors in blue, trust factors in green, and DGT (Data Governance Trilemma) goal factors in red.

The DGT goal factors received quite varied ranks. Both designing methods to protect data & rights and acting in public interest were ranked very high, whereas economic value proposition mattered just above average. The main explanation is that the ranking accurately reflects a sentiment for restructuring of data governance
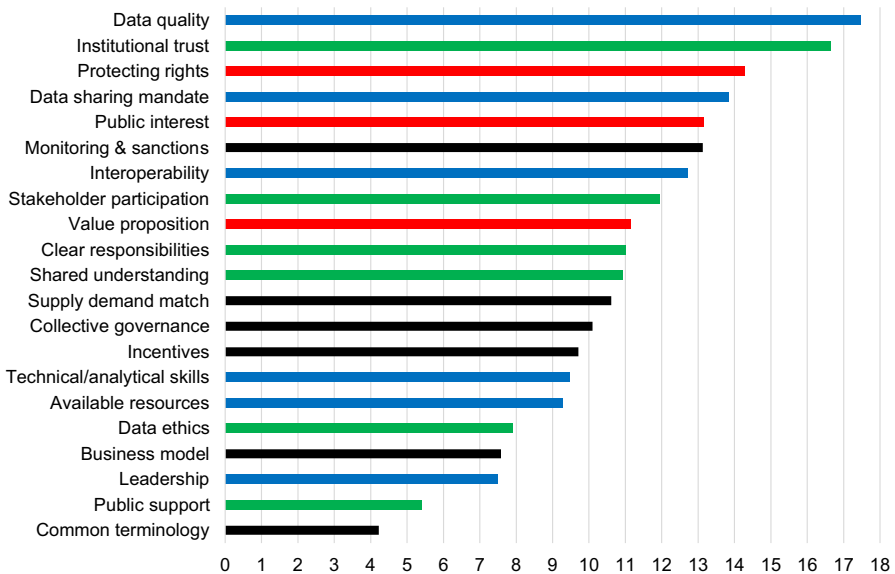

Fig. 2 Importance of critical success factors in building data commons

towards greater protection of fundamental rights and increased recognition of the public interest in data. Both of those goals score far higher than many of quite undeniably crucial factors, such as availability of resources, skills, or even a sound business model. In order to get a better understanding, we also look at the biggest rank changes before and after the Delphi method, presented on Fig. 3.

Bar plot shows the difference between the ranked poll before and after the application of the deliberative Delphi method. The bars represent the relative change, calculated as the difference divided by the score of the first poll. Only critical success factors (CSFs) with a change exceeding 10% or − 10% were selected for inclusion in the figure. Organizational factors are represented in black, resource factors in blue, trust factors in green, and DGT (Data Governance Trilemma) goal factors in red.

The only factor drastically losing about half of its score is common terminology, although one has to remember that the ranked poll method measures only *relative* importance. One concern may be that we observe "commonswashing," where the semantics of commons are appropriated for commercial purposes without endorsement of the values (Dulong de Rosnay, 2020). Alternative explanation is that stakeholders coming from various background use different terms and the innovative nature of building data commons leads to a certain "lack of language" to describe certain phenomena. Nevertheless, after alignment during the Delphi discussion this impediment was regarded as less detrimental as originally perceived.

We observe that the two most important CSFs—data quality and institutional trust—have gained considerably during group discussion. However, the change for public interest and protection of rights was insignificant, supporting the explanation that these goals were already regarded as important for the participants and remained so in the process. The goal of economic value proposition gained with the Delphi method, possibly reflecting a balancing effect of the discussion for finding
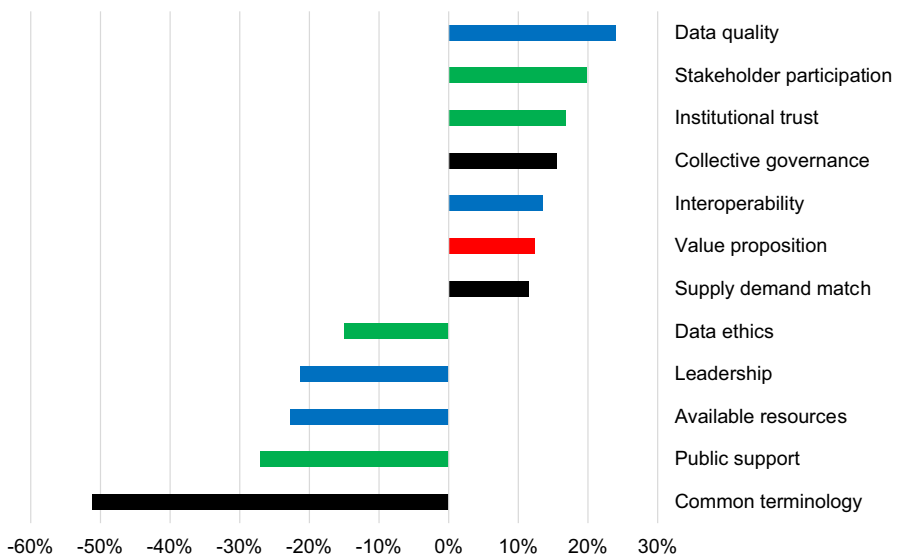
Fig. 3 CSFs rank change after Delphi method relative to first poll (select factors)

trilemma configurations that are not completely out-of-synch with one of the goals. The lowest scoring DGT goal was "rebalanced" after Delphi intervention. Therefore, we conclude that collective deliberation improves decision-making in a way unlikely to be attained in a scenario of individual data management.

5 Discussion and Conclusions

Some of the limitations of this study are the number of participants, their representativeness as experts or delegates of their organizations, and the ex-ante polling, preceding piloting data commons in practice. Our framework for studying CSFs for building data commons could be still improved upon and deployed on larger and varying groups to study the differences and similarities across countries or sector-specific data commons. Future studies could also implement methods such as case studies and comparative analysis to test the accuracy and usefulness of DGT model for designing data governance. Using empirical models such as IAD (Hess & Ostrom, 2007) or GKC (Frischmann et al., 2014) could prove fruitful for studying already existing and operational data commons, which is beyond the scope of this study.

The findings of this study should be interpreted as a preliminary assessment of building data commons by stakeholders interested in governing data better, yet not engaged in an existing project. Their responses emanate from their current data management operations, including open data projects, as well as intuitions what should be the next steps to govern data on a larger, even societal scale with substantive involvement of people as decision-makers and more effective re-usage of data. Whether such results can be extended to make sense of actually existing data commons is a question which merits a separate study.

We confirm Susha's (2020) findings on the very high importance of data quality. But compared to data collaboratives, the factors of trust, stakeholder participation and non-economic value goals, namely public interest and protection of rights, are regarded as critical for a broader challenge of building data commons. There is also a considerable interest in obligatory data sharing mandates for private firms. Overall, the tendency visible in these findings is to restructure, or renegotiate, the systemic configuration of data governance towards data sovereignty, founded on a dignitarian approach and supplemented with public interest-backing.

Based on the results, answering the call for greater operationalization of data commons should begin with sector-specific analysis of data quality, preferably involving the community producing the data. Stewarding trust requires transparency about means, goals of data governance, and clear recognition of rights in data. Data commons are also expected to engage with existing political economy actors by socializing siloed data and supporting public interest use of data.

These findings allow us to better understand what is required to govern data, which is inseparable from society it describes. Based on our results, we claim that unless we build data commons to steward data as a "democratic medium" (Viljoen, 2021), lack of trust will riddle attempts to govern data, and societal benefits will fail to manifest. Data commons may be useful to improve data governance, because they provide the

institutional space to embed data rights, negotiate rules and allow for decision-making over access to data. The considerations of rights, value, and goals are present in data and we should design with them in mind.

Data Availability   All the datasets generated and analyzed during the study are available from the corresponding author on reasonable request.

Declarations

Ethical Approval   Not applicable.

Informed Consent   All authors consent and approve of the content of this manuscript. The participants of the workshops have consented to the results being published before taking part in the workshop, and their identities are not revealed in the article.

Competing Interests   The author declares no competing interests.

References

Abraham, R., Schneider, J., & vom Brocke, J. (2019). Data governance: A conceptual framework, structured review, and research agenda. *International Journal of Information Management, 49*, 424–438. https://doi.org/10.1016/J.IJINFOMGT.2019.07.008

Alemanno, A. (2018). Big data for good: Unlocking privately-held data to the benefit of the many. *European Journal of Risk Regulation, 9*(2), 183–191. https://doi.org/10.1017/err.2018.34

Arrieta-Ibarra, I., Goff, L., Jiménez-Hernández, D., Lanier, J., & Weyl, E. G. (2018). Should we treat data as labor? Moving beyond "free." *AEA Papers and Proceedings, 108*, 38–42. https://doi.org/10.1257/pandp.20181003

Bauwens, M., Kostakis, V., & Pazaitis, A. (2019). *Peer to peer: The commons manifesto*. University of Westminster Press. https://doi.org/10.16997/book33

Bauwens, M., & Niaros, V. (2017). *Value in the commons economy: Developments in open and contributory value accounting*.

Benfeldt, O., Persson, J. S., & Madsen, S. (2020). Data governance as a collective action problem. *Information Systems Frontiers, 22*, 299–313. https://doi.org/10.1007/s10796-019-09923-z

Biga, B., Możdżeń, M., Oramus, M., & Zygmuntowski, J. (2022). Data governance : The state, the market, and the commons in the era of the fourth industrial revolution. In *Industrial Revolution 4.0* (pp. 221–240). Routledge. https://doi.org/10.4324/9781003264170-11

Bodó, B. (2020). Was the open knowledge commons idea a curse in disguise? – Towards sovereign institutions of knowledge. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3502119

Broumas, A. (2020). *Intellectual commons and the law: A normative theory for commons-based peer production*. University of Westminster Press.

Bühler, M. M., Calzada, I., Cane, I., Jelinek, T., Kapoor, A., Mannan, M., Mehta, S., Micheli, M., Mookerje, V., Nübel, K., Pentland, A., Scholz, T., Siddarth, D., Tait, J., Vaitla, B., & Zhu, J. (2023). *Data cooperatives as catalysts for collaboration, data sharing, and the (trans) formation of the digital commons*. https://doi.org/10.20944/PREPRINTS202304.0130.V1

Calzada, I. (2019). Technological sovereignty: Protecting citizens' digital rights in the AI-driven and post-GDPR algorithmic and city-regional European realm. *Regions*. https://doi.org/10.1080/13673882.2018.00001038

Carballa Smichowski, B. (2019). Alternative data governance models: Moving beyond one-size-fits-all solutions. *Intereconomics, 54*(4), 222–227. https://doi.org/10.1007/s10272-019-0828-x

Chander, A., Kaminski, M. E., & McGeveran, W. (2021). Catalyzing privacy law. *Minnesota Law Review, 105*(4), 1733–1802. https://doi.org/10.2139/SSRN.3433922

Colebatch, H. K. (2014). Making sense of governance. *Policy and Society, 33*(4), 307–316. https://doi.org/10.1016/J.POLSOC.2014.10.001

Couldry, N., & Mejias, U. A. (2019). Data colonialism: Rethinking big data's relation to the contemporary subject. *Television & New Media, 20*(4), 336–349. https://doi.org/10.1177/1527476418796632

Coyle, D., Diepeveen, S., Wdowin, J., Kay, L., & Tennison, J. (2020). *The value of data: Policy implications*.

Ctrl-Shift. (2018). *Data mobility: The personal data portability growth opportunity for the UK economy*.

DAMA International. (2017). *DAMA-DMBOK: Data management body of knowledge (2nd edition)*. Technics Publications.

Dulong de Rosnay, M. (2020). Commonswashing – A political communication struggle. *Global Cooperation Research - A Quarterly Magazine, 2*(3). https://hal.science/hal-02986722

Feeny, D., Berkes, F., McCay, B. J., & Acheson, J. M. (1990). The tragedy of the commons: Twenty-two years later. *Human Ecology, 18*(1), 1–19. https://doi.org/10.1007/BF00889070/METRICS

Fia, T. (2021). An alternative to data ownership: Managing access to non-personal data through the commons. *Global Jurist, 21*(1), 181–210. https://doi.org/10.1515/gj-2020-0034

Floridi, L. (2020). The fight for digital sovereignty: What it is, and why it matters, especially for the EU. *Philosophy & Technology 2020 33:3*, *33*(3), 369–378. https://doi.org/10.1007/S13347-020-00423-6

Frischmann, B. (2012). *Infrastructure: The social value of shared resources*. Oxford Univesity Press.

Frischmann, B. M., Madison, M. J., & Strandburg, K. J. (2014). Governing knowledge commons. In *Governing Knowledge Commons*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2490622

Fumagalli, A., Giuliani, A., Lucarelli, S., Vercellone, C., & Negri, A. (2019). Cognitive capitalism, welfare and labour. *Routledge*. https://doi.org/10.4324/9781315623320

Glass, J., McMorran, R., Currie, M., McKee, A., Pinker, A., Reed, M., Meador, E., & Markantoni, M. (2022). Translating community resilience theory into practice: A deliberative Delphi approach. *Sociologia Ruralis*. https://doi.org/10.1111/SORU.12397

Hagolani-Albov, S. E., Chagnon, C. W., Durante, F., Gills, B. K., Hokkanen, S., Kangasluoma, S. M. J., Konttinen, H., Kröger, M., LaFleur, W., Ollinaho, O., & Vuola, M. P. S. (2022). From extractivism to global extractivism: The evolution of an organizing concept. *The Journal of Peasant Studies, 49*(4), 760–792. https://doi.org/10.1080/03066150.2022.2069015

Hardinges, J., Wells, P., Blandford, A., Tennison, J., & Scott, A. (2019). *Data trusts: Lessons from three pilots*. https://docs.google.com/document/d/118RqyUAWP3WIyyCO4iLUT3oOobnYJGibEhspr2v87jg/edit

Hess, C., & Ostrom, E. (2007). *A framework for analyzing the knowledge commons*. https://ieeexplore.ieee.org/abstract/document/6284192/

Hicks, J. (2022). The future of data ownership: An uncommon research agenda. *The Sociological Review*, *0*(0). https://doi.org/10.1177/00380261221088120

Hummel, P., Braun, M., & Dabrock, P. (2021a). Own data? Ethical reflections on data ownership. *Philosophy and Technology, 34*(3), 545–572. https://doi.org/10.1007/S13347-020-00404-9

Hummel, P., Braun, M., Tretter, M., & Dabrock, P. (2021b). Data sovereignty: A review. *Big Data & Society, 8*(1), 205395172098201. https://doi.org/10.1177/2053951720982012

Jarke, M. (2020). Data Sovereignty and the Internet of Production. *Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics), 12127 LNCS*, 549–558. https://doi.org/10.1007/978-3-030-49435-3_34/FIGURES/7

Jemielniak, D., & Przegalinska, A. (2020). *Collaborative society*. MIT Press Essential Knowledge series.

Jessop, B. (1998). The rise of governance and the risks of failure: The case of economic development. *International Social Science Journal, 50*(155).

Johns, F. (2021). Governance by data. *Annual Review of Law and Social Science, 17*, 53–71. https://doi.org/10.1146/ANNUREV-LAWSOCSCI-120920-085138

Kariotis, T., Ball, M. P., Tzovaras, B. G., Dennis, S., Sahama, T., Johnston, C., Almond, H., & Borda, A. (2020). Emerging health data platforms: From individual control to collective data governance. *Data & Policy, 2*. https://doi.org/10.1017/DAP.2020.14

Kenney, M., Zysman, J., & Bearson, D. (2020). What Polanyi teaches us about the platform economy and structural change. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3678967

Khatri, V., & Brown, C. (2010). Designing data governance. *Communications of the ACM, 53*(1), 148–152. https://doi.org/10.1145/1629175.1629210

Kitchin, R., & Lauriault, T. (2014). *Towards critical data studies: Charting and unpacking data assemblages and their work*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2474112

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

Kooiman, J. (2003). *Governing as governance*. Sage Publications. https://philpapers.org/rec/KOOGAG

Laurer, M., & Seidl, T. (2021). Regulating the European data-driven economy: A case study on the general data protection regulation. *Policy & Internet, 13*(2), 257–277. https://doi.org/10.1002/POI3.246

Lehtiniemi, T., & Haapoja, J. (2019). Data agency at stake: My data activism and alternative frames of equal participation. *New Media & Society, 22*(1), 87–104. https://doi.org/10.1177/1461444819861955

Levi-Faur, D. (2012). The Oxford handbook of governance. *The Oxford Handbook of Governance*, 1–832. https://doi.org/10.1093/OXFORDHB/9780199560530.001.0001

Madison, M. (2020). Tools for data governance. *Technology and Regulation, 29*. https://scholarship.law.pitt.edu/fac_articles/394

Marella, M. R. (2016). The Commons as a Legal Concept. *Law and Critique 2016 28:1, 28*(1), 61–86. https://doi.org/10.1007/S10978-016-9193-0

McMahon, A., Buyx, A., & Prainsack, B. (2020). Big data governance needs more collective responsibility: The role of harm mitigation in the governance of data use in medicine and beyond. *Medical Law Review, 28*(1), 155–182. https://doi.org/10.1093/MEDLAW/FWZ016

Micheli, M., Ponti, M., Craglia, M., & Berti Suman, A. (2020). Emerging models of data governance in the age of datafication. *Big Data & Society, 7*(2), 205395172094808. https://doi.org/10.1177/2053951720948087

Mindel, V., Mathiassen, L., & Rai, A. (2018). The sustainability of polycentric information commons. *MIS Quarterly, 42*(2), 607–631. https://doi.org/10.25300/MISQ/2018/14015

Monge, F., Barns, S., Kattel, R., & Bria, F. (2022). *A new data deal: The case of Barcelona*. https://www.ucl.ac.uk/bartlett/public-purpose/wp2022-02

Morozov, E., & Bria, F. (2018). *Rethinking the smart city*. www.rosalux-nyc.org

Mouffe, C. (2005). *On the political*. Routledge.

Nemitz, P. (2018). Constitutional democracy and technology in the age of artificial intelligence. *Philosophical Transactions of the Royal Society a: Mathematical, Physical and Engineering Sciences, 376*(2133), 20180089. https://doi.org/10.1098/rsta.2018.0089

Niklas, J., & Dencik, L. (2020). *European artificial intelligence policy: Mapping the institutional landscape*. https://orca.cardiff.ac.uk/id/eprint/133252

Offe, C. (2000). Civil society and social order: Demarcating and combining market, state and community. *European Journal of Sociology / Archives Européennes De Sociologie, 41*(1), 71–94. https://doi.org/10.1017/S000397560000789X

O'Hara, K. (2019). *Data trusts ethics, architecture and governance for trustworthy data stewardship*. WSI White Paper #1

Ostrom, E. (1990). *Governing the commons: The evolution of institutions for collective action*. Cambridge University Press.

Pelczynski, Z. A. (1984). *The state and civil society: Studies in Hegel's Political Philosophy*. Cambridge University Press.

Prainsack, B. (2019). Logged out: Ownership, exclusion and public value in the digital data and information commons. *Big Data & Society*, *6*(1). https://doi.org/10.1177/2053951719829773

Purtova, N. (2017). Health data for common good: Defining the boundaries and social dilemmas of data commons. In Samantha Adams, Nadezhda Purtova, & Ronald Leenes (Eds.), *Under Observation: The Interplay Between eHealth and Surveillance* (pp. 177–210). Springer, Cham. https://doi.org/10.1007/978-3-319-48342-9_10

Roberts, H., Cowls, J., Casolari, F., Morley, J., Taddeo, M., & Floridi, L. (2021). Safeguarding European values with digital sovereignty: An analysis of statements and policies. *Internet Policy Review*, *10*(3). https://doi.org/10.14763/2021.3.1575

Rosenbaum, S. (2010). Data governance and stewardship: Designing data stewardship entities and advancing data access. In *Health Services Research* (Vol. 45, Issue 5, part 2, pp. 1442–1455). https://doi.org/10.1111/j.1475-6773.2010.01140.x

Rossi, A. (2018). How the Snowden revelations saved the EU general data protection regulation. *International Spectator*, *53*(4), 95–111. https://doi.org/10.1080/03932729.2018.1532705

Ruhaak, A. (2020). When one affects many: The case for collective consent. *Mozilla Foundation*. https://foundation.mozilla.org/en/blog/when-one-affects-many-case-collective-consent/

Sadowski, J. (2019). When data is capital: Datafication, accumulation, and extraction. *Big Data & Society*, *6*(1). https://doi.org/10.1177/2053951718820549

Sadowski, J., Viljoen, S., & Whittaker, M. (2021). Everyone should decide how their digital data are used — Not just tech companies. *Nature 2021 595:7866*, *595*(7866), 169–171. https://doi.org/10.1038/d41586-021-01812-3

Srnicek, N. (2017). *Platform Capitalism*. Polity Press.

Stalla-Bourdillon, S., Thuermer, G., Walker, J., Carmichael, L., & Simperl, E. (2020). Data protection by design: Building the foundations of trustworthy data sharing. *Data & Policy*, *2*. https://doi.org/10.1017/dap.2020.1

Susha, I. (2020). Establishing and implementing data collaborations for public good: A critical factor analysis to scale up the practice. *Information Polity, 25*(1), 3–24. https://doi.org/10.3233/IP-180117

Susha, I., Janssen, M., & Verhulst, S. (2017). Data collaboratives as a new frontier of cross-sector partnerships in the age of open data: Taxonomy development. *Proceedings of the 50th Hawaii International Conference on System Sciences (2017)*. https://doi.org/10.24251/hicss.2017.325

Syvertsen, T. (2020). Digital detox: The politics of disconnecting. *Digital Detox: The Politics of Disconnecting*. https://doi.org/10.1108/9781787693395

Taylor, L. (2017). What is data justice? The case for connecting digital rights and freedoms globally. *Big Data & Society, 4*(2), 205395171773633. https://doi.org/10.1177/2053951717736335

Verhulst, S. G., Young, A., Zahuranec, A. J., Aaronson, S. A., Calderon, A., & Gee, M. (2020a). *The emergence of a third wave of open data*.

Verhulst, S. G., Zahuranec, A. J., Young, A., & Winowatan, M. (2020b). *Wanted: Data stewards. (Re-) Defining The Roles And Responsibilities Of Data Stewards For An Age Of Data Collaboration*.

Viljoen, S. (2021). A relational theory of data governance. *The Yale Law Journal, 131*(2), 573–781. https://doi.org/10.2139/SSRN.3727562

Webster, A., Kuznetsova, O., Ross, C., Berranger, C., Booth, M., Eseonu, T., & Golan, Y. (2021). Local regeneration and community wealth building–place making: Co-operatives as agents of change. *Journal of Place Management and Development, 14*(4), 446–461. https://doi.org/10.1108/JPMD-01-2020-0009/FULL/XML

Wong, J., Henderson, T., & Ball, K. (2022). Data protection for the common good: Developing a framework for a data protection-focused data commons. *Data & Policy*, *4*, e3. https://doi.org/10.1017/DAP.2021.40

Zoboli, L. (2020). Fueling the European digital economy: A regulatory assessment of B2B data sharing. *European Business Law Review, Forthcoming*. https://doi.org/10.2139/ssrn.3521194

Zuboff, S. (2019). *The age of surveillance capitalism: The fight for freedom and power in the age of surveillance capitalism*. Public Affairs New York.

Zygmuntowski, J., Zoboli, L., & Nemitz, P. F. (2021). Embedding European values in data governance: A case for public data commons. *Internet Policy Review*, *10*(3). https://doi.org/10.14763/2021.3.1572

Zygmuntowski, J., & Tarkowski, A. (2022). *Data Commons Primer*. https://openfuture.eu/wp-content/uploads/2022/07/220723data-commons-primer.pdf