Managing public sector data: National challenges in the context of the European Union's new data governance models

Clarissa Valli Buttow* and Sophie Weerts
Swiss School of Public Administration, University of Lausanne, Lausanne, Switzerland

Abstract. In its regulatory enterprise to improve the conditions of data sharing and reuse, the European Union has enacted new legislation: the Data Governance Act (DGA). The DGA envisages new forms of sharing public sector data (PSD). Based on a legal analysis of the DGA and an in-depth study of data governance literature, this paper highlights what is at stake in the new regulatory framework and argues that more than the mere openness of more PSD will be necessary to ensure that the European Union policy goals are achieved, especially those concerning enhancing innovation for the common good. From this perspective, the paper argues that the public data trust model of data governance and the Responsible Research and Innovation approach offer two powerful tools for public sector data governance. In this context, this paper contributes to the debate about new data governance models and discusses tools and frameworks enabling the use of data for the common good. It also provides insights to public administration practitioners aiming to implement a framework for increased and sustainable PSD sharing.

Keywords: Data governance, public sector data, open government data, public data trust, data for common good

Key points for practitioners:

- The implementation of the European Union's Data Governance Act will entail several challenges to national authorities, especially in what concerns new forms of public sector data sharing;
- Public Data Trusts can serve as an institutional model for new public sector data-sharing possibilities;
- Responsible research innovation offers a framework for democratic public sector data management.

1. Introduction

Data is at the core of technological innovation and is perceived as an important source of economic growth and social good (European Commission, 2020b). This potential lies in the possibilities of data being aggregated and combined, enabling new insights, predictions and content. Such aggregation is seen as unlocking economies of scale and scope that "would materialize in efficiency gains for the industry and increased social welfare" (European Parliament, 2021, p. 2). However, in the current state of affairs, data and the technological capacity to explore it are concentrated in the hands of a few actors and benefit first and foremost powerful private groups pursuing exclusively economic interests (Micheli, Ponti, Craglia, & Berti Suman, 2020). This situation, also referred to as the platform economy, reinforces the power imbalance in modern societies and hampers innovation (Pistor, 2020; Taylor, Mukiri-Smith,

*Corresponding author: Clarissa Valli Buttow, Swiss School of Public Administration, University of Lausanne, Lausanne, Switzerland. E-mail: clarissa.vallibuttow@unil.ch.

Petročnik, Savolainen, & Martin, 2022). Moreover, the most important data uses are linked with targeted advertising, personalised insurance premiums and credit score ratings (Ohm & Frankle, 2018; Viljoen, 2021). Although significant from an economic point of view, these applications are of limited relevance to improving social welfare (Viljoen, Goldenfein, & McGuigan, 2021). In this context, the concentration of technological capacity limits the bottom-up processes of social innovation through which civil society, social entrepreneurs and small businesses tackle social needs that are generally not addressed by the market or by players motivated solely by profit (Taylor et al., 2022).

In this framework, governments hold crucial sources of data. Despite this, the bulk of data controlled by governments remains underexploited by the government itself and continues inaccessible to private entrepreneurs due to, on one hand, legal and technical constraints to making it accessible and, on the other hand, the lack of internal technical structure and human skills to exploit this data fully (Bharosa, 2022; Taylor, 2017). The openness of what is called 'Public Sector Data' (PSD) has been an object of public policy and regulation for a long time (Valli Buttow & Weerts, 2022b), and it has been promoted by the Open Government Data (OGD) movement. The liberalisation of PSD was first motivated by arguments such as its potential for improving democracy and boosting the economy (Valli Buttow & Weerts, 2022a); now it is used as an important tool to fight data concentration and to foster innovation that is aligned with the public good (European Commission, 2020c).

To make data available while tackling imbalances of power, the European Union (EU) has engaged in a regulatory effort to foster a digital transformation that promotes human flourishing and is aligned with European values (European Commission, 2020d). As part of this broader effort to regulate new technologies and the digital market, the Data Governance Act (DGA) proposes increased data openness. It aims to improve the conditions of data sharing and, by so doing, to boost innovation and ensure that society as a whole profits from technological advances (European Commission, 2020c). With regard to PSD, the DGA aims to complement the regulatory framework established by the Open Data Directive (European Union, 2019), which is closely linked with the OGD movement (Valli Buttow & Weerts, 2022b). On this subject, the European regulation defines data as the digital representation of acts, facts or information (European Union, 2022, art.2 (1)) and PSD as the data produced or collected as a result of the execution of public tasks over which governments exercise legal and de facto control (European Union, 2022). The PSD open definition is justified because, within the EU law, the scope of public tasks is defined by national states. The attribution of data as public can therefore vary in line with different national legal regimes (European Union, 2022, recital 8).

To sum up, this new legal framework expands and deepens the possibilities of data sharing based on the adaptation of the OGD model. However, it is generally agreed that the OGD on which the new European regulation is built has had limited positive results. Scholars have noted an important gap between OGD promises and real-world impact (Jetzek, Avital, & Bjørn-Andersen, 2013; Reggi, Dawes, & Gil-Garcia, 2022; Ruijer & Meijer, 2020; Safarov, Meijer, & Grimmelikhuijsen, 2017; Schwoerer, 2022). The literature identifies several barriers that hinder OGD potential such as a lack of opportunities, motivation and ability to use public data (demand side), a lack of appropriate institutional incentives to share (supply side), and a lack of adequate governance mechanisms to gather and act upon, users' feedback (Jetzek, 2016; Reggi et al., 2022).

Achieving the DGA's ambitious policy goals poses a critical challenge to scholars, policymakers and practitioners in the field. In this context, this paper aims to investigate the following research question: What instruments and tools could be used to meet the DGA's objectives, i.e., to ensure that PSD is open to the various political and economic players while guaranteeing that it can be reused for the common good? To answer this question, we conducted an in-depth analysis of the literature on data governance, OGD

and Responsible Research and Innovation (RRI) and undertook an analysis of the DGA preparatory and pre-legislative work complemented by a legal analysis of the DGA final text in the light of a law-in-context approach. We argue that the adoption of collective and democratic management of PSD, based on a public data trust institutional design rooted in the RRI framework, is an adequate response to the challenges posed by the DGA implementation. This article addresses what we believe is a gap in the current literature concerning the study of new institutional settings and policy instruments that enable the use of data in more democratic ways designed to ensure that society as a whole benefits from data use (Beaulieu & Leonelli, 2022; Pagallo, 2022; Prainsack, 2020). This article also contributes from a practical point of view to the debate on the DGA implementation by national governments.

The following section (two) presents the conceptual background needed to understand the issue of data governance and the theoretical basis of RRI. Section three outlines the DGA and considers it in its policy context, emphasising the novelties concerning PSD governance. Section four links the theoretical framework described in Section two to the legal framework investigated in Section three and explores possible policy options to realise the potential of PSD for the common good. Section five presents our conclusions.

2. Data governance and responsible research and innovation

This section presents the conceptual backgrounds for PSD governance, specifically focused on data governance and RRI. Each conceptual background is constructed using the relevant scholarly literature consulted by browsing the main scientific journals in the fields of technology, regulation, and innovation. We studied relevant articles and chapters and checked references to confirm and thicken the concepts. This background provides a deeper understanding of the issues at stake when discussing PSD and provides insights on tools (institutional and organisational) that could be useful when implementing the DGA and pursuing the DGA policy goals.

2.1. The concept of data governance and its dominant models

The concept of data governance has gained the attention of different fields in the literature, but its conceptual boundaries are still being determined. The expression was first used in the realm of Information Management, and it was thought of as the system of governance used to manage data inside different organisations (Abraham, Schneider, & vom Brocke, 2019). In Science and Technology Studies (STS), authors emphasise that data governance entails the "power relations between all the actors affected by, or having an effect on, the way data is accessed, controlled, shared and used, the various socio-technical arrangements set in place to generate value from data, and how such value is redistributed between actors" (Micheli et al., 2020, p. 3). In critical data studies, Kitchin explains that the concept of data governance designates the idea of a system of actors with different entitlements and several legal and organisational layers that define how value is extracted from data (Kitchin, 2022, pp. 278–279). In socio-legal scholarship, the concept of data governance captures both the managerial element and the question of power/entitlements. In this sense, Ducuing explains that "from a policy and regulatory perspective, data governance can be defined as a system of rights and responsibilities that determine who can take what actions with what data." (Ducuing, 2020, p. 59). Additionally, from a policy perspective, data governance can be understood broadly as the ensemble of technical tools and mechanisms, legislative and regulatory actions that aim to "enhance the creation of value from data" (Graef, 2020, p. 25). In brief, all these perspectives, and more specifically, the perspectives of critical data studies and the social-legal
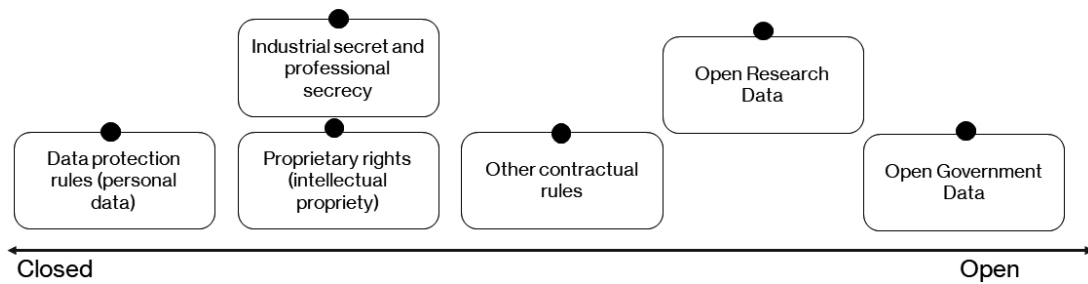
Fig. 1. Openness spectrum.

scholarship, highlight the social constructions that enable the use and reuse of data and, consequently, the extraction of value from them.

Two crucial and entangled issues are at the heart of the issue of data governance: the concentration of data in the hands of just a few actors and the stewardship of data-driven innovation toward the common good. Data governance models determine how and by whom data can be accessed, used and reused (Beaulieu & Leonelli, 2022). There is a variety of rules and regulations that define the conditions in which data governance takes place. These regulations encompass data protection rules, proprietary rights, security and secrecy rules, contractual conditions, and the still unsolved question of data ownership, among others. Different types of data will be subject to different rules and regulations, that will also depend on the context of their use. Broadly speaking, the different possibilities of data governance will be found anywhere along the spectrum between entirely closed and completely open access. Figure 1 illustrates how the regulatory framework determines the place of data under this spectrum. It is important to note that any data in this spectrum can be held or controlled by either public entities (such as governments of different spheres, local, regional, national or even international authorities) or private actors.

The current legal and political framework of data governance enabled the rise of the platform economy and data-driven markets with the winner-take-all structural logic, leading to the concentration of data and technical resources in a few giant tech companies (Pistor, 2020; Stilgoe, 2020). The exclusive access enabled by the current legal framework locks data into silos and hampers innovation. Indeed, small actors with good business ideas could find insurmountable difficulties in going further with their ventures if access to data is too difficult or, in some cases, impossible (Beaulieu & Leonelli, 2022; Verdegem, 2022). The second aspect that becomes salient in the discussion on data governance concerns the final use of data, the desirability of innovation and the legitimacy of the decisions concerning the direction and speed of innovation. Much of the recent technological development has been driven by giant corporations that retain the power to decide the areas of innovation, having as their core value the pursuit of economic gains (Birch, Chiappetta, & Artyushina, 2020; Taylor, 2021; Viljoen, 2021). In this sense, the question of the purpose of data-driven innovation underlines the discussion about increased access to data and improved data-sharing conditions (Taylor et al., 2022).

While the 'closed' side of data governance had led to data and technology concentration in the hands of a few powerful actors, leaving small space for social innovation, the open regime of freely and widely accessible data embodied by the OGD movement has produced only modest outcomes. In this context, research into OGD's features, barriers and effects provides insightful inputs concerning the process of opening up data and its consequences. Within this literature, there is an emerging consensus that the OGD's initial ambitious promises are difficult to realise in the real world (Janssen, 2012; Jetzek et al., 2013; Murillo, 2015; Reggi et al., 2022; Schwoerer, 2022). Some authors have suggested that the real potential of OGD is quite limited (Ruijer & Meijer, 2020). Others suggest that the results of widespread

OGD availability are not meeting expectations (Murillo, 2015; Susha, Grönlund, & Janssen, 2015). Safarov and colleagues indicate a general lack of evidence of the causal links between OGD and positive outcomes and point out that "it is not enough to assume that the effects will occur, and that this will occur automatically" (Safarov et al., 2017, pp. 18–19). Khayyat and Bannister highlight the weakness of misguided ideas about openness that ignore the complexity of license regimes and argue that an adequate license regime could offer a better solution to the management of PSD (Khayyat & Bannister, 2015). Jaakola and colleagues note that it is not enough to open data, it is necessary to promote and encourage the use of that data (Jaakola, Kekkonen, Lahti, & Manninen, 2015).

The literature thus shows that the simple provision of access to data does not ensure its reuse nor that its potential reuse will produce results aligned with social goals. Research to date corroborates this view, emphasizing the need for improving mechanisms of feedback (Reggi et al., 2022) and for constructing a closer collaboration between public sector bodies and potential PSD users (Ruijer & Meijer, 2020; Smith & Sandberg, 2018). Furthermore, one should not be blind to the commercial interests behind the movement to open up PSD (Bates, 2014; Birch et al., 2020; Bodó, 2019; Longo, 2011). The question of open (or closed) access is only the beginning of the discussion on better uses of data. New forms of data governance, forms that were more socially responsible and sustainable, would entail a discussion on systems of participation and representation within a democracy, not only on the possibilities of access, but also on what kind of uses are enabled and what innovation is both needed and necessary (Beaulieu & Leonelli, 2022; Bodó, 2019).

2.2. Alternative models of data governance

The discussion about alternative data governance models aims to find new ways to best articulate the different rights related to data to maximise its potential value, considering equally social and economic goods (Graef, 2020). Scholars have proposed several alternatives to resolve the problem of data concentration and promote more sustainable use of data (Madison, 2020; Micheli et al., 2020). Sustainable use of data can be understood as the type of use that preserves and supports communities and their values (Taylor & Purtova, 2019). Another argument in this debate is that regimes incentivizing data sharing could be a more economically efficient answer and would boost innovation, even if it is not without costs (Graef & Prüfer, 2021; Pistor, 2020). Some authors call for public data infrastructures or data spaces that foster innovation and sustain better political decisions (Zygmuntowski, Laura Zoboli,& Nemitz, 2021). Most alternative models proposed depend on creating new legal arrangements and policy adaptations.

From this perspective, Micheli et al. propose a taxonomy of alternative data governance models that policymakers could foster to counter-balance the problems created by the platform economy (Micheli et al., 2020). Among the different models put forward, the public data trust model entails establishing a new kind of institution that manages, preserves privacy and improves the public value of data (Kitchin, 2022). A public data trust would be an independent institution empowered to receive, store, and share data from public bodies (but also eventually from private actors); it would establish the conditions for sharing such data with private actors and other interested public bodies. The institution would also oversee data reuse, preserving control over data and preventing abuse. The key idea underlying such a proposition is that data should provide insights and support for policymaking and promote innovation and social good (Micheli et al., 2020). Under the model described by Micheli et al., there is no detailed specification of how the governance of such structures would be organized, but the authors emphasize that innovative forms of participation and democratic decision-making should be deployed (Micheli et al., 2020). A few pilot projects of public data trust have been tried and are discussed in the literature. One such is the city
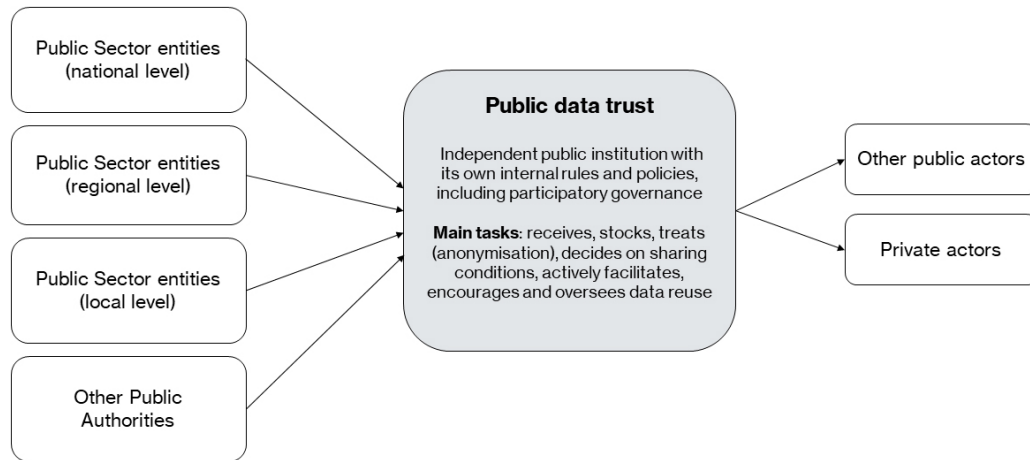
Fig. 2. Public data trust.

of Barcelona and others are to be found in the United Kingdom in different areas, for example urban data, the illegal wildlife trade, and food waste (Hardinges, Wells, Blandford, Tennison, & Scott, 2019; Morozov & Bria, 2018).

The public data trust model echoes the idea of exploring data through a regime of commons, producing a more plural and collective data governance approach (Madison, 2020; Taylor & Purtova, 2019). However, the public data trust contrasts with the idea of purely open commons, which is at the foundation of the OGD movement and the roots of PSD liberalisation (Bodó, 2019). Indeed, while in the OGD regime, PSD is widely and freely available, without any control or oversight of its use, a public data trust regime would entail democratic control over data access and reuse. The management of PSD would be closer to a closed commons regime. In this sense, Madison explains that data could be considered a collective resource, and their extraction and management should involve the participation of all the interested and affected actors. This closed commons approach relies on the idea that the resources are not completely open to being exploited, but they should be subject to collective governance through the enactment of collective institutional rules addressing coordination and social (and ethical) dilemmas. Such a governance system could co-exist with private and market laws (such as proprietary rights) (Madison, 2020). Figure 2 illustrates data flow among public sector entities and public or public actors through public data trust.

Public data trust and data governance through the commons approach raises the question of what kind of data-driven innovation is desirable and who decides about the desirability of such innovation? The idea is to add public and democratic reflection on the purpose and the form of data sharing and reuse. This approach, it is argued, would also bring public scrutiny to the innovation process by making the link between the input (data) and the output (final product or service) more transparent and shedding light on the overall impact of innovation. Adding this reflection layer could provide a space for discussion and negotiation on the possible harms and benefits of technological advances and how those harms and benefits are distributed throughout society (Pfotenhauer, Laurent, Papageorgiou, Stilgoe, & Jack, 2022). It also provides more transparency and public awareness of the trade-offs of innovation, making more explicit how, for example, gains in efficiency could also lead to an increase in surveillance or aggravate discrimination against certain vulnerable groups (Taylor, 2017). Moreover, a democratic debate about the desirability of innovation would contest the rhetoric of the inevitability of technological advances (Taylor, 2017).

The discussion as to how to make the most of data's potential while aligning data-driven innovation to society's needs and priorities envisages, inter alia, engagement with the broader question of how to increase the potential of scientific knowledge and technical innovation so as to address societal problems and challenges (such as climate change, demographic-ageing crisis, etc.) (Beaulieu & Leonelli, 2022; Mazzucato, 2016). In this respect, the literature on RRI proposes a possible path to reduce the gap between society's needs and innovation. In the next section, we explore the RRI to bring into the discussion on data governance the lessons that can be learned from the RRI practice and literature.

2.3. Responsible research and innovation

The possibility of bridging the needs of society and technological advances is a challenge and has received much attention from scholars and policymakers (Jasanoff, 2016). Several mechanisms have been proposed and in some cases tried (with different levels of success). Some of these aim to enable society to have an active voice in the discussion of the technological future. STS provides the shared theoretical base of these approaches. It considers that technological advances are not the deterministic result of technical features. Instead, it assumes that political and social choices have a direct influence over the technological trajectory (Stilgoe, Owen, & Macnaghten, 2013). This premise gains renewed relevance, considering that technological innovation, once deployed, becomes structural and shapes social sets by enabling and constraining action (Orlikowski, 1992). In this sense, considering that the trajectory of technological advances is shaped by social structures, discussions on the governance of innovation are also discussions about power and decisions over our collective future (Pfotenhauer et al., 2022).

In this context, the RRI framework emerged as a possible approach to tackle the challenge of promoting socially desirable innovation (Jirotka, Grimpe, Stahl, Eden, & Hartswood, 2017). The framework emerged simultaneously as a policy discourse and a research field (Owen & Pansera, 2019). As a policy discourse, it was adopted by the EU to make science and innovation more responsive to contemporary society's challenges (Owen & Pansera, 2019). The framework aims to bring together on one side, science and innovation and, on the other side, the expectations and values of society (Owen, Pansera, Macnaghten, & Randles, 2021). Several researchers have collaborated on the elaboration and development of the concept. Scholars emphasize that the process of innovation creates the future and entails changes that may affect society at large; hence, it should be the subject of democratic reflection and decision (Owen & Pansera, 2019). The most current definition of RRI in the literature is inspired by a policy document published by the European Commission (EC) in 2011 (Burget, Bardone, & Pedaste, 2017). The definition goes as follows:

> Responsible Research and Innovation is a transparent, interactive process by which societal actors and innovators become mutually responsive to each other with a view on the (ethical) acceptability, sustainability and societal desirability of the innovation process and its marketable products. (European Commission, 2011, p. 9).

The RRI approach engages with a broad realm of stakeholders in the discussion concerning the direction and speed of research and innovation (Randles, Tancoigne, & Joly, 2022). It proposes a democratic and inclusive decision-making process concerning new technologies. At the same time, it is a valuable tool for managing risks and uncertainties concerning technological development (Jirotka et al., 2017).

The RRI's foundations are the notions of reflexibility, anticipation, inclusion, and responsiveness (Burget et al., 2017; Stilgoe et al., 2013). Some authors include the dimensions of care and sustainability as complementary concepts to RRI (Burget et al., 2017). Reflexibility aims to inculcate a culture of self-reflection, awareness and moral responsibility in researchers and innovators (Stilgoe et al., 2013). RRI

seeks to respond to the idea that innovation is creating the future. For this reason, within the RRI process, there is an exercise of imaging possible futures and anticipating the risks and benefits of a given invention. By implementing an inclusive process, RRI fosters a collective debate and promotes the engagement of a larger public in finding innovative and technical solutions (Stilgoe et al., 2013). Responsiveness is linked to the transparency and accessibility of the innovation process (Burget et al., 2017). It conveys the idea that the innovation process should be open-ended and adjust the innovation accordingly to the inputs of stakeholders (Stilgoe et al., 2013). The emerging dimensions of RRI refer to the resource-efficiency of innovations (sustainability) and the idea that the decisions about our future should be collectively constructed, making citizens equally responsible for them (Burget et al., 2017). The RRI framework does not aim to build a utopian world where the risks of innovation are all foreseen, and only harmless innovation is allowed to be developed. Instead, the RRI is anchored in concrete practices and inclusive dialogue that seeks to foresee and democratically evaluate all the possible risks of a new technology (Burget et al., 2017). Authors have argued that this practice is valuable as a compliance exercise and that it adds value to innovation, improving acceptability and increasing public trust in technology (Ribeiro, Smith, & Millar, 2017).

The RRI approach has faced criticism (Owen & Pansera, 2019). The most important criticism relates to the challenge of transforming RRI from discourse into practice (Ribeiro et al., 2017). More substantially, some critics accuse RRI of slowing innovation and unduly interfering in the autonomy of science. Critics have also warned about the limits of predicting the impact of innovation due to the impossibility of foreseeing all potential uses of a given technology (Owen et al., 2021). However, RRI is an evolving concept and policy, and it has demonstrated real potential for fostering reflection and responsiveness to research and innovation (Ribeiro et al., 2017). Indeed, the RRI has become a horizontal theme in the work of the EC, which has driven the development of the RRI framework by integrating it into its research programs, encouraging its integration into the academic curricula and promoting training and public forums for discussions on the topic (Owen & Pansera, 2019). Moreover, the EC has developed tools to help governments and businesses implement the approach (https://rri-tools.eu/). In what concerns the actions of governments, the EC recommends that the RRI approach should be deployed in the design of policies covering research and innovation.[1] It also recommends incorporating RRI into the structure of research funding, proposing that governments condition the allocation of funds on projects that embrace RRI.[2] All in all, RRI could be a valuable framework for complementing the data governance propositions designed by the DGA. To better understand how this could work, the following section presents an in-depth analysis of the DGA policy context with a focus on the governance of PSD.

2.4. The EU's alternative approach to data

To address the challenges of data concentration and restrained innovation, the EU has adopted a data strategy and developed its regulatory framework to provide the legal conditions for harnessing the potential of data. The DGA is one of the fundamental pieces of this new regulatory framework. This section presents the policy context and then outlines the main features of the DGA, emphasizing the novelties in PSD governance and highlighting national implementation issues. It is built on an in-depth analysis of European Commission preparatory documents, including the consultation process. The

---

[1]https://rri-tools.eu/how-to-stk-pm-incorporate-rri-in-policy/funding-institutions last time consulted on 30 September 2022.
[2]See: https://rri-tools.eu/how-to-stk-pm-incorporate-the-rri-principles-in-a-funding-call last time consulted on 30 September 2022.
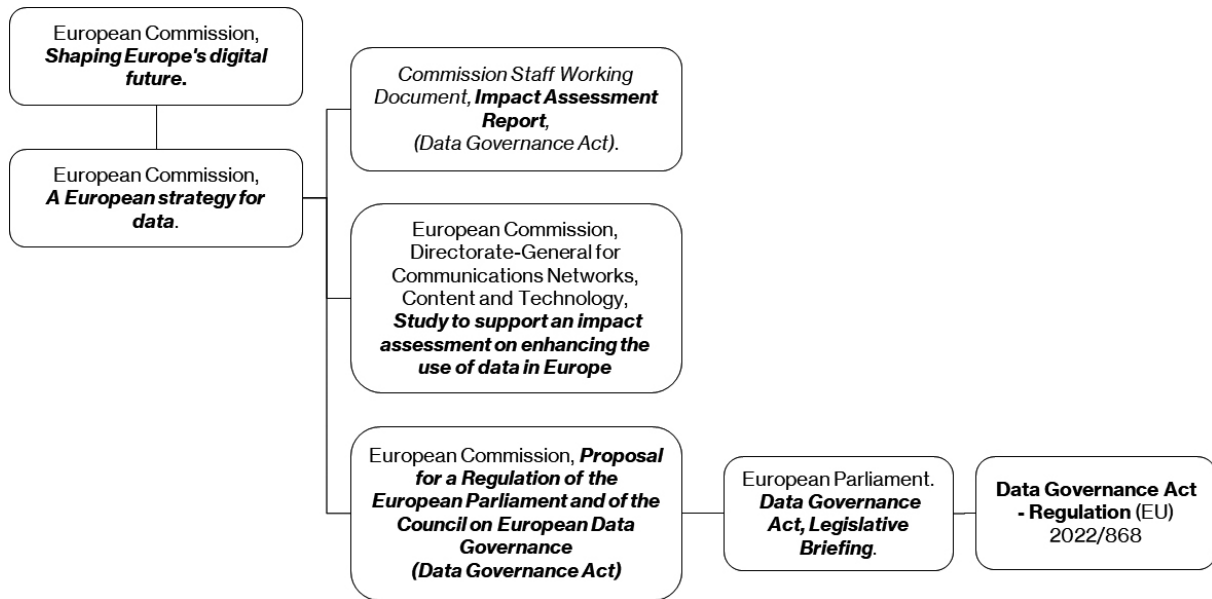
Fig. 3. European data governance policy chain.

selection of these preparatory documents traces back to the policy context of the DGA and considers the relevant documents that help in understanding the policy's rationale and goals. Figure 3 illustrates the DGA's policy path with the most relevant policy documents to its adoption. The analysis followed an inductive strategy that highlights the constitutive function of discourse (Howarth & Griggs, 2015; Yanow, 2015). The legal analysis of the DGA final text used a classic legal hermeneutic analysis that considers the policy context in which the legal act was adopted, following a law-in-context approach (Selznick, 2003).

2.4.1. The European data strategy

In 2020, the EC released its data strategy with the goal of increasing both the supply and demand for data by improving the technical and legal conditions of data sharing and use in the EU market (European Commission, 2020b). This goal responds to a broad political agenda that aims to build a comprehensive approach to the data economy (European Commission, 2020d). Such an approach was expressly presented as an alternative to the platform model (European Commission, 2020a). The platform model of data economy is described as one in which giant tech companies are present in the entire data value chain and have established a quasi-monopoly over data (European Commission, 2020a; Micheli et al.,2020). The vision encapsulated in the European Data Strategy aims to empower society to exploit data's potential for creating economic and social goods and to promote innovation that responds to individual and collective needs (European Commission, 2020b). It is, therefore, part of a regulatory effort that aims to change the current data governance model in order to level the playing field for a broader and more diverse network of actors.

The barriers identified by the EC as holding back data sharing can be summarised in three categories: (i) lack of trust in data sharing (data holders do not trust that re-users will use data in accordance with contractual rules or general guidance) (ii) lack of structures and processes for data sharing, including legal uncertainty and (iii) technical obstacles (European Commission, 2020a). In the EC's view, data should flow within the EU and across sectors. By putting forward this plea for openness, the EC aims

to put in place the conditions for "society to get the most out of innovation and competition and ensure that everyone benefits from the digital dividend." (European Commission, 2020b, p. 1). Moreover, the improvement in data sharing conditions combined with the increased availability of data is deemed essential for "tackling societal, climate and environment-related challenges" (European Commission, 2020b, p. 3). For that purpose, the EC intends to develop a clear and fair legal framework enabling data sharing and reuse. Finally, the EC aims to build a legal and technical environment that combines a wide flow of data with high levels of privacy, security, safety and ethical principles (European Commission et al., 2022, p. 12).

The EC's DGA proposal partially formalised the goal of deploying structural conditions for improving data sharing and reuse. The preparatory documents reveal that EC has significant expectations concerning this new regulatory framework. It is expected to achieve economic growth and improve social well-being. The impact assessment of the DGA's proposal foresaw the economic potential but did not present any evidence of it. For example, it mentions the difficulty in accessing the real potential of PSD and refers to an OECD estimate that enhanced access to sensitive PSD would bring economic and social benefits equivalent to something between 0.1% to 1.5% of GDP (European Commission et al., 2022, p. 19). Besides its economic ambition, the European policy on data emphasizes the concern about using PSD accordingly in the public interest and to promote the common good (European Commission, 2020b). In the same way, when summarising the consultation process, the impact assessment prepared by the EC notes that the "decision to allow reuse should be based on the public interest (which needs to be defined) and use-case specific risk assessment." (European Commission, 2020a, p. 71).

Although the EC's discourse reveals an effort to establish the link between data reuse and the promotion of social-oriented technologies, it remains silent concerning the problem of what constitutes the public interest and common good or how to establish it. In this sense, what constitutes the common good depends on political discussion. Within the EU policy context, the RRI approach, adopted in the realm of public funding of research, provides a suitable tool for overcoming this gap in the regulation. The following section shows how the final text of the DGA addresses the issue of bridging the gap between innovations and societal needs. It emphasizes the absence of mechanisms that ensure a technological development more closely aligned with the common good and public interest.

2.5. *Public sector data in the data governance act*

The DGA was adopted by the European Council and the European Parliament in March 2022. It is built on four pillars: (i) a regime for the reuse of certain categories of protected data held by public sector bodies (subject to rights of others); (ii) a certification framework for data intermediaries (conceived as a neutral agent that provides data sharing services by connecting data holders to data re-users); (iii) measures facilitating data altruism; (iv) the institution of an expert group, the European Data Innovation Board, to facilitate and coordinate European Member States' (EMS) efforts to support data-driven innovation, to lower transaction costs and to prevent sectoral fragmentation.

Concerning the governance of PSD, the DGA aims to complement the Open Data Directive (European Union, 2019) by establishing a data-sharing regime for PSD that is not yet covered by 2019's Directive.

The DGA creates a mechanism for reusing specific categories of protected PSD. These kinds of data held by the public sector are subject to others' rights, such as personal data protection, intellectual property rights and commercial confidentiality. Because of the incidence of other kinds of rights, these data sets are excluded from open data regimes. In European policy, this particular category of PSD is also referred to as 'sensitive data' (European Commission et al., 2022). Based on the assumption that

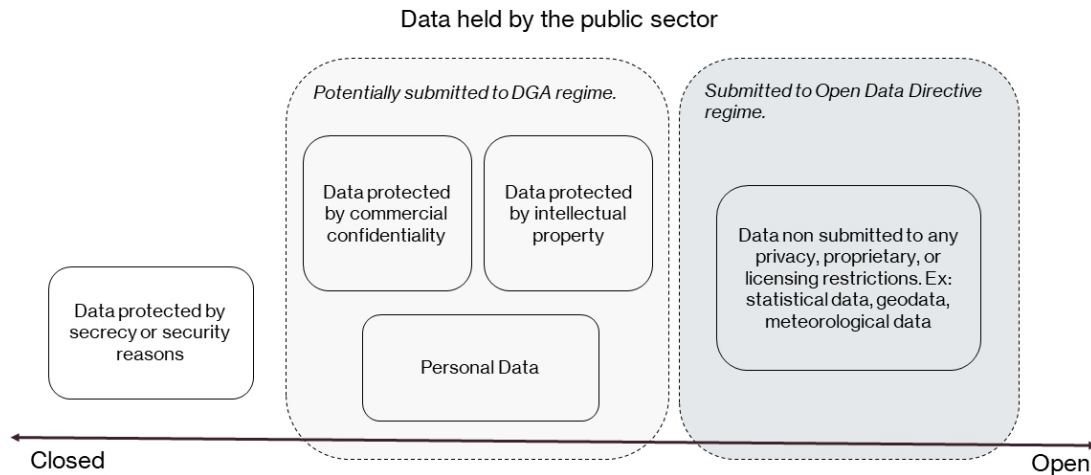Data held by the public sector

Fig. 4. PSD in the openness spectrum.

these data hold important economic and social potential, the idea is to make them accessible, despite the additional work this entails for public sector bodies.

The DGA establishes harmonized rules under which the EMS may allow and enable the reuse of public sector sensitive data. These rules set out some substantial principles. First, EMS are required to be transparent concerning the conditions under which the reuse could be granted (Article 5(1)). The DGA also prohibits – with exceptions – exclusive agreements and restrictions of access for third parties (Article 4). Moreover, the sharing conditions should be non-discriminatory, transparent, proportionate and objectively justified (Article 5 (2)). EMS should designate competent bodies to assist public sector bodies in the activities of granting access to data (Article 7 (1)). Such institutional infrastructure should be provided with the legal, financial and human resources necessary to undergo these tasks (Article 7 (3)). EMS have to put in place a single information point that will offer easy information (providing searchable asset lists and available data sources) (Article 8(1)). This single information point will receive and transmit data reuse requests to the competent bodies (Article 8(2)). The DGA also contains provisions concerning remote access to data and data processing in secured environments (secured physical locations or remote access) (Article 5(3 (b and c))). In these specific cases (processing in secured environments), the public sector body should reserve the right to exercise control through the assessment of processes, means and results of data processing to restrict any kind of use that results in prejudice to the rights and interests of third parties (Article 5(4)). There are also specific restrictions concerning the transmission of data to third countries (Article 5(9 and 10). Finally, the DGA allows public sector bodies to charge for allowing the reuse of data. The fees shall nevertheless be transparent, non-discriminatory, proportionate, objectively justified and shall not restrict competition (Article 6(1 and 2)). Such fees should be derived from the costs involved in the procedure to enable access and reuse (such as costs with reproduction, anonymisation, and maintenance of the secure processing environment, among others) (Article 6(5)). Special fee conditions should be provided for scientific research purposes, small and medium enterprises and start-ups, civil society and educational establishments (Article 6(4)).

However, it is important to note that governments have no obligation to provide data. EMS should "be able to decide whether data is made accessible for reuse, also in terms of the purposes and scope of such access." (Recital 11). Within the DGA, the EMS should establish the specific conditions under which PSD is shared. At the same time, the conditions "should be limited to what is necessary to preserve

the rights and interests of third parties in the data and the integrity of the information technology and communication systems of the public sector bodies" (Recital 15).

Overall, the DGA gives the EMS some leeway to implement a public policy covering the sharing of public sector sensitive data that is aligned with the general policy goal of implementing a model of data governance that is truly an alternative to platform models and that prioritizes the common good.

3. Discussion: The role of the EMS in achieving the policy goal

The governance of PSD remains in the hands of EMS, and governments will have to adopt and adapt to the DGA's new rules. This will be a challenging task. Governments will have to adapt public infrastructures to execute tasks such as managing requests, taking decisions concerning data reuse and preparing data sets for reuse (e.g., anonymisation procedures, ensuring the protection of intellectual rights, etc.). Additionally, governments will need to implement secured environments (remote or on-site) and be able to supervise data processing in such environments. Acquiring human and technical competencies in this domain will certainly require considerable effort. Nonetheless, the more critical challenge will be achieving the EU policy goal of building a data-sharing model that works for the common good (European Commission, 2020b). Building on the literature presented in Section 2, we now discuss how the DGA not only imposes some demanding tasks on the EMS, but also opens opportunities to experiment with new models of PSD governance.

The regulation does not establish mechanisms to access or evaluate how the final use of data contributes to the common good or how well the innovation resulting from data reuse is aligned with the public interest. A possible path to ensure that PSD liberalisation is responsive to society's needs and values would be to integrate the institutional design of public data trust embedded in the approach proposed by RRI.

From an institutional level, creating a public data trust appears to be an adequate solution to manage PSD liberation. The DGA establishes that states should create a single point of contact, but does not specify how such an authority should be organized nor the type of legal form it should have. Creating a single point of contact raises the idea of a central authority or institution managing the data supply and demand. In this sense, a public data trust would be an intermediary between public sector bodies and potential re-users of PSD providing data management and ensuring security and compliance with privacy and data protection rules (Delacroix & Lawrence, 2019). A public data trust should be the custodian of PSD enforcing a governance policy that ensures data reuse aligns with social goals and needs (Zygmuntowski et al., 2021). Following the typology proposed by Micheli et al., a public data trust would ensure the collective and democratic governance of PSD (Micheli et al., 2020). Civil society's active and meaningful involvement is important to build trust and to define what the common good is. While there exist several possibilities and frameworks for participatory governance, the integration of the RRI approach has the advantage of providing a framework already tested for the governance of public resources, as is the case with public research funding. This institutional design of public data trust embedded in the RRI approach would integrate the RRI principles into the internal work of the public data trust, applying RRI tools to its internal policies and practices. The combination of public data trust institutional features and the RRI approach would ensure independent and democratic data management from an institutional point of view.

Besides an institutional design that ensures the democratic and participative decision-making process concerning PSD reuse, another substantial tool inspired by the European Research and Innovation policies could also be deployed. In addition to applying the RRI to its own internal work, public data trusts could
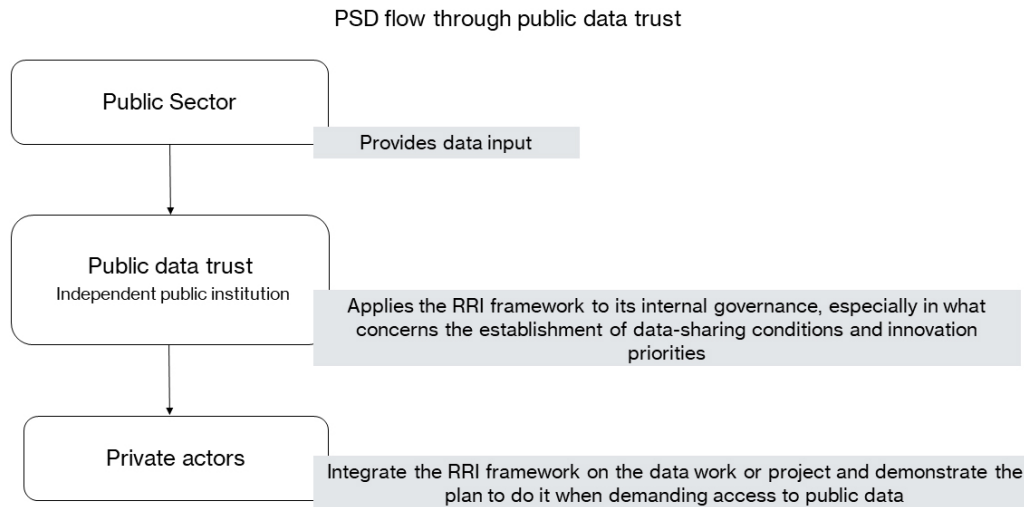
PSD flow through public data trust

Fig. 5. Public data trust and the RRI institutional design.

impose, as a requirement for PSD reuse, the adoption of the RRI approach by innovators and potential data re-users. By imposing the RRI approach as a requirement in the reuse conditions, governments would encourage developers to implement a reflexive and responsive posture. It would foster discussion between innovators and society concerning what innovations are needed and desired. It would also enable a more considered discussion on technology's harms and benefits and the respective distribution in society, bringing more transparency and accountability to innovation processes. The oversight and follow-up of these projects should be an institutional responsibility of the public data trust, as illustrated in Fig. 5.

A well-designed RRI approach, one that enables participation both in the policy discussion and in the design, use and impact of technology, would provide entry points for the public to discuss innovation policy and technological outputs. It would offer the opportunity for the co-construction of technological futures. Indeed, the RRI approach proposes more than simply balancing the harms and benefits of a given technology. It puts forward a process of co-creating technology that is deeply infused with public values. In this sense, the democratic discussion enabled by RRI, questions what kind of technology is truly desirable and could help to develop technology that effectively contributes to human flourishment by fostering autonomy and helping to tackle social challenges.

4. Conclusion

The current data governance model, also referred to as the platform model, is far from ideal. It entails the concentration of data in segregated silos, intensifying power misbalances in society. The concentration of technological means in the hands of a few actors hampers innovation and inhibits democratic discussion on the collective technological future. At the same time, action must be taken to guarantee that innovation is embedded in public values and stewarded for the common good. In the EU, the recently enacted DGA aims to tackle this challenge and envisages the governance of PSD as part of the solution to the existing problems of the platform economy.

This article explores how the DGA impacts the governance of PSD. In this context, the DGA provides some general principles based on the extension of the OGD model, and entrusts EMS with designing the specificities of PSD sharing. The literature on OGD suggests that the simple liberalisation of PSD

is unlikely to achieve the DGA's policy goal, but further measures can be taken. Indeed, the evidence has shown that the mere fostering of the market is insufficient to promote innovation aligned with social values. More than that, the OGD's limited positive results indicate that simple data liberalisation will not suffice to produce technological advances that enhance the common good. In this context, we suggest that the institutional model of public data trust combined with the RRI approach provides a viable path for EMS. The institutional design of public data trust and the RRI framework would enable a more democratic discussion on innovation policy and at least some level of co-construction of technological futures. EMS should be aware that the decision to enable access to PSD is also a question of what kind of innovation is desirable. What are the costs and benefits involved in constructing these innovations? How are these costs and benefits distributed throughout society? These questions are not easy to answer, and different actors may respond differently. However, these are important questions and governments need to provide the framework for discussing them. Implementing public data trust rooted in the RRI approach would enable this discussion and be a step towards restoring a democratic balance to the current concentration of technological power.

This is much more the beginning of the path than the final answer. Experience of RRI to date shows that the approach is a work-in-progress framework, with much to be experimented and improved. This will demand investments both from scholars and practitioners. Furthermore, the implementation of the DGA will certainly revive the research in the field of PSD, which, so far, has been focused on OGD. This research could focus on how PSD liberalisation could improve democratic control over new technologies. From an epistemological point of view, the investigation of the implications of using PSD as the input of innovation and decision-making is also an important field of further research. The intersection between the DGA and other European legislation in data and digital law (such as the Data Act, the Digital Services Act, Digital Market Act, and the Artificial Intelligence Act) also opens an important field of research for scholars working on law, regulation and governance. Concerning the DGA implementation, it is important to note that in the following years, the EU will proceed with the evaluation of the GDPR (2023) and the Open Data Directive (2024). Both evaluations could provide important insights on how to tackle challenges to national implementation, besides delivering crucial information on the effectiveness of the policy. When planning and implementing the DGA, policymakers should also learn from these evaluations.

Acknowledgments

References

Abraham, R., Schneider, J., & vom Brocke, J. (2019). Data governance: Aconceptual framework, structured review, and research agenda. *International Journal of Information Management*, *49*, 424-438. doi: 10.1016/j.ijinfomgt.2019.07.008.

Bates, J. (2014). The strategic importance of information policy for thecontemporary neoliberal state: The case of Open Government Data in theUnited Kingdom. *Government Information Quarterly*, *31*(3), 388-395. doi: 10.1016/j.giq.2014.02.009.

Beaulieu, A., & Leonelli, S. (2022). *Data and Society: A Critical Introduction*. Los Angeles, CA: SAGE PublicationsSage CA.

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

Birch, K., Chiappetta, M., & Artyushina, A. (2020). The problem ofinnovation in technoscientific capitalism: data rentiership and the policyimplications of turning personal digital data into a private asset. *Policy Studies*, *41*(5), 468-487. doi: 10.1080/01442872.2020.1748264.

Bodó, B. (2019). Was the open knowledge commons idea a curse indisguise? – Towards sovereign institutions of knowledge. doi: 10.2139/ssrn.3502119.

Burget, M., Bardone, E., & Pedaste, M. (2017). Definitions and Conceptual Dimensions of Responsible Research and Innovation: A Literature Review. *Sci Eng Ethics*, *23*(1), 1-19. doi: 10.1007/s11948-016-9782-1.

Delacroix, S., & Lawrence, N. D. (2019). Bottom-up data Trusts: Disturbing the 'one size fits all' approach to data governance. *International Data Privacy Law*, *9*(4), 236-252. doi: 10.1093/idpl/ipz014.

Ducuing, C. (2020). Beyond the data flow paradigm: governing data requires to look beyond data. *Technology and Regulation*, *12*(1), 57-64. doi: 10.26116/techreg.2020.006.

European Commission. Research, D.-G. f., & Innovation. (2011). *Towards responsible research and innovation in the information and communication technologies and security technologies fields*: Publications Office.

European Commission. (2020a). *Commission Staff Working Document, Impact Assessment Report, Accompanying the document Proposal for a Regulation of the European Parliament and the Council on European data governance (Data Governance Act)*. Brussels, Belgium.

European Commission. (2020b). *A European strategy for data* (COM(2020) 66 final). Brussels, Belgium.

European Commission. (2020c). *Proposal for a Regulation of the European Parliament and of the Council on European data governance (Data Governance Act)*. (COM(2020) 767 final). Brussels, Belgium.

European Commission. (2020d). *Shaping Europe's digital future*. (COM(2020) 67 final). Brussels, Belgium.

European Commission, Directorate-General for Communications Networks, Technology C., Peijl, S., Denny, E., Koring, E., . . . Walker, A. (2022). *Study to support an impact assessment on enhancing the use of data in Europe*: Publications Office of the European Union. Brussels, Belgium.

European Parliament. (2021). *Data Governance Act, Legislative Briefing*. Brussels, Belgium.

European Union, Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the re-use of public sector information (recast), 2019/1024 C.F.R. (2019). Brussels, Belgium.

European Union, Regulation (EU) 2022/868 of the European Parliament and of the Council of 30 May 2022 on European data governance and amending Regulation (EU) 2018/1724 (Data Governance Act) (Text with EEA relevance) OJ L 152, (2022). Brussels, Belgium.

Graef, I. (2020). Paving the Way Forward for Data Governance: a Story ofChecks and Balances: Editorial. Technology and Regulation, 24-28. doi: 10.26116/techreg.2020.003.

Graef, I., & Prüfer, J. (2021). Governance of data sharing: A law & economics proposal. *Research Policy*, *50*(9). doi: 10.1016/j.respol.2021.104330.

Hardinges, J., Wells, P., Blandford, A., Tennison, J., & Scott, A. (2019). Data trusts: lessons from three pilots. *London: Open Data Institute*.

Howarth, D., & Griggs, S. (2015). Poststructuralist discourse theory and critical policy studies: interests, identities and policy change. In F. Fischer (Ed.), Handbook of Critical Policy Studies: Edward Elgar Publishing.

Jaakola, A., Kekkonen, H., Lahti, T., & Manninen, A. (2015). Open data, open cities: Experiences from the helsinki metropolitan area. case helsinki region infoshare www. hri. fi. *Statistical Journal of the IAOS*, *31*(1), 117-122. doi: 10.3233/SJI-150873.

Janssen, K. (2012). Open Government Data and the Right to Information: Opportunities and Obstacles. *Journal of Community Informatics*, *8*(2), 1-9. Retrieved from http://ci-journal.net/index.php/ciej/article/view/952/954.

Jasanoff, S. (2016). *The ethics of invention: technology and the human future*: WW Norton & Company.

Jetzek, T. (2016). Managing complexity across multiple dimensions of liquid open data: The case of the Danish Basic Data Program. *Government Information Quarterly*, *33*(1), 89-104. doi: 10.1016/j.giq.2015.11.003.

Jetzek, T., Avital, M., & Bjørn-Andersen, N. (2013). *Generating Value from Open Government Data*. Paper presented at the 34th International Conference on Information Systems (ICIS), Milan.

Jirotka, M., Grimpe, B., Stahl, B., Eden, G., & Hartswood, M. (2017). Responsible research and innovation in the digital age. *Communications of the ACM*, *60*(5), 62-68. doi: 10.1145/3064940.

Khayyat, M., & Bannister, F. (2015). Open data licensing: More than meets the eye. *Information Polity*, *20*, 231-252. doi: 10.3233/IP-150357.

Kitchin, R. (2022). *The Data Revolution: A critical analysis of Big Data, Open Data, Data Infrastructures*, London: SAGE.

Longo, J. (2011). OpenData: Digital-Era Governance Thoroughbred or New Public Management Trojan Horse? *Public Policy & Governance Review*, *2*(2), 38-51. doi: 10.1080/00480169.2011.633308.

Madison, M. (2020). Tools for Data Governance. *Technology and Regulation*, *2020*, 29-43. doi: 10.26116/techreg.2020.004.

Mazzucato, M. (2016). From market fixing to market-creating: a new framework for innovation policy. *Industry and Innovation*, *23*(2), 140-156. doi: 10.1080/13662716.2016.1146124.

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), 205395172094808-205395172094808. doi: 10.1177/2053951720948087.

Morozov, E., & Bria, F. (2018). Rethinking the smart city. In (Vol. 2): Rosa Luxemburg Foundation New York, NY, USA.

Murillo, M. J. (2015). Evaluating the role of online data availability: The case of economic and institutional transparency in sixteen Latin American nations. *International Political Science Review*, *36*(1), 42-59. doi: 10.1177/0192512114541163.

Ohm, P., & Frankle, J. (2018). Desirable Inefficiency. *Florida Law Review*, *70*(4), 777-777. doi: https://scholarship.law. ufl.edu/flr/vol70/iss4/2.

Orlikowski, W. J. (1992). The Duality of Technology: Rethinking the Concept of Technology in Organizations. *Organization Science*, *3*(3), 398-427. Retrieved from http://www.jstor.org/stable/2635280.

Owen, R., & Pansera, M. (2019). Responsible Innovation and Responsible Research and Innovation. In D. Simon, S. Kuhlmann, J. Stamm, & W. Canzler (Eds.), *Handbook on Science and Public Policy* (pp. 26-48). Chelteham: Edward Elgar Publishing Limited.

Owen, R., Pansera, M., Macnaghten, P., & Randles, S. (2021). Organisational institutionalisation of responsible innovation. *Research Policy*, *50*(1-1). doi: 10.1016/j.respol.2020.104132.

Pagallo, U. (2022). The Politics of Data in EU Law: Will It Succeed? *Digital Society*, *1*(3), 20-20. doi: 10.1007/s44206-022-00021-3.

Pfotenhauer, S., Laurent, B., Papageorgiou, K., Stilgoe, & Jack. (2022). The politics of scaling. *Social Studies of Science*, *52*(1), 3-34. doi: 10.1177/03063127211048945.

Pistor, K. (2020). Rule by Data: The End of Markets? *Law and Contemporary Problems*, *83*, 101-124.

Prainsack, B. (2020). The political economy of digital data: introduction to the special issue. *Policy Studies*, *41*(5), 439-446. doi: 10.1080/01442872.2020.1723519.

Randles, S., Tancoigne, E., & Joly, P.-B. (2022). Two tribes or more? The historical emergence of discourse coalitions of responsible research and innovation (rri) and Responsible Research and Innovation (RRI). *Journal of Responsible Innovation*, 1-27. doi: 10.1080/23299460.2022.2061306.

Reggi, L., Dawes, S. S., & Gil-Garcia, J. R. (2022). The effects of open government data on the inclusiveness of governance networks: Identifying management strategies and success factors. *Information Polity*, *27*, 473-490. doi: 10.3233/IP-220004.

Ribeiro, B. E., Smith, R. D. J., & Millar, K. (2017). A Mobilising Concept? Unpacking Academic Representations of Responsible Research and Innovation. *Science and Engineering Ethics23*(1), 81-103. doi: 10.1007/s11948-016-9761-6.

Ruijer, E., & Meijer, A. (2020). Open Government Data as an Innovation Process: Lessons from a Living Lab Experiment. *Public Performance & Management Review*, *43*(3), 613-635. doi: 10.1080/15309576.2019.1568884.

Safarov, I., Meijer, A., & Grimmelikhuijsen, S. (2017). Utilization of open government data: A systematic literature review of types, conditions, effects and users. *Information Polity*, *22*, 1-24. doi: 10.3233/IP-160012.

Schwoerer, K. (2022). Whose open data is it anyway? An exploratory study of open government data relevance and implications for democratic inclusion. *Information Polity*, *27*, 491-515. doi: 10.3233/IP-220008.

Selznick, P. (2003). 'Law in Context' Revisited. *Journal of Law and Society*, *30*(2), 177-186. doi: 10.1111/1467-6478.00252.

Smith, G., & Sandberg, J. (2018). Barriers to innovating with open government data: Exploring experiences across service phases and user types. *Information Polity*, *23*, 249-265. doi: 10.3233/IP-170045.

Stilgoe, J. (2020). Who's driving innovation: New Technologies and the Collaborative State. *Palgrave Macmillan*.

Stilgoe, J., Owen, R., & Macnaghten, P. (2013). Developing a framework for responsible innovation. *Research Policy*, *42*(9), 1568-1580. doi: 10.1016/j.respol.2013.05.008.

Susha, I., Grönlund, Å., & Janssen, M. (2015). Driving factors of service innovation using open government data: An exploratory study of entrepreneurs in two countries. *Information Polity*, *20*(1), 19-34. doi: 10.3233/IP-150353.

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

Taylor, L. (2021). Public Actors Without Public Values: Legitimacy, Domination and the Regulation of the Technology Sector. *Philosophy & Technology*, *34*, 897-922. doi: 10.1007/s13347-020-00441-4.

Taylor, L., Mukiri-Smith, H., Petročnik, T., Savolainen, L., & Martin, A. (2022). (Re)making data markets: an exploration of the regulatory challenges. *Law, Innovation and Technology*, *14*(2), 355-394. doi: 10.1080/17579961.2022.2113671.

Taylor, L., & Purtova, N. (2019). What is responsible and sustainable data science? *Big Data & Society*, *6*(2-2). doi: 10.1177/2053951719858114.

Valli Buttow, C., & Weerts, S. (2022a). Open Government Data: the OECD's Swiss Army Knife in the transformation of government [Manuscript submitted for publication]. *Policy & Internet*, *14*(1), 219-234. doi: 10.1002/poi3.275.

Valli Buttow, C., & Weerts, S. (2022b). Public sector information in the European Union policy: The misbalance between economy and individuals. *Big Data & Society*, *9*(2-2). doi: 10.1177/20539517221124587.

Verdegem, P. (2022). Dismantling AI capitalism: the commons as an alternative to the power concentration of Big Tech. *AI & SOCIETY*. doi: 10.1007/s00146-022-01437-8.

Viljoen, S. (2021). A relational theory of data governance. *Yale LJ*, *131*, 573.

Viljoen, S., Goldenfein, J., & McGuigan, L. (2021). Design choices: Mechanism design and platform capitalism. *Big Data & Society*, *8*(2-2). doi: 10.1177/20539517211034312.

Yanow, D. (2015). Making sense of policy practices: interpretation and meaning. In F. Fischer (Ed.), *Handbook of Critical Policy Studies*. Cheltenham, UK: Edward Elgar Publishing.

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