

Data governance for public transparency

Agustí Cerrillo-Martínez; Anahí Casadesús-de-Mingo

How to cite this article:

Cerrillo-Martínez, Agustí; Casadesús-de-Mingo, Anahí (2021). "Data governance for public transparency". *Profesional de la información*, v. 30, n. 4, e300402.
<https://doi.org/10.3145/epi.2021.jul.02>

Article received on 28-01-2021
Final acceptance: 25-05-2021



Agustí Cerrillo-Martínez ✉
<https://orcid.org/0000-0002-1577-4194>

Universitat Oberta de Catalunya
Avda.Tibidabo, 39
08035 Barcelona, Spain
acerrillo@uoc.edu



Anahí Casadesús-de-Mingo
<https://orcid.org/0000-0002-2682-6025>

Universitat Autònoma de Barcelona
Escuela Superior de Archivística y Gestión
de Documentos (Esaged)
Edifici Blanc, Campus de la UAB
08193 Cerdanyola del Vallès
(Barcelona), Spain
Anahi.Casadesus@uab.cat

Abstract

Public transparency is becoming increasingly complex due to the volume of data generated by government, the plurality of uses given to public data, their dispersal over different organizations, bodies and units and the diversity of mechanisms through which they are channelled. All this requires government agencies not only to improve data management but also to adopt procedures and structures that facilitate decision-making regarding data's use and quality. In this context, this study defines data governance as the set of principles, values and standards that guide interaction in decision-making among stakeholders who create, manage and use data. This study uses the analysis of three data governance cases to identify the defining characteristics of data governance (data governance's design, the institutional position on data governance in the organizational structure, the stakeholders involved in data governance, the interaction channels provided and the functions attributed to them). Based on these elements, three models of data governance promoted by government agencies are observed. In the light of the data governance models analysed, the final reflection identifies how data governance can contribute to improve public transparency.

Keywords

Data; Data governance; Transparency; Public administration; Local administration; Big data; Public sector; Public policy; Accountability; Data management; Data re-use; Open government; Open data; Legal frameworks; Policies; Compliance; Case studies.

Funding

This article has been prepared under the project "Personalization of Public Services, Bias and Artificial Intelligence". State Programme for R&D&I Focusing on the Challenges Facing Society (PID2020-115774RB-I00), Spanish Ministry of Science, Innovation and Universities.

1. Introduction

Transparency seeks to help citizens effectively find out what is happening in government (Janssen *et al.*, 2017, p. 3; Meijer, 2013). It is based on knowledge of data generated by the activity of government agencies, opening it up to the scrutiny of the citizens to whom government is accountable for the decisions it makes, how it makes them and how it manages public funds. Thus, citizens can participate actively in public decision-making and monitor government action (Hood; Heald, 2006).

Transparency is based on data hold by government agencies.

Traditionally, data management has had a great impact on the improvement of transparency in government agencies (Cerrillo-Martínez; Casadesús-de-Mingo, 2018, p. 8; Janssen *et al.*, 2017, p. 4).

As data resides in documents, databases, and other media, transparency needs data management to identify, categorize, preserve, and retrieve data when required for business purposes or regulatory needs (Bhansali, 2014, p. 2). Data management also considers definition of formats and, when necessary, migration and conversion, to permit interoperability among different government agencies and reuse by citizens and companies, as well as guaranteeing long-term digital

preservation. Indeed, not all data (or records) should be kept forever, especially in open data or transparency portals. In the long run, such perpetual preservation of datasets will only produce noise in searches and information retrieval and will occupy an enormous amount of space with obsolete and useless information.

Data management has also been useful to guarantee quality of information, i.e. information that is accurate, trustworthy, complete, usable, reusable, up-to-date and reliable (Koltay, 2016, p. 306). Quality of data is necessary, if they are to truly meet public transparency duties by opening transparency channels and to increase the value of data provided to the public. This is achieved through appropriate data and records management systems that ensure these qualities are maintained over time. This requires guarantees that the information held by government agencies meets certain standards and is managed in such a way as to ensure and maintain the quality of the information over time. It is also important that information and records management does not focus exclusively on internal control, oriented mainly to the operation of the public administration and its employees, without also seeking to ensure public information can be provided to citizens and guaranteeing its authenticity, integrity, usability, reliability and reuse. Furthermore, as the value of information as the basis for government decisions increases, records management systems need to ensure that information is both the result of public decisions and the basis for them.

However, since transparency is becoming more complex, data management tools and frameworks need to be strengthened with new mechanisms to improve the design and implementation of data policies in government agencies.¹ In the past, transparency was only channelled through access to public records and information held by public organizations. In present times, transparency is the result of complex processes that involve different sources, people, data or channels. There are different reasons.

Firstly, transparency is now more complex because the volume of information generated by government agencies has grown exponentially, and with it the amount of information that has to be provided to the public (Biancone; Secinaro; Brescia, 2018, p. 494). Furthermore, it also has to be taken into account that big data has arrived to many government agencies (Giest; Ng, 2018; Hong *et al.*, 2019).

Secondly, transparency increases in complexity due to the fact that the possible uses of data generated and disseminated by government agencies have also increased significantly (Pozen, 2019, p. 1). And this is not only in government agencies, through data analysis and the development of predictive policies, but also in companies and among the general public through reuse (Magalhaes; Roseira, 2020, p. 2; Zuiderwijk *et al.*, 2014).

Thirdly, transparency is more complex because public data are spread among different bodies and units (Meijer, 2013, p. 429). At present, public data are often spread over different organizations or, if centralized, held in different units. It is also not unusual for different units in an organization to be unaware of whether certain data exist and which unit holds them. Data in different units might have also been stored in different formats, impeding interoperability or even, occasionally, automatic processing. Nor is it strange for a unit to fail to use its data or share them with others in the same organization. Organizations may even hold data that could be used to provide new public services, improve decision-making, and offer new services on the market, but they fail to do so because no one has seriously considered what to do with such data. It may also be the case that the metadata, as well as the underlying data quality for datasets, is known to be deficient. For instance, many open datasets have duplicate, inconsistent, and missing data generally lack easily accessible schema descriptions (Sadiq; Indulska, 2017, p. 1) to help recover the information. Such gaps directly affect the quality of data and how they can be used by both government and citizens.

Fourthly, transparency is more complex because it is mediated by different actors, not only within the government but also by civil society and companies (Cerrillo-Martínez, 2012, p. 790; Meijer, 2009).

Fifthly, transparency is more complex because it is channelled through a variety of mechanisms that enable citizens to access public data. The traditional channel was individual access to information through formal requests to government bodies. However, since the rise of the Internet, transparency has been channelled through government dissemination of large amounts of information on websites and transparency portals and also on open data portals. In the last decade, thanks to the development of 2.0 technologies, transparency has also been channelled through the reuse of public information by citizens, thus also helping to make information available to all (Lassinantti, 2019, p. 4).

Finally, transparency is more complex because there are at stake different rights and interests over public data. In particular, government agencies hold large amounts of personal data, which requires that transparency is managed according to the *General Data Protection Regulation*.

For all such reasons, to improve transparency it is necessary to enhance mechanisms and procedures to be able to satisfy this complexity. Government agencies need to adopt policies, procedures and structures to facilitate decision-making about data uses and data quality taking into account all interests at stake. In addition, government agencies need to design mechanisms that facilitate interaction and interoperability between those stakeholders that create and maintain data, those that define access and reuse, and those that use data, to guarantee these principles are observed and procedures are respected. This is achieved through adequate data governance, which goes beyond mere management to adopt a holistic approach that permits information in public authorities to be governed correctly. This involves greater

coordination as well as an interdisciplinary approach that brings together different professionals (public managers, records managers, technologists and jurists, among others), which helps manage the risks facing the organization by using the information and data it manages as well as optimize the value of its information (Sloan, 2014, p. 2).

These are the reasons why governing data is vital (Alhassan; Sammon; Daly, 2019, p. 1) and it also is inevitable (Abella-García, 2020). The main aim of these mechanisms is to define and implement data policies that increase the quality and the value of data within an organization or its environment. The poorer the quality of data and datasets, the worse the use and reuse. In addition, the aim should also be to design and implement solid data policies according to the views and interests of all data stakeholders.

All these involves more than data management systems and technologies; it affects the model of governance itself (Benfeldt-Nielsen, 2017). As it has been said,

“the changing nature of data management and data use, the evolving technological context, and the shifting meaning of core governance concepts, place today’s systems for data governance under stress” (*British Academy; Royal Society*, 2017, p. 6).

As we will see in the next sections, the plurality and diversity of data held by government agencies and the complexity of their uses require new mechanisms to channel the interaction between different organizations, units and people producing, owning, managing, using and re-using data and thus increase the value of data generated by government activity. We will analyse different models of data governance to arise this diversity and to see which model could be more effective in order to achieve a better governance and a better transparency in public organisms.

In this context, the remainder of this paper is organized as follows. Section 2 shows the methodology used in this paper. Section 3 examines the notion of data governance as a mechanism for promoting, strengthening and improving data creation, management, use and reuse to improve the quality of public decision-making, provide greater transparency and increase the value of public data. Next, section 4 analyses three different data governance cases that have been implemented by different public agencies in Spain. The discussion section identifies the defining elements of data governance and sets out three data governance models. Finally, the paper ends with some concluding remarks relating data governance and transparency.

2. Methodology

The aim of this study is to identify the key characteristics of data governance in public administration and how they can improve transparency.

In order to achieve this objective, we have analysed different data governance cases. In particular, we have explored the three data governance cases that had been formalized in different rules published in official journals. These cases were selected because they were the only available cases for study in Spain when the article was written.²

The case study is based on the formal analysis of the rules that regulate data governance. Nevertheless, the analysis has some limitations. Unfortunately, the short time that has elapsed since the approval of the regulations and the lack of sufficient data about their enforcement have prevented us from carrying out a study on its actual performance. However, from our perspective, this is not an obstacle to getting enough information to abstract the defining characteristics of data governance, identifying different data governance models and to assess its impact on transparency.

Despite the fact that the study was carried out from the analysis of three models put into operation in Spanish public administrations, in our opinion the conclusions reached may be useful beyond Spanish borders. Indeed, the characteristics that are identified can be used in the design of data governance models in other countries to improve public transparency.

3. Conceptualizing data governance: An overview of the literature

The rise of governance in public administration is an indication of the profound changes in Western societies in recent decades, characterized by their complexity, dynamism, fragmentation and diversity. A single vision is no longer possible and a plural approach is needed, leading to the emergence of new models of government of the public sphere (Kooiman, 2003, p. 11).

From a general point of view, governance is identified as a set of principles, values and standards that guide interaction between actors involved in developing a given public policy (Koliba *et al.*, 2019).

These reflections have been transferred to public data policies, where data governance has arisen as a management model intended to transform government of data (Benfeldt-Nielsen; Persson; Madsen, 2018). Data governance aims to respond to the plurality, complexity and dispersal of data held by government agencies, required for both public decision-making and creating public value and needed for transparency and accountability. Thus, data governance seeks to provide a strategic vision for data management in an organization, with the aim of making maximum use of the data it holds.

In recent years, various authors have attempted to define data governance. As shown below, the concepts can be grouped into those that see data governance as identifying rights and responsibilities related to data, and those that link the concept to the design of data management procedures.

Authors in the former group include **Benfeldt-Nielsen, Persson and Madsen**, who define data governance as

“the allocation of decision-making rights and related responsibilities to encourage desirable behaviour in the use of data” (**Benfeldt-Nielsen, 2017; Benfeldt-Nielsen; Persson; Madsen, 2018**).

Al-Ruithe, Benkhelifa, & Hameed (2019) or **Alhassan, Sammon, & Daly (2019)** define data governance along the same lines. Also Abraham, Schneider, & Vom-Brocke consider that

“data governance refers to the exercise of authority and control over the management of data” (**Abraham; Schneider; Vom-Brocke, 2019**).

As well, Janssen *et al.* define data governance as

“organizations and their personnel defining, applying and monitoring the patterns of rules and authorities for directing the proper functioning of, and ensuring the accountability for, the entire life-cycle of data and algorithms within and across organizations” (**Janssen et al., 2020**).

Finally, Thompson, Ravindran, & Nicosia state that

“data governance initiative provide the checks and balances” (**Thompson; Ravindran; Nicosia, 2015**)

Among authors in the latter group, Pierce, Dismute, & Yonke consider data governance as

“the collective set of decision-making processes for the use and value-maximization of an organization’s data assets” (**Pierce; Dismute; Yonke, 2008**).

Similarly, Paskaleva argues that

“data governance should be an inclusive and iterative process of data development by the stakeholders for the shared benefits of both people and city” (**Paskaleva et al., 2017**).

Furthermore, Korhonen defines data governance as

“an organizational approach to data and information management that formalizes a set of policies and procedures to encompass the full data life cycle, from acquisition to use and to disposal” (**Korhonen et al., 2013**).

This definition lacks a reference to data creation, the first step in data life-cycle management. Finally, Dumthanasarn considers that

“data governance is a set of processes that assures that data, as assets, are formally and properly managed” (**Dumthanasarn; Thammaboosadee; Ayuthaya, 2018**).

All these definitions share a lowest common denominator that revolves around three elements: the definition of roles and responsibilities of participants in data governance; the definition of principles and policies; and accountability.³

These elements are what differentiates data governance from data management, understood as

“the development, execution, and supervision of plans, policies, programs, and practices that control, protect, deliver, and enhance the value of data and information assets” (*Data Management Association, DAMA*).

Data governance goes further, focusing above all on

“exercising authority, control and shared decisions” and on “high-level planning and control of data management” (**Al-Ruithe; Benkhelifa; Hameed, 2019**).

Data governance refers to decision-making, while data management refers to implementation. Whatever the case, they are complementary rather than replaceable.

However, we believe the above definitions should be complemented by a fourth element, one that recognizes nowadays complexity in data management which summarizes in the recognition of interaction between interdependent actors to whom are attributed rights of decision and responsibilities related to the data in an organization and outside it. This incorporates the complexity facing organizations in light of the plurality of data at their disposal spread over different units.

That is the reason why in the following pages we consider data governance as the system of principles, values and standards that guide interaction in decision-making on data among people who create, manage and preserve the data, who define access and use, and who access, use and reuse data, all of them providing different visions and resources (relating to strategy, archiving, technology, standards, economics, and more).

Despite the potential of data governance, literature on the topic is still scarce and generally excessively theoretical (**Benfeldt-Nielsen; Persson; Madsen, 2018; Al-Ruithe; Benkhelifa; Hameed, 2019; Paskaleva et al., 2017; Alhassan; Sammon; Daly, 2016**). However, it has now been identified as an emerging area (**Al-Badi; Tarhini; Khan, 2018**). Also, experiences are still not very usual. Indeed, at present few government agencies have implemented data governance.

A number of reasons for this have been identified. Thus, **Benfeldt-Nielsen, Persson, & Madsen (2018)**, following **Begg & Caira (2011)**, asks why government agencies find it difficult to explore and exploit their datasets through data governan-

ce and discuss the frequency with which organizations fail to recognize the inherent value of their data or see data as something different from technology. They also observe that some organizations, despite recognizing the value of data governance, may consider that the effort of designing and implementing it outweighs the potential benefits. Nor should we ignore that up until very recently data governance was informal (Al-Ruithe; Benkhelifa; Hameed, 2019).

Table 1. Challenges of data governance

Topic	Challenge
Value of data and joint vision	Short-term view of data use
	Difficulty understanding the value of data initiatives
	Lack of joint vision on existing data
Data practice and collaboration	Autonomy between different departments
	Lack of trust in data in the social sphere
	Lack of inter-organization collaboration
Data capacities and policies	Variable levels of data maturity among different departments
	Lack of top-down support for data initiatives
	Lack of political focus on data use

Source: Benfeldt-Nielsen, Persson, and Madsen (2018)

This is also due to the lack of consideration given to using previously implemented controls, processes and methodologies with a long tradition, such as records and information management systems. Such an approach would provide the methodological foundations for data governance, ensuring integrity, reliability, access and usability for as long as necessary and, therefore, the possibility of reuse. It should not be forgotten that a key aspect of data life cycle relates to decisions over what should be kept, for how long and what can be deleted.

4. Approaching to data governance models through three case studies

As we have seen, data governance is not understood and managed the same by different authors, but there is a diversity and plurality. Taking this into account, there are also different data governance models within public administration. Traditionally, data governance in public administrations has been founded on hierarchy and a command-and-control basis. However, in the last years new models of data governance are based in the design of networks with the participation of different stakeholders.

In the following pages we will present three cases of data governance set up by different Spanish public bodies.⁴

In particular, we have explored the three data governance cases that had been formalized in different rules published in official journals when the article was written. They were the only available cases for study in Spain when the article was written. However, as will be see, they are representative cases of three different data governance models. In the first case, a single person is responsible for data governance of a specific government body. The second case involves a collegiate body that has been created as a hub for data governance. Finally, the third case is an agency specifically created to data governance.

From the analysis of these cases, we will be able to abstract the defining characteristics of data governance and identify three data governance models in the following section.

4.1. Case study 1: Information and Knowledge Management Authority (IKMA)

In Spain, the *Ministry of Defense* has set up the *Information and Knowledge Management Authority (Autoridad de Gestión de la Información y el Conocimiento)* to manage and oversee compliance with the principles and objectives of the ministry's *ICT strategy*. This function is attributed to the Secretary of State for Defense. The *IKMA* must coordinate the information management structure at corporate level, supervising compliance with applicable policies and developing the ministerial strategy.⁵

IKMA designates the Chief data officer.⁶ The Chief data officer is the official responsible for guaranteeing quality, reliability and no duplicity of data. He has also to guarantee data precision, consistency, credibility and timeliness, and completeness.⁷

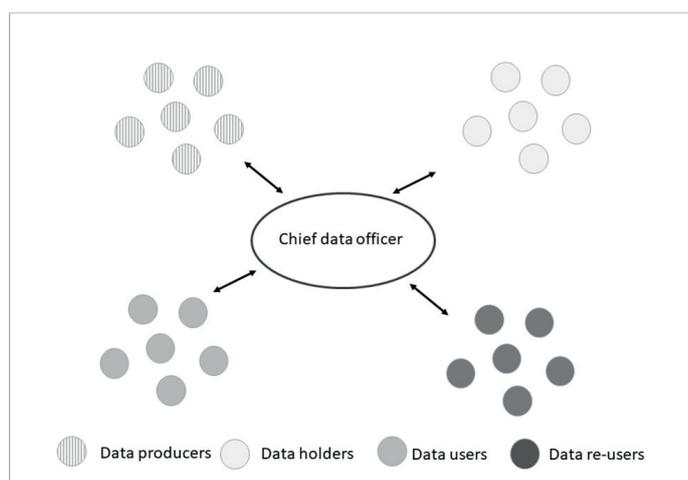


Figure 1. Centralized model: Chief data officer

The *IKMA* has several functions including the classification and qualification of information, the implementation of the organization, governance and control structures of the information management system; the identification of the managers of the information, the identification of the users of the information, and the evaluation of the effectiveness and efficiency of data management policies and processes.

In the public sphere, few government agencies have appointed chief data officers (Tito, 2017; Wiseman, 2017). With the growth of big data, many organizations have created the position of chief data officer, responsible for guaranteeing data quality and security (Deloitte, 2016; Wiseman, 2018). The chief data officer should be responsible for finding ways of using current data in the organization increasing the value of data and guaranteeing data privacy and security.

The chief data officer interacts with data producers, data holders, data managers and data users and reusers but they do not directly interact among them. They only contact through informal channels (see Figure 1).

4.2. Case study 2: Data Governance Committee

In Spain, the Spanish Government (specifically the part of the executive branch of government known as the *Administración General del Estado, AGE*) has passed the *Plan of digitalization of public administrations 2021-2025. Digital administration and digital public services strategy* which establishes the objectives, principles and actions to develop the e-government of *AGE* and provides the basis for drawing up ministerial action plans for digital transformation.⁸

The aims of the *Plan of digitalization of public administrations 2021-2025* include to turn Spanish public administration into a modern data guided one where data is used efficiently to design public policies according to Spanish social, economic and territorial reality.

To achieve this objective, the *Strategy* defines several measures to improve data management in an efficient and transparent way and to develop public services based on data.⁹ In particular, the *Strategy* pursues to have secure and quality data; to encourage innovation and experimentation using data in Public Administration; and promote the creation of accessible data repositories.¹⁰

Design of ICT governance in *AGE* has been transformed in the last decade. Among other objectives, the model aims to ensure that *AGE* is ready to

“collect, generate and process large volumes of digital information on its operations, processes and results, which could usefully be made available to citizens to promote transparency and to companies and social agents to foster reuse of public sector information”.¹¹

The ICT governance model outlined by *Royal decree 806/2014* seeks to centralize areas of authority and resources for implementing them in a single collegiate body that includes all the ICT units of *AGE* and its public authorities. This body interacts with other areas of government through new collegiate bodies set up as a streamlined information channels, sharing needs and opportunities for the rational and efficient reuse of ICT resources (see Figure 2).¹²

The *ICT Strategic Committee* is the collegiate body responsible for defining and supervising application of the *ICT strategy* for *AGE* and its public authorities (hereinafter referred to as the *ICT strategy*).¹³

The *ICT Strategic Committee* has a number of functions, including:¹⁴

- Establishing strategic lines in ICT, in accordance with government policy, to boost e-government in *AGE* and its public authorities.
- Approving the proposed *ICT strategy* of *AGE* and its public authorities to be discussed in the *Council of Ministers* by the heads of the ministries of *Finance, Territorial Policy, and the Presidency*.

The *ICT Strategic Committee* is organized into a *Plenary* and an *Executive Committee*. The *Plenary Committee* consists of the state secretariats for *Public Administrations, Telecommunications and the Information Society, and Social Security*, together with the under-secretaries or the secretary of a higher body in the different ministerial departments, and the ICT director-general. The *Plenary Committee* is chaired by the minister of *Finance and Public Administration*, with the ICT director-general as secretary. Importantly, the chair may invite representatives from other public or private institutions to participate without voting rights.¹⁵ For its part, the *Executive Committee* of the *ICT Strategic Committee* is the instrument that ensures efficient and effective implementation of the *ICT strategy* in *AGE* and its public authorities.¹⁶

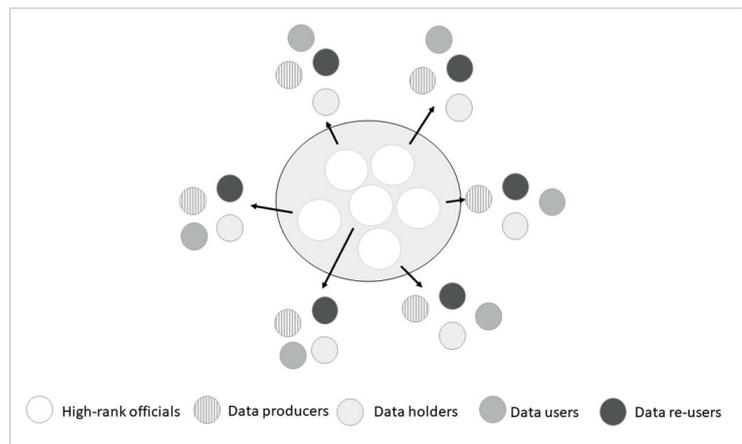


Figure 2. Coordination model: collective body

Along with these agencies, the *Ministerial Digital Administration Committees (Comisiones Ministeriales de Administración Digital, CMADs)* are departmental collegiate bodies responsible for fostering and coordinating e-government internally in each department.¹⁷

The *CMADs* are chaired by the under-secretary and are made up of representatives, with a minimum rank of deputy director-general, from the participant functional areas and bodies and the ministerial ICT unit managers. Their functions include:

- Acting as the liaison body to ensure coordination with the criteria and policies defined by the latter.
- Promoting, executing and supervising, within each department, compliance with directives and monitoring the action guidelines included in the *ICT strategy for AGE* based on the proposals of the *ICT Strategic Committee*.
- Drawing up the departmental action plan for digital transformation, developing the criteria established by the *ICT Directorate*, based on the *ICT strategy*.
- Promoting digitalization of departmental services and procedures with the aim of harmonizing and simplifying them, improving their quality and facilitating their use, and the services provided to citizens and companies, optimizing the use of available ICT resources.

4.3. Case study 3: Data Office

Barcelona City Council designed its data governance model in 2018.¹⁸ According to this model,

“governance of municipal data must permit the supervision and coordination of all actions related to data, avoiding duplication, standardizing processes, and generating synergies that permit the availability of more, higher quality, and more accessible data”.¹⁹

Barcelona City Council data governance consists of decision-making areas and cross-departmental coordination. Specifically, these areas are the *Data Executive Committee*, the *Data Coordination Board*, the sector data managers and the *Municipal Data Office*.²⁰ The *Data Office* integrates several areas and departments with responsibilities for corporate data (see Figure 3). It is part of municipal government, it reports to the *Technology and Digital Innovation Committee*, and is headed by a chief data officer. Thus, it follows the model first promoted by cities such as New York, Chicago and Boston, and subsequently Paris and London.

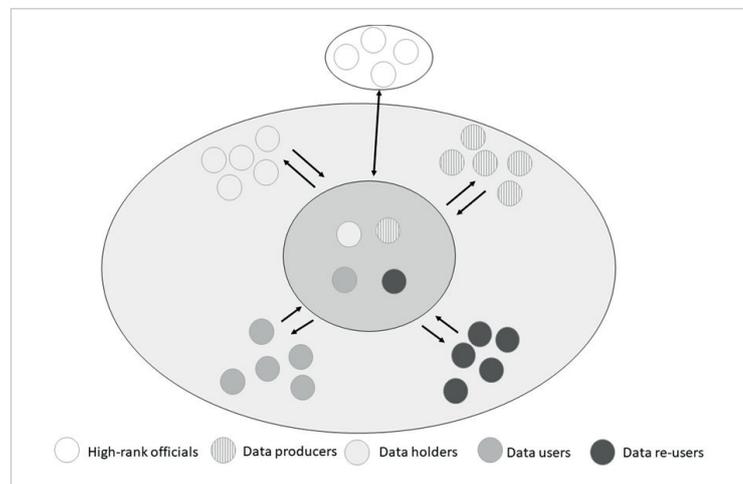


Figure 3. Decentralized model: Data office

The *Municipal Data Office* is a unified data governance and management point, which aims to streamline municipal structures and improve the efficiency and transparency of public policy.

<https://ajuntament.barcelona.cat/digital/es/transformacion-digital/city-data-commons/oficina-municipal-de-datos>

It consists of a multidisciplinary team of 40 people, as well as an information and records manager, and experts in efficiency, transparency, accountability, and life-cycle management.

<https://www.slideshare.net/francescabria/barcelona-municipal-data-office>

Its mission is to exercise governance of data owned or held by *Barcelona City Council* from three axes: digital transformation, digital innovation and digital empowerment.

The objectives of the *Municipal Data Office* are:

- Defining and coordinating a municipal data governance model, based on a defined corporate data management strategy.
- Guaranteeing compliance with the standards and regulations expressed in the data strategy.
- Facilitating alignment of technological tools with requirements of use.
- Intensively introducing data science in the *City Council* as a tool to provide greater knowledge of citizens and their needs, management of the city, internal management, and risk management.
- Strengthening data sovereignty, dissemination, availability and transparency of all municipal data, and the promotion of open data.
- Exchanging knowledge on policies, standards and good government data practice through active dialogue with citizens and data consumers, other producers, public/private corporations and research centers.
- Promoting care, reuse, improved accessibility, searchability, preservation and enrichment of data.

Among other functions, the *Municipal Data Office* establishes the vision, strategy and governance of initiatives, projects and actions in the council data system, ensures the smooth functioning and management of data systems and guarantees ethical, responsible and lawful use of data.

The *Municipal Data Office* uses public data to provide potentially useful information for meeting the challenges facing the city and providing a service to different municipal units.

5. Discussion

The cases analysed in the previous section show that data governance is important for guaranteeing the quality of data in an organization (Al-Ruithe; Benkhelifa; Hameed, 2018) and maximizing their value. It also fosters use and reuse of data in an organization to improve decision-making and generate value. In this latter instance, data governance is necessary for aligning an organization's data policy with its goals and general objectives (Benfeldt-Nielsen; Persson; Madsen, 2018). However, it is not usual that they have specific functions related to transparency. In fact, only the *Municipal Data Office* of Barcelona holds duties on transparency.

From a general point of view, in the previous section we have seen that there are different data governance models through which decisions about data can be taken in a public administration.

Data governance can be based on hierarchy and a command-and-control model which means that a central actor takes decisions about data and guarantees their implementation (Weber; Otto; Österle, 2009). This is the case of the *Information Management Authority*, the data chief officer of the *Ministry of Defense*.

However, other models of data governance are based in the design of networks which are frameworks where interdependent stakeholders (data producers, data owners, data managers, data users and data reusers) interact through their numerous interests, and the variety of resources they have (regulatory, economic, strategic, cognitive, etc.) are put into force (Kickert; Klijn; Koppenjan, 1997; Peters, 1998). In the cases analysed in the previous section, we have seen that the participation of different stakeholders and the level of interaction among them are different, depending on the design of the data governance model.

According to the information gathered from the cases analysed, the characteristics of data governance are identified below and three data governance models will be defined later.

6. Defining characteristics of data governance models

Different governance models are based in five characteristics that define them.

First of all, the design of data governance. As seen, data governance can be based on a hierarchical design or a collaborative one. On one hand, a hierarchical data governance model uses command and control mechanisms and it is based in an up-down relationship among stakeholders. Decisions are taken unilaterally. On the other hand, collaborative data governance model functions by non-hierarchical coordination based on an exchange of resources among stakeholders. Decisions are based on the agreement and collaboration among stakeholders.

Secondly, the institutional position of data governance in the organizational structure. Data governance models should be given a position within the institution that ensures interaction among stakeholders without undue interference from the organization. The design and, if necessary, implementation of data policies should be the result of the interaction between interdependent stakeholders through the channels designed for this purpose. Along these lines, it has been proposed, for example, that this body should be independent or given a large degree of autonomy from the governing authorities (*British Academy; Royal Society*, 2017, p. 5). Data governance can be linked to the top of an organization but also to the executive or managerial levels.

Thirdly, stakeholders involved in data governance. Inside stakeholders can be the only participants in data governance. But data governance may also involve representatives from other public or private institutions. Data governance models can recognize the participation of all stakeholders, those who create and maintain data, those who define their access and use, and those who use and reuse data, or only the contribution of some of them.

Fourthly, the interaction channels provided. The data governance model should guarantee interaction among all stakeholders. Through interaction, the different stakeholders establish channels for communication, the exchange of information, experiences and other resources. Interaction can be formal or informal. Collegiate bodies are a powerful instrument for channelling the complexity and plurality of stakeholders, interests and resources coexisting within a public policy towards decision-making. Formally established bodies can be complemented by informal mechanisms of interaction among stakeholders. The data governance model should ensure that interaction among stakeholders is conducted under the right conditions. In particular, data governance should be a transparent, and participatory process that guarantees accountability.

Finally, the functions attributed. Different data governance models develop different functions. The functions can be different depending on the level of interdependence between stakeholders. The higher the level of interdependence among stakeholders, the bigger will be the need of data governance to integrate their different interests, views and resources. The data governance model should be able to ensure data quality and increase the value of the data, but also improve public transparency.

7. Data governance models

Combining the elements defined in the previous sections allows us to identify three data governance models: centralized model, coordinated model and decentralized model:

- Centralized model of data governance. This model assigns the coordination of the data decision-making to a single figure. It functions on a command and control basis. Generally, the chief data officer is a technical profile and reports to the highest governing authority in the institution. Although the chief data officer interacts with stakeholders, the model does not facilitate interaction among them.
- Coordinated model of data governance. This model is based on the creation of a collegiate body, which holds regular meetings among representatives from all bodies in the institution that hold, manage and/or use data. Generally, the composition of this collegiate body matches the functions assigned to it. Thus, when its functions relate to the design of data policies, the collegiate body's members are the people in charge of the organization's data policy. When the functions are related to policy implementation these members have a more executive profile. The composition of the collegiate body frequently determines its institutional position.
- Decentralized model of data governance. This model consists of creating a body responsible for data policy's design and implementation that also interacts, not only with all bodies and with units in the institution that hold, manage or use information, but also with the external bodies and persons who use the information. As has been stated, "this stewardship body would be expected to conduct inclusive dialogue and expert investigation into novel questions and issues, and to enable new ways to anticipate the future consequences of today's decisions" (*British Academy; Royal Society, 2017*). This body generally reports to the highest governing authority in the organization.

8. Concluding remarks

Data governance has an important role in the definition and implementation of public data policies. According to Abraham,

"the purpose of data governance is to increase the value of data and minimize data-related cost and risk" (**Abraham; Schneider; Vom-Brocke, 2019**).

In this line, in November 2020, the *European Commission* has presented a proposal for a regulation of the *European Parliament* and of the *Council on European data governance (Data Governance Act)*.²¹ The proposal is a measure included in the *European Strategy for Data*. The aim of the proposal is

"to foster the availability of data for use by increasing trust in data intermediaries and by strengthening data-sharing mechanisms across the EU."

As it is stated in the explanatory memorandum of the proposal,

"setting up the governance structures and mechanisms that will create a coordinated approach to using data across sectors and Member States would help stakeholders in the data economy to capitalize on the scale of the single market."

The proposal does not establish a specific governance model.

Either way, it is safe to say that the existence of different data governance models shows that there is not only one approach to the issue, as noted in the statement

"effective data governance strongly resists a one-size-fits-all approach" (*British Academy; Royal Society, 2017, p. 7*).

Despite the potential of data governance in organizations and that it is gaining popularity (**Kooper, Maes, & Lindgreen, 2011**), as we have seen, data governance models don't usually have specifically taken into account transparency when defining its configuration, organization and functioning.

However, as transparency is based on data held by government agencies, it should be considered when defining a data governance model in a public administration.

As we have seen, transparency is now more complex because the volume of information generated by government agencies has grown exponentially, the possible uses of data generated and disseminated by government agencies have also increased, public data are spread among different bodies and units. It is mediated by different actors, not only within the government but also by civil society and private companies. It is channeled through a variety of mechanisms that enable citizens to access public data and there are different rights and interest over public data at risk.

That means that if data governance wants to have an impact on transparency it should integrate this complexity to facilitate interaction between those who create and maintain data, those who define their access and use, and those who use and reuse data, with the aim of guaranteeing data quality, improving public transparency and increasing the value of the data. To do this, a key aspect of data governance is identifying relevant stakeholders (**Al-Badi; Tarhini; Khan, 2018**). As stated above,

"data governance should be an inclusive and iterative process of data development by the stakeholders for the shared benefits of both people and city" (**Paskaleva et al., 2017**).

That is the reason why public administrations should not have only one actor for defining data policies or guaranteeing their implementation, particularly those related to transparency. Indeed, transparency has become so complex that it requires the participation of everyone who produces, uses (both internally and externally) and manages data. There must be an interdependent network in place for interactions to ensure data's quality along with their use, dissemination and reuse.

In this context, this study has identified three data governance models. Each one has a different design and institutional position. Furthermore, the interaction between stakeholders participating in data governance also differs. Attributed functions vary depending on these elements. Their contribution to improve transparency will be different as far as the channel, more or less interests, visions and resources into data decision-making.

9. Notes

0. A previous version of this article was presented at the *6th Global conference on transparency research* held in Rio de Janeiro on the 26th and 27th of June, 2019.

1. This is clear in the *Digital transformation plan for the Spanish General Administration and Public Agencies (ICT strategy 2015-2020)*, which states that

“ICT development has led to the production of a huge amount of information as a result of the activity of the Administration in connection with internal procedures, social media, interaction with citizens, and so on. All this information opens up new perspectives and enables the development of innovative services based on emerging technologies, such as the processing of large volumes of information, data mining, predictive analytics, and others”.

https://administracionelectronica.gob.es/pae_Home/dam/jcr:0d4cfaad-3df4-46a1-8b87-aa3dc602e90b/Plan_de_trans_Estrategia-TIC_ingles.pdf

2. After finishing this paper, the Spanish government created a new data governance body: the *Division of Data Office (Order ETD/803/2020, July 31, establishing the Data Office Division and the Planning and Program Execution Division within the Secretary of State for Digitalization and Artificial Intelligence)*. The *Data Office Division* will be responsible for designing data management strategies and frameworks, creating secure and governable data sharing spaces for companies, citizens, and public administrations, designing governance and management and analysis data policies and developing a competence center on data analytics.

3. However, it should be borne in mind that when some authors define data governance, they are referring to “policies and procedures adopted to manage data in an organization” (Al-Badi; Tarhini; Khan, 2018).

4. Along with the data governance bodies analysed in the text, some government agencies have created transparency offices responsible for designing, coordinating, assessing and monitoring transparency policies developed by governmental agencies. In addition, transparency laws frequently define the bodies responsible for coordinating responsibilities in applying transparency standards in one or more government agencies.

5. Article 9 *Order DEF/1196/2017, of 27 November*.

6. *Sixth Instruction 37/2019, of 9 July, from Secretary of State of Defense, to coordinate information and knowledge management in the Ministry of Defense*.

7. Article 10 *Order DEF/1196/2017, of 27 November*.

8. Accessible in:

https://administracionelectronica.gob.es/pae_Home/pae_Estrategias/Estrategia-TIC/Plan-Digitalizacion-AAPP.html

The *Plan of digitalization of public administrations 2021-2025* is included in the *Recovery, transformation and resilience plan* passed in April 27, 2021.

https://administracionelectronica.gob.es/pae_Home/en/dam/jcr:8dca8417-29b8-4b2f-9fa1-f3ed3e29d29e/30042021-Plan_Recuperacion_Transformacion_Resiliencia.pdf

9. Strategic axis 1.

10. Measure 6.

11. Preamble to *Royal decree 806/2014, of 19 September*, that defines ICT organization and operational instruments in the *Administración General del Estado* and its public authorities. Author's note: author's own translation.

12. The deployment of the *Plan of digitalization of public administrations 2021-2025* will need a new governance model to guarantee efficiency in the control, direction, setting and decision-taking. The *Plan* states that the governance model will have two levels (plan and projects). The model will be regulated in a royal decree that will replace *Royal decree 806/2014*. The new governance model will include two bodies: the *Steering Committee for Administration Digitalization (Comité de Dirección para la Digitalización de la Administración)* and the *Executive Committee for Administration Digitalization (Comité Ejecutivo para la Digitalización de la Administración)*.

13. Article 3 of *Royal decree 806/2014*. Author's note: author's own translation.
14. Article 4 of *Royal decree 806/2014*. Author's note: author's own translation.
15. Article 5 of *Royal decree 806/2014*. Author's note: author's own translation.
16. Article 6 of *Royal decree 806/2014*. Author's note: author's own translation.
17. Article 7 of *Royal decree 806/2014*. Author's note: author's own translation.
18. *Municipal Gazette* of 18 April 2018.
19. Author's note: author's own translation.
20. As well as the areas mentioned in the text, the governance model also includes the *Data Protection Board* and the data protection officer.
21. *COM(2020) 767 final*.

10. References

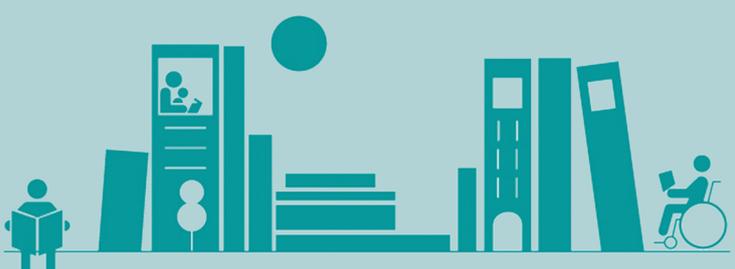
- Abella-García, Alberto** (2020). "Gobernanza inteligente a través de los datos. Modelo de gobernanza y arquitectura de los datos". *Consultor de los ayuntamientos y de los juzgados: Revista técnica especializada en administración local y justicia municipal*, n. Extra 3, pp. 167-194.
<https://dialnet.unirioja.es/servlet/articulo?codigo=7480710>
- Abraham, Rene; Schneider, Johannes; Vom-Brocke, Jan** (2019). "Data governance: A conceptual framework, structured review, and research agenda". *International journal of information management*, v. 49, pp. 424-438.
<https://doi.org/10.1016/j.ijinfomgt.2019.07.008>
- Al-Badi, Ali; Tarhini, Ali; Khan, Asharul-Islam** (2018). "Exploring big data governance frameworks". *Procedia computer science*, v. 141, pp. 271-277.
<https://doi.org/10.1016/j.procs.2018.10.181>
- Alhassan, Ibrahim; Sammon, David; Daly, Mary** (2016). "Data governance activities: an analysis of the literature". *Journal of decision systems*, v. 25, supl. 1, pp. 64-75.
<https://doi.org/10.1080/12460125.2016.1187397>
- Alhassan, Ibrahim; Sammon, David; Daly, Mary** (2019). "Critical success factors for data governance: a theory building approach". *Information systems management*, v. 36, n. 2, pp. 98-110.
<https://doi.org/10.1080/10580530.2019.1589670>
- Al-Ruithe, Majid; Benkhelifa, Elhadj; Hameed, Khawar** (2019). "A systematic literature review of data governance and cloud data governance". *Personal and ubiquitous computing*, n. 23, pp. 839-859.
<https://doi.org/10.1007/s00779-017-1104-3>
- Bhansali, Neera** (2014). *Data governance: Creating value from information assets*. CRC Press. ISBN: 9781439879139
- Begg, Carolyn; Caira, Tom** (2011). "Data governance in practice: The SME quandary reflections on the reality of data governance in the small to medium enterprise (SME) sector". In: *The European conference on information systems management*.
- Benfeldt-Nielsen, Olivia** (2017). "A comprehensive review of data governance literature". *Selected papers of the IRIS*, n. 8.
<https://core.ac.uk/download/pdf/301373908.pdf>
- Benfeldt-Nielsen, Olivia; Persson, John-Stouby; Madsen, Sabine** (2018). "Why governing data is difficult: Findings from Danish local government". In: Elbanna A.; Dwivedi Y.; Bunker D.; Wastell D. (eds.). *Smart working, living and organising. TDIT 2018. IFIP advances in information and communication technology*, v. 533, pp. 15-29.
https://doi.org/10.1007/978-3-030-04315-5_2
- Biancone, Paolo-Pietro; Secinaro, Silvana; Brescia, Valerio** (2018). "The innovation of local public-sector companies: Processing big data for transparency and accountability". *African journal of business management*, v. 12, n. 15, pp. 486-500.
<https://doi.org/10.5897/AJBM2018.8596>
- British Academy; Royal Society* (2017). *Data management and use: Governance in the 21st century*.
<https://royalsociety.org/topics-policy/projects/data-governance>
- Cerrillo-Martínez, Agustí** (2012). "The re-use of public sector information in Europe and its impact on transparency". *European law journal*, v. 18, n. 6, pp. 770-792.
<https://doi.org/10.1111/eulj.12003>

- Cerrillo-Martínez, Agustí; Casadesús-de-Mingo, Anahí** (2018). "El impacto de la gestión documental en la transparencia de las administraciones públicas: la transparencia por diseño". *Gestión y análisis de políticas públicas*, v. 19, pp. 6-16.
<https://doi.org/10.24965/gapp.v0i19.10515>
- Deloitte (2016). *The evolving role of the chief data officer in financial services: From marshal and steward to business strategist*.
<https://www2.deloitte.com/tr/en/pages/financial-services/articles/the-evolving-role-chief-data-officer-financial-services-turkey.html>
- Dumthanasarn, Natcha; Thammaboosadee, Sotarati; Ayuthaya, S. Darakorn-Na** (2018). "Comparative study of open government data law towards data governance legal frameworks". *ITMSOC Transactions on information technology management*, v. 3, n. 1, pp. 25-31.
<http://www.itmsoc.org/index.php/ITM/article/download/65/54>
- Giest, Sarah; Ng, Reuben** (2018). "Big data applications in governance and policy". *Politics and governance*, v. 6, n. 4.
<https://doi.org/10.17645/pag.v6i4.1810>
- Hong, Souman; Kim, Sun-Hyoung; Kim, Youngrok; Park, Jeongin** (2019). "Big data and government: Evidence of the role of big data for smart cities". *Big data & society*, v. 6, n. 1.
<https://doi.org/10.1177/2053951719842543>
- Hood, Christopher; Heald, David** (2006). *Transparency. The key to better governance?* Oxford: Oxford University Press for The British Academy. ISBN: 978 0 197263839
- Janssen, Marijn; Brous, Paul; Estevez, Elsa; Barbosa, Luis S.; Janowski, Tomasz** (2020). "Data governance: Organizing data for trustworthy Artificial Intelligence". *Government information quarterly*, v. 37, n. 3.
<https://doi.org/10.1016/j.giq.2020.101493>
- Janssen, Marijn; Matheus, Ricardo; Longo, Justin; Weerakkody, Vishanth** (2017). "Transparency-by-design as a foundation for open government". *Transforming government: People, process and policy*, v. 11, n. 1, pp. 2-8.
<https://doi.org/10.1108/TG-02-2017-0015>
- Kickert, Walter J. M.; Klijn, Erik-Hans; Koppenjan, Joop F. M.** (1997). "Managing networks in the public sector: findings and reflections". In: Kickert, Walter J. M.; Klijn, Erik-Hans; Koppenjan, Joop F. M. (eds.). *Managing complex networks: strategies for the public sector*. London: Sage.
<https://doi.org/10.4135/9781446217658.n10>
- Koliba, Christopher J.; Meek, Jack W.; Zia, Asim; Mills, Russell W.** (2019). *Governance networks in public administration and public policy* (2nd ed.). Boca Raton, FL: CRC Press. ISBN: 978 1 138286108
- Koltay, Tibor** (2016). "Data governance, data literacy and the management of data quality". *IFLA journal*, v. 42, n. 4, pp. 303-312.
<https://doi.org/10.1177/0340035216672238>
- Kooiman, Jan** (2003). *Governing as governance*. London: Sage. ISBN: 978 0 761940364
- Korhonen, Janne J.; Melleri, Ilkka; Hiekkänen, Kari; Helenius, Mika** (2013). "Designing data governance structure: an organizational perspective". *GSTF journal on computing*, v. 2, n. 4.
<https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.890.7944&rep=rep1&type=pdf>
- Lassinantti, Josefin** (2019). "Re-use of public sector open data - Characterising the phenomena". *International journal of public information systems*, v. 13, n. 1.
<http://www.ijpis.net/index.php/IJPIS/article/view/158>
- Magalhaes, Gustavo; Roseira, Catarina** (2020). "Open government data and the private sector: An empirical view on business models and value creation". *Government information quarterly*, v. 37, n. 3, 101248.
<https://doi.org/10.1016/j.giq.2017.08.004>
- Meijer, Albert** (2009). "Understanding modern transparency". *International review of administrative sciences*, v. 75, n. 2, pp. 255-269.
<https://doi.org/10.1177/0020852309104175>
- Meijer, Albert** (2013). "Understanding the complex dynamics of transparency". *Public administration review*, v. 73, n. 3, pp. 429-439.
<https://doi.org/10.1111/puar.12032>
- Paskaleva, Krassimira; Evans, James; Martin, Christopher; Linjordet, Trond; Yang, Dajuan; Karvonen, Andrew** (2017). "Data governance in the sustainable smart city". *Informatics*, v. 4, n. 41.
<http://www.mdpi.com/2227-9709/4/4/41>

- Peters, B. Guy** (1998). "Policy networks: myth, metaphor and reality". In: Marsh, David (ed.). *Comparing policy networks*. Buckingham: Open University Press. ISBN: 978 0 335196463
- Pierce, Elizabeth; Dismute, Wendell-Scott; Yonke, C. Lwanga** (2008). *The state of information and data governance—understanding how organizations govern their information and data assets*. IQ International publications, Industry report. <https://www.iqint.org/publication2011/pierce-2008-04.shtml>
- Pozen, David E.** (2019). "Seeing transparency more clearly". *Public administration review*, v. 80, n. 2, pp. 326-331. <https://doi.org/10.1111/puar.13137>
- Sadiq, Shazia; Indulska, Marta** (2017). "Open data: Quality over quantity". *International journal of information management*, v. 37, n. 3, pp. 150-154. <https://doi.org/10.1016/j.ijinfomgt.2017.01.003>
- Sloan, Peter** (2014). "The compliance case for information governance". *Richmond journal of law & technology*, v. 20, n. 2. <https://scholarship.richmond.edu/cgi/viewcontent.cgi?article=1388&context=jolt>
- Thompson, Nik; Ravindran, Ravi; Nicosia, Salvatore** (2015). "Government data does not mean data governance: Lessons learned from a public sector application audit". *Government information quarterly*, v. 32, n. 3, pp. 316-322. <https://doi.org/10.1016/j.giq.2015.05.001>
- Tito, Joel** (2017). *Destination unknown: Exploring the impact of artificial intelligence on government*, Working paper. Centre for Public Impact. <https://www.centreforpublicimpact.org/insights/ai-government-working-paper>
- Weber, Kristin; Otto, Boris; Österle, Hubert** (2009). "One size does not fit all. A contingency approach to data governance". *Journal of data and information quality*, v. 1, n. 1. <https://doi.org/10.1145/1515693.1515696>
- Wiseman, Jane M.** (2017). "Lessons from leading CDOs: A framework for better civic analytics". *Ash Center policy briefs series*. <https://ash.harvard.edu/publications/lessons-leading-cdos-framework-better-civic-analytics>
- Wiseman, Jane M.** (2018). *Data-driven government: The role of chief data officers*. IBM Center for the Business of Government. <https://www.innovations.harvard.edu/data-driven-government-role-chief-data-officers>
- Zuiderwijk, Anneke; Helbig, Natalie; Gil-García, J. Ramón; Janssen, Marijn** (2014). "Special issue on innovation through open data. A review of the state-of-the-art and an emerging research agenda. Guest editors' introduction". *Journal of theoretical and applied electronic commerce research*, v. 9, n. 2, pp. 1-2. <https://doi.org/10.4067/S0718-1876201400020000>

Realiza tu proyecto

Por una biblioteca pública social e inclusiva



En nuestra web te ofrecemos unas pautas para facilitar la formulación de tu proyecto
fundacionbibliotecasocial.org



FUNDACIÓN
BIBLIOTECA SOCIAL

info@fundacionbibliotecasocial.org
 @Biblio_Social
 FundaciónBibliotecasSocial



Hazte amigo/amiga
de la Fundación