Designing data governance in Brazil: an institutional analysis

Fernando Filgueiras & Lizandro Lui

Published online: 15 Apr 2022.

Submit your article to this journal ⌕

View related articles ⌕

View Crossmark data ⌕

Routledge
Taylor & Francis Group

RESEARCH ARTICLE

Designing data governance in Brazil: an institutional analysis

Fernando Filgueiras[a,b] and Lizandro Lui[c]

aSchool of Social Sciences, Federal University of Goiás, Goiânia, Brazil; bNational School of Public Administration, Brasília, Brazil; cSchool of Public Policy and Government, Getulio Vargas Foundation, Brasilia, Brazil

ABSTRACT

Data governance is a decision-making process focused on authority building to specify decision rights and accountability that encourage desired behaviors regarding data use, security, integrity, and availability. The emergence of big data technologies to design public policy and deliver public services requires governments to design data policy and governance. This paper analyzes the dynamics of data governance design in the Brazilian Federal Government, based on institutional analysis and policy development. The paper reports a series of interviews with Brazilian Federal Government policymakers to frame the data policy design dynamics. The paper concludes that the policy design dynamics to data governance are path dependent and shape actions situations that reinforces previous institutions. In the case of the Brazilian Federal Government, the institutional framework is ambiguous, creating situations of conflict and ineffectiveness in the design of the data policy.

1. Introduction

This paper aims to analyze the design dynamics that shape governance and data policy in the Brazilian Federal Government. To analyze these design dynamics, we resort to understanding the action situations that delimit policymakers' interactions to shape data policy and data governance institutions. The institutional analysis makes it possible to understand the dilemmas of action that surround policymakers when they are asked to design public policy and how the development of policies over time is essential to understanding the choices of tools and perspectives surrounding policy design.

Big data is a prominent topic these days and there is no consensus regarding its concept, scope, and properties (Ekbia et al. 2014). A broader definition considers that big data is the set of methodologies and technologies applied to collect, store, process, and

share data from various sources and expand the domain of information and deployment of computational technologies (Kitchin 2014). Governments increasingly use big data tools in policy design and implementation. And big data's use allows governments to detect expanded information about various issues in different policy domains (Giest 2017). Data implies the logic of shared resources in a networked place, with a private and public nature. Governments use these data to develop applications to provide services and policies via information control (Filgueiras and Almeida 2021).

Data governance requirements arise from emerging problems related to using digital technologies based on massive volumes of data, shaping tools to reduce sharing risks, and define patterns for stewards and analysts actions (Madison 2020; Abraham, Schneider, and Vom Brocke 2019; Dawes, Cresswell, and Pardo 2009). Big data and its expansion creates problems concerning esthetic, methodological, epistemological, technological, legal, ethical, and economic issues (Ekbia et al. 2014). Working with big data requires data governance to promote policies aimed at the use and qualification of data to promote security, standards, guidelines, and rules. Data governance guides the work of data analysts and stewards to comply with rules, standards, and strategies. To change data stewards' and analysts' behavior, data governance in governments requires policy design dynamics that can contend with problems related to data collection, storage, processing, sharing, and the use, reuse, and disposal of data in their lifecycle. The design of data governance emerges from institutional frameworks that shape the actors' interactions related to massive data for public policy.

Data governance design is still an emerging theme that has mobilized several international organizations, including the Organization for Economic Co-operation and Development, or OECD (2019), and the World Bank (2021). In this paper, we set about analyzing the dynamics of data governance design, focusing on policies aimed at regulatory and procedural instruments for the safe and effective use and sharing of data across different policy domains.

To achieve this goal, we analyze the case of Brazilian Federal Government, exploring the dynamics that emerge and showing the dilemmas and problems encountered within data governance design. We start from the assumption that policy design results from choices of instruments and mixes made by policymakers to achieve policy objectives. Thus, in this article, we analyze how policymakers are designing data policy among the different organizations of the Brazilian Federal Government. This analysis makes it possible to understand policymakers' choice dilemmas regarding the different tool mixes and processes that inform policy design dynamics. This article intends two contributions. First, from a theoretical point of view, to examine how policy design are path-dependent, where design dynamics are constrained by previous choices that define a stream of feedbacks and a pathway. Second, from a practical point of view, we examine the centrality of data policy and governance and how policymakers face institutional dilemmas in digital transformation. As part of this effort, we explore interactions between the Central Data Governance Committee (CCGD) members of the Brazilian Federal Government to design data policy.

2. Background—policy design dynamics and time

To better understand the nuance of design dynamics and conflicts in the practice of data governance design in Brazil, first we provide some background information.

2.1. Policy design dynamics

Policy design reflects a dynamic combination of goals and instruments that develop over time (Howlett 2019). The interests and concerns of policy design researchers and those who work in neo-institutionalism theory sometimes converge because both are committed to explaining policy changes, outcomes, and implications (Van Geet, Lenferink, and Leendertse 2019). Before delving into these dynamics, though, the policy analysis must explain the process of maintaining tailored policy design elements by enduring goal coherence, instrument consistency, and the congruence of goals and instruments (Capano and Howlett 2020). A new branch of design research focuses on finding the appropriate mix of goals and instruments used during the policy design process (Howlett 2019). Policy design's effectiveness is based on the fit between the specter of goals and instruments that are socially and politically accepted.

Policy design orbits around successfully fulfilling policy objectives by calibrating policy goals and policy instruments (Van Geet, Lenferink, and Leendertse 2019). The advancement of policy design theory as a dynamic activity involves the choice of instruments, the composition of policy mixes, calibration, and achievement of results. Policy design theory can benefit from several aspects of the Institutional Analysis and Development Framework (IAD). To analyze policy design, it is essential to understand that it is not a fixed activity but an action situation. Policy design as an action situation mirrors the disputes and conflicts that shape a policy by actors' interactions, considering rules in use, disputed resources, and actions (Ostrom 2005). Policy design is an almost activity involving institutional dynamics that may vary over time according to policy instrument mixes' calibration and policy coherence requirements (Capano and Howlett 2020). In this way, public policies are subject to design and redesign situations when policymakers want to change different aspects related to the instruments, aiming to achieve policy objectives.

Institutional Analysis and Development Framework (IAD) is a method for organizing policy analysis. IAD considers the role of institutions and actors in the construction of outcomes and the policy design (Ostrom 2005; Polski and Ostrom 1999). The IAD framework's starting point is to consider that public policies are institutional arrangements that organize the rules of the game for society to produce public goods and solve collective action dilemmas (Heikkila and Andersson 2018). Based on this premise, the IAD framework seeks to understand the action situations and how these situations guide the actors' interactions in the policy choices. Analytically, the IAD framework is employed to understand actors' motivations to devise or change institutions. IAD does not provide a tool to analyze the design and its results per se, but the complexity of the dynamics that shape and surround institutional choices (Heikkila and Andersson 2018). Institutions govern policy situations, which will define demands, places, and people. Individuals and groups deliberately craft these institutions to make the actors' interactions more predictable to reduce uncertainties and risks (Polski and Ostrom 1999). Thus,

understanding action situations is at the heart of the IAD framework, producing the link between institutions and how they guide individuals and groups' behavior and the outcome achieved by policy (Ostrom 2011).

IAD framework should be understood as a framework for assessing policy design by bringing together actors and institutions in complex situations, which generally involve situations nested in multilevel and polycentric decision systems, including constitutional issues, collective choices, and operational level (Carlisle and Gruby 2019; Carney, Heikkila, and Wood 2019). In this conceptual definition, we have an introductory first statement: policy design involves decision-making by multiple actors who may be engaged in policy in a cooperative or conflicting manner. The design that the policy will perform depends on political factors and institutional dynamics. Thus, the IAD framework is engaged by a policy design perspective (Dunlop, Kamkhaji, and Radaelli 2019; Siddiki 2020), enabling us to understand design dynamics and how they shape policy design.

2.2. Policy development

In a sterilized world, policy designs' deployment would be made from the beginning. Nevertheless, this hardly ever happens in real life. Instead, policy goals and instruments tend to change over time by building on prior choices. As a result, policy designs often develop into suboptimal arrangements, which compromises their effectiveness (Rayner, Howlett, and Wellstead 2017). In policy design, three concepts are fundamental to the analysis: coherence, consistency, and calibration. Policy goals are considered coherent if they connect to the same whole policy intentions and can be chased simultaneously without problems (Kern and Howlett 2009). Instruments are considered consistent if they mutually support and work together to achieve the same goal by creating related incentives and constraints (Van Geet, Lenferink, and Leendertse 2019). Finally, the calibration process ensues from the permanent monitoring of policy coherence and consistency. In this sense, calibration is a process that adjusts goals and instruments over time. Calibration is necessary because the policy changes the social reality, solves the original problems, collects learning about the process, and formulates new objectives.

Policy mixes change over time in response to institutional dynamics requiring instrument calibration and objectives adjustment (Lima, Aguiar, and Lui 2021). Observing these processes, the literature focuses on the mixes through specific characteristics such as coherence, consistency, and congruence, treated as proxies of robustness (Lima, Aguiar, and Lui 2021). Howlett, Mukherjee, and Rayner (2018) argue that decisions, in general, are about repairing or restructuring elements rather than proposing new alternatives. Again, the policy design theory and new institutionalist theory achieve a commonplace that tries to explain the empirical world by policy development. It is apparent that the institutional theory wields significant influence because the arguments attached to the idea of legacies from past decision making are often persistent and hard to modify. Understanding how this issue fits into the design dynamics is essential to advancing the design approach.

There is a tendency for policy design theories to observe their dynamics as moments of policy choices made by singular actors empowered in the policy process. Policy

design analysis tends to be actor-based functionalist, where strategic actors select policy designs to achieve an objective. The analysis of design dynamics can benefit from a policy development perspective, in which actors' choices are not only strategic but constrained by temporal developments that shape present choices. Once public policy changes over time, they present a development process, making the actors' choices path dependent (Pierson 2004).

Our theoretical argument in this article (when looking to explain data governance design in Brazil) is that policy design tends to be path dependent (Pierson 2004; Krasner 1988). Thus, often these initial choices will not represent equilibrium solutions—such as the Nash equilibrium—but suboptimal results or inefficient outcomes that require changes in inertial dynamics (Lewis and Steinmo 2012). In summary, this article analyzes the design dynamics converging two approaches. The first approach is actor-centered, seeking to understand their policy choices through action situations that drive policy design—IAD framework (Ostrom 2005; Howlett 2019; Dunlop, Kamkhaji, and Radaelli 2019). Second, how these action situations are shaped by path dependence factors (Pierson 2004).

In this case, we observe two theoretical-analytical dimensions. The first concerns the current conflicts and dynamics of design data governance in Brazil's Federal Government; the second is regarding the approaches and instrumentation to data sharing. In both cases, we identify a path-dependent process focusing on how institutional background influences actors' behavior in the process of designing data governance and data sharing. The analysis of these two dimensions of the design dynamics of data policy takes place through the action situations that policymakers faced to design the policy and its implications, which we contextualize below.

3. Brazil's data governance challenges

Brazil has a long tradition of electronic government and data collection, storage, processing, and sharing processes. The formation of the Federal Service of Data Processing (SERPRO) in 1964 enabled the creation of technological infrastructure and the use of information in policy formulation. In 1974, the Federal Government created the Company for Technologies and Information on Social Protection (DATAPREV), expanding the computational infrastructure for data collection, storage, and processing. This technological infrastructure facilitated the constitution of an electronic government in Brazil from the 1970s to the 2000s, with various information services and data storage and processing.

Priorities regarding data collection and storage circumscribe the government's fiscal and tax policy and social protection policies (Filgueiras, Flávio, and Palotti 2019). Incorporating this technology infrastructure always transpires inside each federal agency or department, with these public companies operating business models for data collection and processing. The Federal Government never has a centralized strategy for all the public sector, and the incorporation, adaptability, and learnings about how to use the technology and apply it to the demands constantly are restricted to the agencies.

Brazil's path of electronic government conditioned changes over time, adjusting and calibrating the digital government's objectives and instruments (Filgueiras, Flávio, and Palotti 2019). The institutional bases of e-government in Brazil stem from a series of policies to facilitate the digitization and support of information and communication technology (ICT) infrastructure. The rules in use that shape the action situations regarding data collected, stored, and processed by ICT infrastructure emerged in 1966. After ICT infrastructure began to be created in the federal government in 1964, in 1966, during the Military Dictatorship, the National Tax Code (Law 5172 of October 25, 1966) was enacted, shaping many elements of data policy in Brazil. In its articles 198 and 199, the National Tax Code (CTN) prohibits the State or its public servants from disclosing "information obtained by official letter on the economic or financial situation of the taxable person or third parties and the nature and status of their business or activities." Data managers interpreted the CTN rules as prohibiting data sharing. This prohibition led to two fundamental institutional statements: first, data is a resource of the collecting organization, which must create and comply with its security standards; second, sharing typically is prohibited, because tax secrecy covers all personal data.

Article 198 of the CTN created a broad concept of tax secrecy for all citizens and companies that resonated with the use of personal data within government. The exception to the breach of tax secrecy is the request of judicial authority concerning criminal investigation procedures. This comprehensive concept of tax secrecy created a rule prohibiting the government from disclosing citizens' and companies' information. The possibility of data sharing involving different public organizations, regulated by article 199, depends on a specific law that authorizes sharing or an agreement between organizations that specifies security standards and responsibilities. The initial choices regarding data policy were to regulate data sharing restrictively. In all policy situations, personal data's confidentiality was an extension of tax secrecy and shaped the action situations concerning data-driven policy.

In 2016, there was a change in the Federal Government's policy regarding e-government and data policy. The Brazilian government launched the Digital Governance Strategy (EGD), incorporating innovation in public services, establishing partnerships, and building a single platform for public services, capable of incorporating the entire structure of services provided to citizens and companies in an integrated manner. From the EGD, the policy of digital transformation was configured, whose central focus was constructing the Gov.br Platform. The Gov.br Platform is a Federal Government strategy to consistently consolidate the entire structure of public services delivered to citizens and companies, promoting the redesign of the entire service structure, digitization, and implementation of big data instruments to promote digital governance goals. EGD's objective was to change the use of ICT in government from a conception of e-government to a conception of digital governance, adopting big data tools.

Using big data in public policies is one of EGD's central objectives. Data from all transactions carried out by citizens and companies are collected, stored, and processed by the Gov.br Platform, which uses surveillance mechanisms to increase data volume, variety, and velocity. The Federal Government, aiming to expand big data's use in the

entire public service structure and in policy formulation, published Decree 10,046 of 2019, facilitating data collection and sharing between federal public sector organizations and between public and private organizations. EGD's objective on data is to create a "knowledge commons" perspective. This perspective is sustained in "the institutionalized community governance of the sharing, and, in some cases, creation, of information, science, knowledge, data, and other types of intellectual and cultural resources" (Frischmann, Madison, and Strandburg 2014).

Decree 10,046 also created the CCGD, whose function is to design the data policy, facilitate the sharing of data collected on the Gov.br Platform and public databases, resolve conflicts, and create guidelines and standards that shape data analysts' and stewards' behavior in different public sector organizations. The central objective of the CCGD is to reduce the risk perception of data stewards to share data among various public sector organizations and facilitate data-driven policy.

In this context—in which the objectives of data policy are to facilitate data collection, storage, processing, and sharing, including public-private relations—the National Congress approved the General Data Protection Law, or LGPD (Law 13,709 of August 14, 2018). LGPD expanded the possibilities of privacy and protection of personal data, while also creating procedures related to data collection, storing, processing, and sharing. LGPD represented a critical juncture that required policymakers to undergo an entire process of calibrating instruments and adjusting the Federal Government's data policy objectives and practices.

LGPD's emergence created a situation of political conflict that shapes the design dynamics of data governance in the Brazilian Federal Government. Addressing this conflict reveals temporal dimensions and instrument calibrations that shape data governance. The CCGD aims to facilitate data sharing in all government structures, but LGPD created new rules that constrain data collection and sharing. In this research, we investigate the data policy's and data governance's design dynamics, focusing on how CCGD's members perform conflict resolution and define procedures, guidelines, and standards for working with data and using big data tools in the Brazilian Federal Government. We therefore focus on how CCGD members design data policy and related data governance tools.

4. Research design and methodology

The CCGD was created by Decree 10,046 of 2019 with powers to deliberate guidelines and standards for broad, restricted, and specific sharing of government data regarding personal data laws' protection. The Committee is also responsible for specifying guidelines and standards for integrating organizations and entities with the Citizens' Base Registry (CBC). Comprised of members from different governmental organizations, the Committee has one representative each from seven public organizations (thus, seven members total): the Secretariat for De-bureaucratization, Management and Digital Government, which chairs it; Federal Revenue; Secretariat of State Modernization of Presidency of the Republic; Office of the General Comptroller; Secretariat of the Presidency of the Republic; Union's General Attorney; and National Institute of Social Protection.

The Committee's objectives for performance are to facilitate data sharing among public organizations to use big data tools in policy formulation. In addition, the Committee provides guidelines for all federal agencies that deal with data. In this sense, the Committee is a proper research object in policy design because the objectives, instruments, and expectations about data governance and policy are formulated in this arena. The Committee's members make decisions regarding data governance drawing on the norms, practices, and guidelines that support other decision making by data stewards in all public sector organizations on data sharing. Also, the Committee's objectives include institutionalizing the CBC, a public database of all Brazilian citizens comprising their essential personal data. The CBC includes names, addresses, identification documents, and health and education data.

The Committee brings together members of the center of government and the two organizations that most collect data in Brazil. The Federal Revenue collects real-time data on all commercial and economic transactions carried out by citizens and companies to facilitate tax collection. The National Institute of Social Protection collects data related to all social policies, especially pension systems and labor relations.

The research we conducted was based on interviews with CCGD members. Seven interviews were collected with the Committee members, following the script in Annex 1. The interviews were carried out between August 21 and October 7, 2021. The interview script was designed to understand the institutional framework that organizes action situations when designing data policy, objectives, and instrumentalization of the data sharing policy, along with outcomes achieved.

The members of the CCGD interviewed occupy positions of secretaries or directors within the Federal Government. Interviews with government elites are essential tools to support case studies like this one, enabling the "investigator to fill in pieces of a puzzle or confirm the proper alignment of pieces already in place" (Aberbach and Rockman 2002, 673). The interviews were analyzed to understand this data policy design puzzle in the Brazilian Federal Government. We extract from these interviews the way these policymakers think and interpret the issue of data sharing (essential in the data policy designed in the EGD) and how they think about data policy instrumentation and policy design. For the construction of the questionnaire, we carried out a previous study of the legislation to identify the rules in use that constrain the design of the data policy. Associated with this previous study of the legislation, we identified in policy documents information about the instruments used for the data sharing process, particularly the CBC and procedures and guides delivered by the CCGD for data stewards.

The research project was registered in the Ethical Committee in the Getulio Vargas Foundation in Brazil, register in the number 176/2021. From these interviews, we examine the action situations that guide CCGD members in designing and redesigning the data sharing policy and related governance structures.

5. Research findings

Based on these interviews, here we report the research findings. This section especially focuses on how action situations regarding big data in public policies shape the design and redesign dynamics of data governance in Brazil.

5.1. Conflict and dynamics of data governance

The interviews revealed conflicts within the design dynamics of the Brazilian Federal Government's data governance and data policy. The CCGD's composition excludes the two public companies of the Federal Government that carry out the data collection, storage, and processing: SERPRO and DATAPREV. In turn, this exclusion means that the state is unable to utilize these companies to create data storage and security protocols, as well as interoperability among databases. The Committee's composition considered organizations from the center of government, giving political scale with the Presidency of the Republic, the Secretariat of Digital Government, and organizations dedicated to auditing and risk-management functions. This Committee composition aimed to amplify the objectives present in the EGD, especially with the expansion of data sharing to support digital transformation and fulfilling related objectives by strengthening big data methodologies applied in policy formulation and public services.

According to interviewee E2, the Committee's composition was exclusively aimed at facilitating data sharing to fulfill EGD's objectives. The intention of having the Office of the General Comptroller participate was to create a low-risk perception among data stewards about data sharing. However, this composition led to a conflicting dynamic for designing data governance and policy. While the Secretariat for Digital Government intended to expand the practice of data sharing, organizations such as the Federal Revenue—the most prominent government data collector—expressed concerns about this practice. Federal Revenue highlights how governmental organizations did not have unique and coherent data security protocols, risk mitigation instruments, and a prohibitive normative background concerning data sharing because of the CTN, created in the 60's. According to interviewee E2, the central character of data governance in Brazil, over time, is that data are collected, stored, processed, and shared by each public organization, creating an idea of ownership and custody because data are thought of as organizational resources and not as part of the commons. Thus, the data governance path assumes that each public organization is responsible for its data, without usage rules creating incentives for sharing, along with difficulties related to different and fragmented data security protocols. According to interviewee E2:

> Look, I share data with anyone who has my level of security or higher. If someone says: "Ah, but this makes it infeasible for the small agency," then I say: they cannot receive the data! […]

> [If] I have a big problem, I mark all the personal registration (CPF) bases that I share with other agencies, because I do not trust anyone. If you analyze this from the CCGD's point of view, it is heresy, it is one agency distrusting the other. But I distrust them because I am the one that appears in a newspaper (if something goes wrong, like a leak). So, this part of the security was a stalemate. They wanted to create a single security standard for the government. No, but the CCGD says it will, and the decree says it will, then I said great, I will stick to the law of the personal registration (CPF) and the CTN, which are legislations above and I will not do it. And I will generate an impasse, so in terms of security as the only real shared base with critical [data], that [for] the other two countries, this has no criticality. The only critical data are mine. It is my security policy. The agencies swear on their feet that they meet the requirements. I doubt it, but it is signed by the highest office holder of the agency saying that it meets the requirements, so I am not going to get into that. I can later be published in the newspaper, but then I will point out that it was you, your fault, you said you attended.[1]

In this sense, data are understood as each agency's own assets, which prevents the idea from evolving into something related to the data commons. The dynamic of conflict emerges when, on the one hand, the Secretariat for Digital Government wants to amplify data sharing among public organizations, resorting to the CTN, either restricting sharing or unifying data governance protocols and standards. On the other hand, according to interviewee E2, the Federal Revenue managed to approve a decision procedure of the Committee, which requires that all decisions be taken by consensus of all representatives, intentionally conflicting with the data policy's objective and creating difficulties for effective instrumentation of data governance. Thus, even though the Committee is an instrument to facilitate data sharing, the institutional choices made are conflicting and, in many points, ambiguous regarding the choice and instruments' calibration of data policy. Interview E1 highlights conflicts getting resolved in the design dynamics, but they demanded much institutional work, even mobilizing the Presidency of the Republic.

> Interviewee: Today the managers thinks they are the data owners. He does not see it as part of the government. The public servant does not see the entire government, let alone the citizen. So, we need to overcome the resistance that comes from [this] culture. However, this resistance is not unfounded. If you look at the LGPD, you will see all the responsibilities of the comptroller, which is the guy who decides about data's treatment. Sharing is one phase of data processing. So, to be authorized, sharing must be done, since the enactment of the law until now [comes] with much responsibility, with many criteria, with obvious questions about the data use, the need, the adequacy, the security and privacy controls, the risk levels, and the transparency instruments. Thus, you combine this concern with a culture, and a history that is not very favorable. So, you create resistance that is gradually overcome.

The conflict of objectives required the government to create standards and regulations, so that on the one hand, sharing is seen as essential for developing digital governance. On the other hand, data protection emerges as an impediment to digital governance but is driven by reduced state surveillance and more realistic cybersecurity standards.

The action situation shaped by the rules in use—CTN, EGD, LGPD, and Decree 10,046—become data as disputed resources by actors, which creates an ambiguous context for data policy. While some of the Committee members act to facilitate data sharing, other members create barriers with the argument of security and protection. The context of ambiguity that builds up over time creates a fragmented, incoherent, and inconsistent data policy design, leading to ineffective instrumentation.

5.2. Approaches and instrumentation for data sharing

The Committee created instrument mixes to address the objectives of data policy and governance in the Brazilian Federal Government. The instruments were selected by the Committee by consensus, following decision procedures specified in Decree 10,046. Furthermore, the definition of instrument mixes follows the procedures set out in this decree, requiring that, together, they are considered institutions and reproduce social values, identities, and worldviews that affect, support, or conflict over their choice (Lascoumes and Le Galès 2007). In this sense, the choice of data policy and governance

tools—with the particular objective of expanding data sharing in the Federal Government structure—follows a complex pattern of instrumentation and control.

The instrumentation pattern started with government databases classifying risk, defining databases as high, medium, or low sharing risk. The classification was considered essential, to meet the LGPD's requirements regarding protection of personal data. Furthermore, the categorization was used as a tool—first, so the CCGD could understand which bases existed in the Brazilian public sector and be aware of their content; and second, to create a formal classification that would allow sharing and overcome agencies' resistance to cooperating.

Associated with the risk classification initiated, the Committee defined guidelines and procedures related to the LGPD, particularly on data sharing. These guides and procedures established a permissive understanding regarding data sharing, intending to reduce the data stewards' perception of risk. The role of the Office of General Comptroller within the Committee was fundamental in creating guidelines and interpretations of the LGPD that reinforce the data sharing policy. These guides emerged in response to the LGPD and required the Committee's members to adapt to data collecting, storing, processing, and sharing procedures.

The CCGD resolutions reflect a standard of procedural instruments related to data sharing, with a particular focus on database interoperability and monitoring. Likewise, the Committee established procedures for constructing the CBC, containing diverse data shared between public sector and private sector. The CBC raises many situations of disputes and conflicts, making incremental progress. However, some interviewees' perceptions are that the CBC does not advance, frustrating one of the objectives of the EGD. This fact stems from the Committee's few advances with the sharing rules and the LGPD's emergence as a risk factor for data stewards, creating barriers and red tape effects because of previous rules in use, particularly the tax secrecy present in the CTN.

Overall, the interviews point to a conflicting dynamic within the CCGD and inconsistency regarding the objectives of data governance design and instrumentation to achieve these objectives. Data governance's design in the Brazilian Federal Government occurs in a context of ambiguous action, marked by disparate rules, making it challenging to implement policy. In this context, actors interpret the data-sharing situation to recover and reinforce previous interpretations, restricting data sharing because of ambiguity. The actors reinterpret the CTN to restrict data sharing and implement this restriction through ad hoc rules—for example, the security requirements pointed out by interviewee E2.

6. Discussion and conclusions

The ambiguity of the rules in use concerning data governance in the Brazilian Federal Government shapes an action situation with conflicting trends, followed by inconsistent instrumentation. This ambiguity reinforces path dependence, causing initial choices to shape current choices. Different interviewees stated the behavior of data stewards as non-adherent to data sharing, creating barriers and difficulties with the justification of the prohibitive rules of sharing. For example, although the CBC intends to create a

single personal database, the incentives remain for each public organization to maintain its data ownership, disparate security protocols, and unique risk management.

This statement provides a first theoretical implication for policy design. The action situations from which policy designs emerge through the interactions and decision making are not "snapshots" conditioned by the moment of decision. Instead, the institutional environment built up over time defines a path dependent policy design, built based on interpretations of the rules in use made by policymakers. Policy designs emerge from policymakers' intentional actions and interests but reiterate temporal factors that shape decisions and choices in conflicting arenas. The inclusion of temporal factors in the construction of design choices, as revealed by the case study and the interviews, confirms that IAD frameworks can be improved with elements of policy development.

The justification given in interviews for this ambiguity is that the LGPD and non-revision of the CTN expand the risk for data stewards. This action situation is shaped by the first choices made in the initial data policy. Specifically, according to the case study, the long pathway of e-government in Brazil created a prohibitive interpretation of data sharing. In this action situation, data governance's and policy's instrumentation will be conflictive and path dependent. In Brazil, data stewards tend to resort to the CTN to justify restrictive decisions regarding data sharing. Although EGD strives to deepen data sharing, data stewards' choices are shaped by interactions that restrict and interpret data sharing as prohibited and risky. This behavior of data stewards, reported in the interviews as the main target of the policy designed by the CCGD, demonstrates that action situations with ambiguous rules in use tend to reinforce path dependence. This situation creates ineffective, incoherent, and inconsistent instrumentation of data governance that emerges inside the CCGD.

Brazilian legislation defines rules in use that have failed to create clear incentives for data sharing. Simultaneously, institutional framework does not grant Committee members space to cooperate over time. In contrast, the history of each public organization is permeated by its development of information storage and security technologies. In this sense, challenges and difficulties persist in building a governance structure for data policy. This history becomes critical when we observe that each institution was responsible (and still is) for guaranteeing the integrity and data secrecy in its governance process.

The perspective opened by the Digital Governance Strategy is one of deep data sharing to facilitate and qualify digital applications and technologies with big data methodologies. However, the design introduced with the EGD were not enough to change data stewards' behavior, keeping the choices initiated with the CTN and critically changed with the LGPD. Although the CCGD has made changes, they are incremental, constrained by tax secrecy, and generate ineffective outcomes because of inconsistent and incoherent instrumentation with the objectives of the data policy and the data governance design. Because institutional choices are path dependent, the design dynamics of data policy in Brazil reinforce conflict situations shaped by data ownership and custody and the fragmentation of data storage and use.

From a practical point of view, policymakers chose a data policy design whose objective is to strengthen data sharing between organizations to support big data in

digital transformation. This choice expressed in the Brazilian EGD was strengthened by creating rules and a committee that would instrumentalize and encourage data sharing. However, internal conflicts in the committee arena provided a process of institutional choices driven by the choices of the past. The path dependence reinforced a prohibitive interpretation of data sharing. This difficulty regarding data sharing creates a significant barrier to digital transformation, in many cases making it impossible for data policy designs to be effective. Even if digital transformation processes want to be disruptive, policymakers need to consider and adapt to the institutional frameworks that condition their choices regarding policy design. In many ways, the findings of this research make it possible to think about why policy designs that coherently express their objectives with the available instruments fail to achieve outcomes.

From theory, this research contributes to explaining policy design over time, considering how path dependence conditions the designers' behaviors and use of tools in the policy process. Policy design and governance theory, concerning the use of data in public policy, can benefit from understanding the historical paths about institutional choices, which shape and make it possible to understand policy designers' behavior to achieve policy objectives, as well as the incentives and constraints that shape their choices. Thus, the policy design dynamic is shaped in contexts of institutional change, enabling the understanding of design over time.

This research initiates a future agenda on policy design crossed with institutional design, understanding the action situations that shape the dynamics of designers' behavior, and the possibilities of understanding policy mixes' effectiveness. The findings of this research makes it possible to understand that the instrument calibration, the definition of policy objectives, and the internal dynamics and choices explain policy design effectiveness and governance. The literature that studies policy design predominantly has focused on understanding public policies' instruments. However, little attention is given to how problems are constructed in the body of policy and how these change over time. Furthermore, this research can refresh the research agenda on digital policy, creating a unique perspective on the use of technologies—such as big data—as essential tools in the policy process, but dependent on choices and institutional frameworks that explain their use, reach, and results.

Note

1. The authors translated quotes of interviewees.

Consent to participate

Survey respondents consented to participate in the research under confidential treatment of their identity.

Ethics approval

The research is registered in Ethical Committee in the Getulio Vargas Foundation in Brazil, register in the number 176/2021.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

ORCID

Fernando Filgueiras 🄳 http://orcid.org/0000-0001-9570-8113
Lizandro Lui 🄳 http://orcid.org/0000-0002-9276-247X

Availability of data and material

Not applicable. Data from this research were collected through interviews and are confidential.

References

Aberbach, J. D., and B. A. Rockman. 2002. "Conducting and Coding Elite Interviews." *Political Science and Politics* 35 (4): 673–676. doi:10.1017/S1049096502001142.

Abraham, R., J. Schneider, and J. Vom Brocke. 2019. "Data Governance: A Conceptual Framework, Structured Review, and Research Agenda." *International Journal of Information Management* 49: 424–438. doi:10.1016/j.ijinfomgt.2019.07.008.

Capano, G., and M. Howlett. 2020. "The Knowns and Unknowns of Policy Instrument Analysis: Policy Tools and the Current Research Agenda on Policy Mixes." *SAGE Open* 10 (1): 1–13. doi:10.1177/2158244019900568.

Carlisle, K., and R. L. Gruby. 2019. "Polycentric Systems of Governance: A Theoretical Model for the Commons." *Policy Studies Journal* 47 (4): 927–952. doi:10.1111/psj.12212.

Carney, P., T. Heikkila, and M. Wood. 2019. *Making Policy in a Complex World*. Cambridge: Cambridge University Press.

Dawes, S. S., A. M. Cresswell, and T. Pardo. 2009. "From "Need to Know" to "Need to Share": Tangled Problems, Information Boundaries, and the Building of Public Sector Knowledge Networks." *Public Administration Review* 69 (3): 392–402. doi:10.1111/j.1540-6210.2009.01987_2.x.

Dunlop, C. A., J. C. Kamkhaji, and C. M. Radaelli. 2019. "A Sleeping Giant Awakes? The Rise of the Institutional Grammar Tool (IGT) in Policy Research." *Journal of Chinese Governance* 4 (2): 163–180. doi:10.1080/23812346.2019.1575502.

Ekbia, H., M. Mattioli, I. Kouper, G. Arave, A. Ghazinejad, A. Bowman, V. R. Suri, A. Tsou, S. Weingart, and C. Sugimoto. 2014. "Big Data, Bigger Dilemmas: A Critical Review." *Advances in Information Science* 66 (8): 1523–1545. doi:10.1002/asi.23294.

Filgueiras, F, and V. Almeida. 2021. *Governance for the Digital World. Neither More State nor More Market*. London: Palgrave.

Filgueiras, Fernando, Cireno Flávio, and Pedro Palotti. 2019. "Digital Transformation and Public Service Delivery in Brazil." *Latin American Policy* 10 (2): 195–219. doi:10.1111/lamp.12169.

Frischmann, B. M. J. Madison, and K. Strandburg. 2014. *Governing Knowledge Commons*. Oxford: Oxford University Press.

Giest, S. 2017. "Big Data for Policymaking: Fad or FastTrack?" *Policy Sciences* 50 (3): 367–382. doi:10.1007/s11077-017-9293-1.

Heikkila, T., and K. Andersson. 2018. "Policy Design and the Added-Value of the Institutional Analysis Development Framework." *Policy & Politics* 46 (2): 309–324. doi:10.1332/030557318X15230060131727.

Howlett, M. 2019. "Procedural Policy Tools and the Temporal Dimensions of Policy Design." *International Review of Public Policy* 1 (1): 27–45. doi:10.4000/irpp.310.

Howlett, M. I. Mukherjee, and J. Rayner. 2018. "Understanding Policy Designs over Time." In *Routledge Handbook of Policy Design*, edited by M. Howlett. London: Routledge.

Kern, F., and M. Howlett. 2009. "Implementing Transition Managements as Policy Reforms: A Case Study of Dutch Energy Sector." *Policy Sciences* 42 (4): 391–408. doi:10.1007/s11077-009-9099-x.

Kitchin, R. 2014. "Big Data, New Epistemologies and Paradigm Shifts." *Big Data & Society* 1 (1): 1–12. doi:10.1177/2053951714528481.

Krasner, S. D. 1988. "Sovereignty: An Institutional Perspective." *Comparative Political Studies* 21 (1): 66–94. doi:10.1177/0010414088021001004.

Lascoumes, P., and P. Le Galès. 2007. "Understanding Public Policy through its Instruments. From the Nature of Instruments to the Sociology of Public Policy Instrumentation." *Governance* 20 (1): 1–21. doi:10.1111/j.1468-0491.2007.00342.x.

Lewis, O., and S. Steinmo. 2012. "How Institutions Evolve: Evolutionary Theory and Historical Institutional Change." *Polity* 44 (3): 314–339. doi:10.1057/pol.2012.10.

Lima, L. L., R. B. Aguiar, and L. Lui. 2021. "Conectando Problemas, Soluções e Expectativas: Mapeando a Literatura sobre Análise do Desenho de Políticas Públicas [Connecting Problems, Solutions, and Expectations: Mapping the Policy Design Literature]." *Revista Brasileira de Ciência Política* [Brazilian Journal of Political Science] 36: 1–41. doi:10.1590/0103-3352.2021.36.246779.

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

OECD. 2019. *The Path to Becoming a Data-Driven Public Sector*. Paris: OECD. https://www.oecd.org/gov/the-path-to-becoming-a-data-driven-public-sector-059814a7-en.htm#:∼:text=It%20presents%20a%20data%2Ddriven,informed%20decisions%20across%20public%20sectors.

Ostrom, E. 2011. "Background on the Institutional Analysis and Development Framework." *Policy Studies Journal* 39 (1): 7–27. doi:10.1111/j.1541-0072.2010.00394.x.

Ostrom, E. 2005. *Understanding Institutional Diversity*. Princeton: Princeton University Press.

Pierson, P. 2004. *Politics in Time: History, Institutions, and Social Analysis*. Princeton: Princeton University Press.

Polski, M, and E. Ostrom. 1999. "An Institutional Framework for Policy Analysis and Design." In *Elinor Ostrom and Bloomington School of Political Economy*, edited by D.H. Cole and M.D. McGinnis, 13–48. Lanham: Lexington Books.

Rayner, J., M. Howlett, and A. Wellstead. 2017. "Policy Mixes and Their Alignment over Time: Patching and Stretching in the Oil Sands Reclamation Regime in Alberta." *Environmental Policy and Governance* 27 (5): 472–483. doi:10.1002/eet.1773.

Siddiki, S. 2020. *Understanding and Analyzing Public Policy Design*. Cambridge: Cambridge University Press.

Van Geet, M. T., S. Lenferink, and W. Leendertse. 2019. "Policy Design Dynamics: Fitting Goals and Instruments in Transport Infrastructure Planning in The Netherlands." *Policy Design and Practice* 2 (4): 324–358. doi:10.1080/25741292.2019.1678232.

World Bank. 2021. *World Development Report—Data for Better Lives*. Washington: World Bank. Retrieved from: https://wdr2021.worldbank.org/the-report/#download.

Annex 1. Interview script

1. The Problem

1.1. How do you understand data governance, and what is its purpose?

1.2. What problems does the Central Data Governance Committee (CCGD) seek to address? Is CCGD aware of policies aimed at using, protecting, and quality of data collected by the Federal Government? How has the CCGD acted on these fronts?

2. Legislation

2.1. How do you see the issue of data protection as regulated in the General Data Protection Law (LGPD)? Is it a factor that facilitates or hinders the use of data in the public sector?

2.2. Has the data governance practiced by the Federal Government sought alignment with the Marco Civil da Internet (Brazilian Civil Rights Framework for the Internet)? How so?

2.3. What are the Brazilian Strategy for Digital Governance's (EGD's) central points, and what are the CCGD's attributions within the strategy?

2.4. How does CCGD interact with the Digital Government Law (Law 14129/2021)?

2.5. Decree 10.046/2019 aims to facilitate the sharing of data from the Federal Public Administration. How is CCGD dealing with this issue of data sharing?

3. Data Governance Instrumentation

3.1. One of the CCGD's tasks is to define resolutions that guide the public administration's different bodies regarding the process of collecting, storing, sharing, and using data. How has CCGD performed these tasks?

3.2. The CCGD issued a resolution defining the categorization of databases. How important is the categorization of databases, and how have agencies performed it?

3.3. Has CCGD been following and monitoring how Federal Government databases are shared and with whom? Who is responsible for this monitoring?

3.4. Has the Federal Government made partnerships with technology companies (such as Google, Amazon, Facebook, Microsoft, among others) involving data storage and sharing? For example, cloud or sharing contracts for developing technology solutions. What is the purpose of these partnerships in terms of public service and public policy innovation?

3.5. Has CCGD carried out campaigns with the agencies to guide the storage, sharing, processing, and use of data in public services and public policies?

3.6. One of the CCGD's attributions is to deliberate on controversies regarding data sharing. How has this dispute settlement been taking place, and what is the Committee's capacity to deal with this issue?

3.7. One of the essential points of Decree 10,046/2019 is the creation of the Citizens' Base Registry (CBC). How has the CBC been built, and what security standards have been adopted to create the Registry?

3.8. Has the CCGD developed any dialogue or partnership with civil society to improve the Federal Government's data governance rules and operational procedures? How has this partnership been?

4. Target Audience and Results

4.1. It is not always easy for data managers from different agencies to interpret the norms that organize data governance. Moreover, data sharing can involve high risk for the manager. Have CCGD's actions contributed to managers reducing the perception of risk from data sharing?

4.2. Have data managers been working to strengthen data collection and qualification mechanisms?

4.3. How do you assess CCGD's performance in guiding data collection, storage, sharing, and use processes in the federal public sector?