

Analysis of open data interoperability through catalogs in Latin America

Roxana Martínez, Gastón Lacuesta and Ezequiel Ricciardi

Universidad Argentina de la Empresa, Buenos Aires, Argentina

Abstract

Today, interoperability between software systems is in full swing. This aspect allows information to be exchanged and reused between different data sources, as well as between software. One of the fields of application is state governments that share open public data and facilitate joint management and transparency between organizations and citizens. This paper addresses the importance of open data interoperability in the governmental context. This article presents a review of various open government initiatives in Latin America, in addition, it analyzes a sample of sets of open public data catalogs extracted from government web portals in Latin America (Brazil, Mexico, Colombia, Uruguay, Chile, and Argentina), to study possible shortcomings in aspects of accessibility, interoperability and contextualization of these to mitigate them in the future. For said analysis, a self-developed software tool is used that uses validation criteria based on initiatives of the Argentine Republic. Finally, an analysis of the findings is carried out to propose aspects of improvement through good practices in their treatment and help in the standardization of certain technical and structural aspects and levels of the exchange processes.

Keywords

Data Interoperability, Open Data Catalogs, Open Government

1. Introduction

In today's global landscape, the digital era has reshaped how we access, distribute, and use information. Information, in this setting, is not only seen as a precious asset but also as a vital tool for making decisions, generating knowledge, and enhancing government transparency. Interoperability, defined as the capacity of various systems and applications to exchange and process information seamlessly, stands out as a central element in this evolving paradigm.

Latin America, despite its rich cultural, political, and economic diversity, several nations in the region, such as Brazil, Mexico, and Colombia, have shown an increasing dedication to enhancing the interoperability of their information infrastructures. However, these praiseworthy initiatives still encounter inherent challenges related to data standardization, updating, and structuring.

This study sets out to provide a thorough and critical examination of open data catalogs from government websites in selected Latin American nations. Using a purpose-built software tool, we will scrutinize these catalogs to pinpoint areas of improvement, shortcomings, and


ICAIW 2023: Workshops at the 6th International Conference on Applied Informatics 2023, October 26–28, 2023, Guayaquil, Ecuador

✉ mariarmartinez@uade.edu.ar (R. Martínez); glacuesta@uade.edu.ar (G. Lacuesta); ericciardi@uade.edu.ar (E. Ricciardi)

ORCID 0000-0001-9318-6411 (R. Martínez)



© 2023 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

 CEUR Workshop Proceedings (CEUR-WS.org)

exemplary practices. Moreover, we will aim to place these discoveries within the broader context of global movements in interoperability and data openness.

Through this research, we hope to add to the international scholarly body on interoperability and open data, delivering valuable perspectives on the current state and future outlook of this topic in Latin America, a region with its own set of unique digital public management challenges and opportunities.

The structure of this work presents a brief introduction to the subject, then the problem is developed in this context, later, a comparative study of some Latin American countries on this situation is shown, and then an analysis of data collected from the countries is explained. government websites of the countries studied, in which some data catalogs are analyzed using the proposed software tool, which makes it possible to identify some shortcomings in the contents of these public files. Finally, it culminates with the analysis and proposal of good practices to facilitate the exchange of data between different organizations.

1.1. Importance of Interoperability

The concept of open government is booming in different countries around the world. This approach consists of an open society, that is, a society in which the state of a country and its citizens are the main actors in which they participate in various activities that benefit different people in social and political aspects. This approach presents an active participation of citizens in public affairs and has some pillars that are essential to consider, these are: citizen transparency, accountability, citizen participation, and the use of various technologies and innovations that favor this context.

Open government refers to the freedom of information that is available to citizens and government authorities, this makes governments more transparent and accountable in their activities. To carry out the latter, it is necessary to develop freedom of information laws to improve transparency, accountability, and public trust [1]. This context allows for an optimal administration in which ethical behavior and professionalism in public administration improve. Therefore, this generates a relationship of greater trust between those involved, thus revealing greater social value and wellbeing [2].

One of the most important associations at the international level is the Open Government Partnership (OGP)¹, which is based on the idea that an open government is more accessible, more receptive, and accountable to citizens and that it improves the relationship between people and their government has long-term benefits for all. OGP includes members from various countries who work together to cocreate two-year action plans with concrete steps (commitments) on a wide range of issues.

Several points are related to the commitment of governments to improve services, support citizen participation, the administration of public resources, and promote public innovation. This last point is essential to increase the availability of information, both to citizens and to various government agencies. It is important to keep in mind the importance of awareness, the dissemination of activities, as well as having digital platforms that provide support in the publication of data. This leads to the need to use open standards to promote civil society access

¹Open Government Partnership (OGP). <https://www.opengovpartnership.org/es/about/approach/>

to public data, as well as to facilitate the interoperability of government information systems². Innovative technologies offer a wonderful opportunity to share information and provide public services with greater transparency.

That information on the activities of public entities is created and is available to the public, on time and in open data formats with no limits to reuse, allowing an initiative-taking work methodology by public entities [3]. Beyond using public policies that support the idea of open government, it is important to work on the technology approach. This work is based on the analysis of the efforts of public organizations to achieve adequate interoperability and availability of public data.

The next section discusses the importance of interoperability and its benefits. Interoperability is indispensable for the practical cooperation of digital public services. It is an essential aspect to improve the efficiency and effectiveness of the provision of services, and at the same time, reusing interoperable government technological solutions could allow for reducing costs in technological platforms. It is important to understand that in addition to the public data that can be offered on public platforms, a strategic analysis must be presented on the digital service infrastructures that are available. It is necessary to analyze the access to digital resources; service infrastructure; platforms that will be provided to the community for the interconnection of data; access to reusable public sector information; and crossborder interoperable online services³.

1.2. Context of the problem

Although the concept of interoperability provides a collaborative space in which governments expose and consume services constantly, it can be defined that the success of this proposal lies in the constant feedback between the parties. As part of the public digital transformation, it is essential to have the appropriate regulatory framework and technology that allows the secure exchange of documents and data between public entities, so this is one of the problems to be considered in each of the governments. I would like to use this approach. Proper use would allow citizens to access the services provided by the administration, avoiding having to submit the same information repeatedly. In other words, one of the shortcomings to work on in this context is data redundancy, since it should prevent the administration from generating data replicas, keeping the authentic source safe, and making them available to those who require groups of data to fulfill your task-interconnection of data; access to reusable public sector information; and cross-border interoperable online services⁴.

The investigations refer to the fact that there are several challenges and impediments to open government management, identified in three categories: impediments to data access (data availability and searchability), data use (related to problems with usability impediments, understandability, quality, compatibility and metadata) and data deposition, that is, difficulties

²Open Government Partnership (OGP). <https://www.opengovpartnership.org/process/joining-ogp/open-government-declaration/>

³Gobierno de España. https://administracionelectronica.gob.es/pae_Home/pae_Estrategias/pae_lineas_ccoperacion/pae_Cooperacion_Internacional/pae_Interoperabilidad_Coop_Inter.html

⁴Gobierno de España. https://administracionelectronica.gob.es/pae_Home/pae_Estrategias/pae_lineas_ccoperacion/pae_Cooperacion_Internacional/pae_Interoperabilidad_Coop_Inter.html

in the interaction with the data provider and the challenges to release the data sets, as well as those of communication between the end user and the data of the entity [4].

One of the approaches proposed by the government regarding the aforementioned aspects is to make datasets openly available to the public on government portals, to disseminate citizen transparency and strengthen the link between government and citizens. One of the main problems encountered with these open data portals is the inconsistency of some data since not all public organizations follow a particular standard for the treatment of dataset content. There are no specific mandatory international standards. This drawback leads to poor data quality, and as a result, various problems can occur when trying to reuse data for review, analysis, or interoperability across data sources. To improve the quality of the open public data provided, it is necessary to introduce a set of standards or aspects that are consistent and sustainable over time, which will give greater possibilities of analysis to the information provided, as well as new knowledge about possible preventive measures. This will allow avoiding flaws in the data analysis process and, in addition, avoiding erroneous or biased results that lead to poor decision-making.

It is important to consider that in open data portals, the availability of these does not always coincide with their quality, unfortunately, today it continues to be a difficulty and is a great challenge for public policies [5]. The analysis of many of the public data sets represents a crucial problem, since they are not standardized and, in some cases, outdated. The publication of public governmental data presents problems of format standardization and connection to heterogeneous data sources [6], so it is essential to have procedures that ensure the quality of the content of the published data.

2. Interoperability in Latin America

In Latin America, the interoperability of government systems continues to be a critical issue. Although projects to improve interoperability have been implemented in many countries in the region, significant challenges still exist. One of the main challenges is the integration of systems at the municipal and regional level since in many cases each municipality or region uses its computer systems, which makes communication and information exchange between government entities difficult.

Another major challenge is the lack of data standardization. Many government systems use different formats and protocols to store and transmit information, making systems integration difficult. Data standardization is critical to effective interoperability, as it allows computer systems to communicate effectively. The lack of common standards can make it difficult to compare data and generate consolidated reports, which can affect informed decision-making. In addition, information technology procurement and contracting processes are often complex and lengthy, making it difficult to adopt more modern and efficient technologies.

To address these challenges, further work on data standardization and systems integration is needed. Governments in Latin America can promote the adoption of common data standards and communication protocols to improve interoperability between government systems. In addition, it is important to improve the training and education of government personnel so that they can effectively use information systems and work together to improve the efficiency and

transparency of public administration. A focus on the education and training of government personnel is needed to ensure that they can effectively use information systems and work together to improve the efficiency and transparency of public administration.

The selection of countries such as Brazil, Mexico, Colombia, Uruguay, Chile, and Argentina for this study is based on rigorous criteria that reflect their prominent position in the Latin American landscape and their commitment to adopting digital policies. Brazil, Mexico, and Colombia, for instance, are leading economies in the region and have proven to be pioneers in the adoption of digital initiatives. Meanwhile, Uruguay, Chile, and Argentina have positioned themselves as models in terms of open data policies, showcasing notable advancements in promoting interoperability. This selection of countries aims to offer a balanced and representative view of current interoperability trends in Latin America. Geographic representativeness, the magnitude of open data initiatives, and the relevance of their policies on the international stage were prioritized. In this way, the study seeks to provide valuable insights and contribute to the international academic discourse on the subject, considering the particularities and challenges presented by the region.

The following section analyzes aspects of interoperability in the government portals of some Latin American countries. The study surveyed below arises from the analyses carried out on the government websites of these countries, considering the most relevant aspects of data interoperability.

2.1. Brazil

The complexity of public administration in Brazil has characteristics typical of a large territory with a large population that stands out in Latin America. Currently, it is made up of more than 5,560 municipalities throughout the entire territory that until 2003 had some autonomy to purchase software and manage computer networks. It was in that year that the Secretary of Logistics and Information Technology of the Ministry of Planning, Budget, and Management decided to focus on interoperability. The decision was immediately supported by the state IT entities represented by the Brazilian Association of State Entities [7]. Moser [8] argues that centralization, although essential, is only a step towards democratizing data. The inclusion of data generated by civil society, academic institutions, and NGOs is crucial to achieving truly inclusive and diverse data management.

In Brazil, there are various open data catalogs at the national, state, and municipal levels. The Brazilian Open Data Portal is a government initiative that has been instrumental in promoting transparency and citizen participation in Brazil. This online resource has stood out as a centralized platform for the publication and access to data sets generated by public institutions in the country⁵. These catalogs typically provide detailed information about each dataset, including its description, format, update date, relevant metadata, and download links. This makes it easier for users to identify and select the data sets of interest, as well as understand their content and structure.

In addition, open data catalogs in Brazil often implement advanced search functions and filters, allowing users to make specific queries and quickly access the desired information. This

⁵Portal Brasileiro de Dados Abertos.<http://dados.gov.br/>

contributes to improving usability and efficiency in accessing data, promoting its reuse and value creation. The data sets available on the portal cover a wide range of areas, including health, education, environment, security, transportation, and economy, among others. This means that users can access a significant amount of information to explore various issues of social and economic importance.

In addition, the Brazilian Open Dice Portal aims to strengthen civic engagement. By providing citizens with the information, they need to understand and evaluate government activities, the platform fosters an active and informed citizenry. Citizens can use the data to monitor public policies, participate in decision-making, and hold the government accountable. The Brazilian government has implemented a series of policies and programs to promote the interoperability of its information systems. One of the most important programs is the National Interoperability System (SIN) [9], which establishes the standards and protocols necessary to guarantee the interoperability of government information systems. Another program is e-PING (Padrões de Interoperabilidade de Governo Eletrônico) which is a set of interoperability norms and standards for Brazilian electronic government, established in 2003. The main objective of e-PING is to guarantee the interoperability of systems and services electronic devices used by the Brazilian government, to improve the efficiency, transparency, and quality of public services provided to citizens. However, it is necessary to address some challenges associated with open data catalogs in Brazil. Among them is the need to improve the quality and constant updating of the data, as well as the standardization of the formats and structures of the data sets to facilitate their interoperability. Likewise, it is essential to promote awareness and training in the use of open data catalogs, both by producers and users, to maximize their potential and benefits [10].

Janssen, Charalabidis, & Zuiderwijk [10] add another dimension to this debate, emphasizing that the adoption of technologies and platforms is only a part of the equation. Education and training are essential to ensure that the benefits of data management extend to the entire population. This perspective is supported by López & Criado [9] and Instituto Nacional de Tecnologia da Informação, ITI⁶, who highlight the importance of data accessibility and reuse for all citizens. Furthermore, the Brazilian government has implemented programs such as the National Interoperability System (SIN) and e-PING, which seek to establish standards and protocols for interoperability. These initiatives reflect a concerted effort to ensure that government computer systems are cohesive and efficient.

While Brazil has made notable advances in data management and interoperability, there is still work to be done. Collaboration between the government, civil society, and the academic sector will be essential to overcome the remaining challenges and ensure that data management in Brazil is truly inclusive and beneficial for all.

2.2. Mexico

The Mexican government has established a series of policies and programs to improve the interoperability of government systems. The National Digital Strategy, for example, establishes a framework for the digital transformation of government, which includes improving the

⁶Instituto Nacional de Tecnologia da Informação. <https://www.gov.br/pt-br/orgaos/instituto-nacional-de-tecnologia-da-informacao>

<https://www.gov.br/pt-br/orgaos/>

interoperability of systems and promoting collaboration between different government agencies and entities⁷.

Within this framework of digital transformation, the National Digital Platform (PDN) stands out, an initiative of the Mexican government that seeks to centralize and standardize government information to improve efficiency and transparency⁸. The PDN focuses on the use of web services and APIs for systems integration, the creation of platforms and tools for data and information management, and the adoption of cybersecurity practices.

Among the specific initiatives and projects that have been carried out to improve the interoperability of systems in the Mexican government, the following can be mentioned: a) The creation of the National Interoperability System (SNI), which aims to allow the integration of different government information systems and interoperability between them. This system is based on the use of common standards and protocols, such as the XML language and the use of web services⁹; b) The adoption of the Unique Digital Identity (IDU) as a common standard for the identification and authentication of users in government systems. This allows citizens to access different government services more easily and securely, without having to provide their personal information each time they register in a new system.

The National Digital Platform (PDN) is a Mexican government initiative to centralize and make government information accessible to foster transparency, improve the efficiency of public services, and promote data-based decision-making. The National Digital Platform uses state-of-the-art technologies, such as the blockchain, to guarantee the security and authenticity of the data. In addition, it promotes system interoperability by using open standards and provides APIs for data access. In this way, the PDN is an essential tool for the digital transformation of the Mexican government.

The creation of the Open Data Portal of the Government of Mexico, allows access and reuse of public information by citizens, companies, and civil society organizations. This portal is based on open data standards and promotes the interoperability of systems through the publication of data in common and reusable formats¹⁰.

The Mexican government has promoted the creation of an open data catalog at the national level to promote transparency and access to public information. This catalog, hosted by the National Transparency Platform (PNT), centralizes defective data sets by various government institutions¹¹. Through the catalog, users can explore and download data in open formats, which encourages the reuse of information for various purposes, such as research, application development, and knowledge generation. The open data catalog in Mexico not only includes data sets from the Federal Government but also from state and municipal governments that have adopted transparency and openness policies [11].

Other authors, Pérez and Rodríguez [12] have pointed out how the National Digital Strategy stands as a pillar in this modernization process. This strategy, supported by official documents from the Government of Mexico¹², not only seeks process digitalization but also the creation of

⁷Portal de Datos Abiertos del Gobierno de México. <https://datos.gob.mx/>

⁸Gobierno de México. <https://www.plataformadigitalnacional.org/about>

⁹Sistema Nacional de Transparencia. <https://www.plataformadetransparencia.org.mx/>

¹⁰Portal de Datos Abiertos del Gobierno de México. <https://datos.gob.mx/>

¹¹Sistema Nacional de Transparencia. <https://www.plataformadetransparencia.org.mx/>

¹²Estrategia Digital Nacional.

an ecosystem where interinstitutional collaboration becomes the norm. The idea is for different government entities to efficiently share information and resources, eliminating redundancies, and improving the citizen's experience. Within this ecosystem, the National Digital Platform (NDP) stands out as a key tool. As noted by Ramírez [13], the NDP not only centralizes information but also employs advanced technologies such as web services, APIs, and blockchain to ensure data security and authenticity. This centralization and securitization are essential for instilling trust in citizens and ensuring that the provided information is current and relevant. However, centralizing information is not enough if systems cannot communicate with each other. In this regard, García and López [14] address the relevance of the National Interoperability System¹³. The NIS aims to have different government systems "speak" the same language, using common languages and protocols like XML and web services. This interoperability is crucial for information to flow efficiently and for citizens not to have to provide the same information multiple times on different platforms.

Furthermore, for digital transformation to be effective, citizens need to interact with the government simply and securely. Martínez and Sánchez [46] emphasize the importance of the Unique Digital Identity (UDI) in this process. The UDI allows citizens to uniformly identify and authenticate themselves in various government services, simplifying access to these services and enhancing information security. On the other hand, transparency is a key element in any government modernization process. Rodríguez and Mendoza [15] discuss the role of the Government of Mexico's Open Data Portal in this regard. By providing data in common and accessible formats, the portal not only promotes transparency but also encourages citizen participation and enables third parties to use this data to generate knowledge and value.

Digital transformation in the Mexican government is a complex process involving multiple facets, from centralizing information to promoting transparency and citizen participation. Although there are challenges along the way, the ongoing initiatives and projects demonstrate a clear commitment to continuous improvement and adaptation to the demands and needs of the digital era. It is essential to maintain this momentum and continue working together to ensure a modern, efficient public administration that serves the citizens.

2.3. Colombia

Interoperability and the open data catalog play a key role in the context of public administration in Colombia. These aspects are essential to promote transparency, efficiency, and citizen participation in access to public information [16]. In Colombia, some various regulations and policies promote the interoperability and integration of computer systems in the public sector. Some of them include the National Information and Communications Technology Policy, Decree 1078 (2015), Law 1341 (2009) and Resolution 2416 (2016). The Colombian government has implemented a series of initiatives to improve technological infrastructure and connectivity throughout the country. This has allowed a greater integration of the computer systems of different government entities. Despite advances in technological infrastructure and regulations, the interoperability of Colombian government information systems remains a challenge. Many government entities operate isolated systems and there is no effective integration between them.

¹³Sistema Nacional de Interoperabilidad. Gobierno de México

The open data catalog in Colombia not only includes information from the public sector but also incorporates data generated by civil society, nongovernmental organizations, and academic institutions. This broadens the diversity and scope of the available data sets, fostering the collaboration and participation of different actors in the generation and dissemination of relevant information. The Datos.gov.co portal is an initiative of the National Government of Colombia, led by the Ministry of Information and Communication Technologies (MinTIC) and the National Digital Government Agency (ANGD). Through this portal, it seeks to promote the openness and transparency of government data, and allow its access and reuse by citizens. The Ministry of Information and Communication Technologies has led the “Online Government” project, which seeks to integrate the computer systems of government entities to improve the provision of public services. One of the most important programs is the Interoperability Reference Framework (IMR), which establishes the standards and protocols necessary to ensure that government information systems can communicate with each other effectively.

Some of the challenges in the interoperability of Colombian government information systems¹⁴ include the lack of standardization that makes it difficult to share data between different government entities since the data is in different formats; lack of coordination between agencies with different priorities and objectives that make collaboration and cooperation difficult; and resistance to change. However, there are opportunities to improve interoperability through the adoption of emerging technologies, training, and strengthening the culture of collaboration among government entities.

However, there are challenges associated with the open data catalog in Colombia that require attention. One of them is to guarantee the quality and constant updating of the data sets, as well as the veracity and integrity of the published information. Likewise, it is necessary to improve the usability and accessibility of catalogs, ensuring that users can perform intuitive searches and easily find the data they need.

Although efforts have been made to promote open data and interoperability in the Colombian public sector, there is still little participation by some public entities in the publication of open data and the implementation of interoperability standards. This limits the scope and availability of open data sets.

2.4. Uruguay

Uruguay has developed an interoperability policy for the public sector that seeks to establish the necessary guidelines so that information systems can share data and services efficiently and securely, and thus improve the quality of citizen care. This interoperability policy is framed in the Law on Access to Public Information (Law No. 18.381), which establishes that all information in the framework of public management is a public good that must be available to citizens and dealt with transparency. Established a series of initiatives and policies to promote data interoperability by creating the Agency for Electronic Government and the Information and Knowledge Society (AGESIC) and has led these efforts, developing the Interoperability and Standards Framework (MIyE) as a guide for the integration of data and systems.

¹⁴Revisión del Gobierno Digital en Colombia: Hacia un Sector Público Impulsado por el Ciudadano. <https://www.oecd.org/publications/revison-del-gobierno-digital-en-colombia-9789264292147-es.htm>

The National System for Productive Transformation and Competitiveness (Transforma Uruguay) is in charge of leading the implementation of the interoperability policy in Uruguay. The program has developed a series of initiatives to promote interoperability between the different State information systems, such as a) A catalog of interoperability services that allows information systems to share services and data more efficiently and securely. This catalog includes services such as user authentication, information consultation, and data validation; b) A methodological guide for the design and implementation of interoperability services, which serves as a reference for developers of information systems in the public sector. This guide describes the steps required to design and implement interoperability services, from defining the requirements to commissioning and monitoring the services; c) A business architecture model that makes it possible to align information systems with the strategic objectives of the government and guarantee interoperability between them. This model defines the key components of the enterprise architecture and describes how these components interact with each other to support government business processes.

Regarding interoperability standards and protocols, the Uruguay Digital Government website uses open standards and technologies for the development of its online services, such as HTML, CSS, JavaScript, and JSON. In addition, it is in the process of implementing the SOA (Service Oriented Architecture) interoperability standard for the integration of different government systems and platforms. The open data catalog in Uruguay is a key tool to facilitate access and reuse of public information by citizens. The Uruguayan Government has developed the Datos.gub.uy portal, which functions as the main open data repository in the country. The open data catalog in Uruguay, through Datos.gub.uy, has fostered transparency and accountability, allowing citizens to access relevant information and use it to make informed decisions, monitor government performance, and develop innovative projects.

This has allowed a greater integration of the computer systems of different government entities. Despite advances in technological infrastructure and regulations, the interoperability of Colombian government information systems remains a challenge. Many government entities operate isolated systems and there is no effective integration between them.

2.5. Chile

One of the main programs for the promotion of interoperability in Chile is the Electronic Government program, which aims to improve the quality of public services and government transparency through the use of information and communication technologies. In this sense, the Chilean government has developed an online services platform, which integrates more than 800 procedures and online services offered by different public entities. This platform allows citizens and companies to conduct procedures and requests electronically, avoiding the need to hide from public offices.

Regarding the interoperability of the Chilean government's computer systems, work has been done on the definition of standards and protocols for the exchange of information between the different government systems and platforms. In this sense, the SOA (Service Oriented Architecture) interoperability standard has been adopted for the integration of different government systems and platforms.

In addition, the Chilean government's open data portal has been created, which allows

citizens and companies to access public information from different government agencies and areas. This portal is in line with open data and government transparency initiatives worldwide. Another important initiative in terms of interoperability in Chile is the National Territorial Information System (SNIT), which is an integrated geospatial information platform that allows access to information from different agencies and government areas related to territory and the environment. This platform uses open standards and technologies for the integration of different government systems and platforms. One of the most prominent catalogs in Chile is *Datos.gob.cl*, managed by the Digital Government Agency (AGD). The catalog houses a wide range of data sets from various government entities in Chile. This includes data produced by ministries, statistical institutes, municipalities, and other agencies. The catalog includes data analysis and visualization tools that allow users to interactively explore and understand information. These tools make it easy to create charts, maps, and visual aids that help identify patterns, trends, and relationships in data. This promotes a deeper understanding of the information and encourages its use in the generation of knowledge and decision-making.

2.6. Argentina

One of the main objectives of government interoperability in Argentina is to simplify and expedite the procedures and services that citizens need to carry out before the State. In Argentina, the government has taken important measures to promote the interoperability of systems and improve the efficiency of the public administration as a whole. One of the main instruments to achieve this is the Law on Access to Public Information No. 27,275, sanctioned in 2016, which establishes a legal framework for the implementation of open data policies and the promotion of interoperability of systems in the public sector. The Distance Procedures Platform (TAD), which allows citizens to carry out procedures online with different state agencies, was created in 2012 and currently has more than 1000 procedures available.

In addition, the government has worked on the implementation of common standards and protocols to ensure interoperability between different government systems. The Biometric Authentication System for the National Registry of People (Renaper) has also been developed, which allows the identification and authentication of people using biometric technology. One of these standards is the Digital Identity System (SID), which allows citizens to have a unique digital identity to access different online services.

Another important initiative is the Electronic Document Management System (GDE), which allows the electronic management of documents and files in the public sector. The system is designed to be interoperable and allows integration with other government systems.

In Argentina, “the data catalogs represent a list of data assets of a state agency with their corresponding metadata, which describes what they are and how they can be used. This helps you find and better understand public data on a given topic. The concept of the data sets, better known as datasets, refers to the fundamental part of the entire data catalog, that is, the group of one or more data resources. Maintaining quality in datasets and their catalogs, makes correct accessibility possible for citizens, for example, what these concepts go together” [5].

In summary, the interoperability analysis of each country fulfills a key aspect in the digital transformation, and as explained above, different initiatives and projects have been conducted to improve it. Although there are still challenges and opportunities for improvement in this area,

the different governments are committed to the continuous improvement of the interoperability of systems to guarantee efficiency and effectiveness in the provision of public services. One of the challenges associated with the open data catalog is ensuring the quality and constant updating of the information. It is necessary to establish clear processes for maintaining and updating data sets, as well as feedback mechanisms and citizen participation to improve the quality and relevance of the data provided.

3. Analysis of Open Data Catalogs

For this work, a proposal of the relevance of open data catalogs that were extracted from the government websites of the analyzed countries is analyzed.

The objective of this study is to validate the content of these files through the software developed by the authors, which presents a series of validation criteria that were considered for the Argentine Republic. The choice of said country was since a series of criteria explained on its government website for their analysis were identified, which allowed the design of a prototype oriented to these validation aspects. It should be noted that the criteria proposed for this country are considered good international practices, developed on its official website¹⁵¹⁶¹⁷.

This research work is an expansion of studies carried out [17, 18]. This new proposal presents an analysis approach especially in:

- a) Current status of the content of open data catalogs used in government portals in Latin America (some countries were analyzed as a sample), which is explained in the section “Open Data Interoperability in Latin America”.
- b) Sample of oriented Open Data catalogs.
- c) Use of a validation software prototype that results in a quality status of the content of the catalog structure for CKAN¹⁸, which is a tool developed by the Open Knowledge Foundation.
- d) Flaws detected and proposals for improvements in the analysis carried out in this investigation.

3.1. Data Sample

There are numerous open data catalogs in various Latin American countries. Various government websites were analyzed, some of these are: a) Brazil¹⁹²⁰; b) Mexico²¹²²; c) Colombia²³²⁴; d)

¹⁵Datos Argentina. <https://datos.gob.ar/>

¹⁶Ministerio de Salud. Datasets. <http://datos.salud.gob.ar/dataset>

¹⁷Perfil de Aplicación Nacional de Metadatos para Datos Abiertos. <https://datosgobar.github.io/paquete-apertura-datos/perfil-metadatos/>

¹⁸The world's leading open source data management system. <https://ckan.org/>

¹⁹Portal da Transparência da Mobilidade Urbana de Natal. <http://dados.natal.br/>

²⁰Portal de Dados Abertos da Cidade do Recife. <http://dados.recife.pe.gov.br/>

²¹La plataforma cívica de datos abiertos de México. <https://datamx.io/>

²²Gobierno de la Ciudad de México. <https://datos.cdmx.gob.mx/>

²³Alcaldía de Santiago de Cali. <http://datos.cali.gov.co/dataset>

²⁴Datos Abiertos Bogotá. <https://datosabiertos.bogota.gov.co/>

Uruguay²⁵²⁶; e) Chile²⁷²⁸; Argentina²⁹³⁰, and it was found that these countries have decentralized repositories by commune, municipality, locality, city, and state as appropriate, in addition to a central repository of federal aspect.

Table 1 shows the number of datasets identified by government websites and the number of data types detected. It is observed that the types of open data formats are various, open data formats allow interoperability between software, but there are public open data that are not in a suitable standard format, for example, the open PDF format.

Table 1

Number of datasets and types of formats by country

Country / Website	Number of Datasets	JSON	PDF	XLS	CSV
Brazil ³¹	19	0	62	0	82
Brazil ³²	150	275	89	0	583
Mexico ³³	9717	813	810	943	4543
Mexico ³⁴	462	1	30	62	427
Colombia ³⁵	651	5	9	4	506
Colombia ³⁶	2164	15	31	33	1127
Uruguay ³⁷	2408	2109	30	110	2262
Uruguay ³⁸	79	0	0	0	46
Chile ³⁹	3736	1	76	993	2382
Chile ⁴⁰	120	0	0	0	114
Argentina ⁴¹	1176	*	*	*	*
Argentina ⁴²	56	21	6	33	46

**You must filter by each category to obtain the total by type of format.*

In addition, in Table 1, it can be seen that the type of format most used in the analysis sample is the CSV type, which remains one of the most standard formats in the selection of data available in Latin American government portals.

It is important to mention that, regarding the periodicity of the published datasets, it was identified that they are updated daily, weekly, monthly, quarterly, and annually. In addition, there is a measurable aspect that is "AccrualPeriodicity", which corresponds to the updating of the data, every 10 years, every 4 years, every 2 years, every 1 year, every 4 months, every 2 months, twice a month, three times a month, once a week, among others. For example, in Argentina, approximately 50% of the datasets are updated monthly.

In the next subsection, the validations carried out with the software prototype made are briefly explained.

²⁵ Catálogo de Datos Abiertos. <https://catalogodatos.gub.uy/>

²⁶ Bienvenidos al Portal de Datos Abiertos de Montevideo. <https://ckan.montevideo.gub.uy/>

²⁷ Datos.gob el repositorio de datos abiertos centralizado del Estado. <https://datos.gob.cl/>

²⁸ Consejo para la transparencia. <https://www.consejotransparencia.cl/datosabiertos/>

²⁹ Datos Argentina. <https://datos.gob.ar/>

³⁰ Ministerio de Salud. Datasets. <http://datos.salud.gob.ar/dataset>

3.2. Validation Software Prototype

A prototype was developed for the validation of part of the sample that, through a processing script in Python, processes the open data catalogs in JSON, that is, it analyzes the metadata in JSON format, usually called “data.json”. This prototype generates two reports as a result in which the analysis of the structure of open data catalogs is shown. This is done using the methods provided by the “pydatajson” library⁴³. It is validated that:

- The URL present in the open data catalog field is available (through the UrlValidator field).
- The script validates that the files associated with the distributions are consistent formats (through the ConsistentDistributionFieldsValidator field).
- It is validated that there are no duplicate elements (through the Theme Ids Not Repeated Validator field).

In the new version that is in full development, graphics were integrated regarding the dataset update frequency and its update to date according to its data type. Data from the DCAT-type scheme was incorporated. Finally, improvements in the detection of processing errors, and performance improvements, for example: pydatajson vendorization for changes and subsequent PR, visualization of booleans instead of integers, messages cut off, bugs, and index array, are in the testing phase. out bounds, null pointers, etc.

In summary, the validation of the content of open data catalog files is carried out from two approaches:

- From the use of the Python library, called “pydatajson”:** A validation is carried out on certain fields in the catalogs that cannot be covered by the JSON Schema library. For example, the landing pages of each dataset. The open data opening package indicates that it is mandatory that this link be present, and it must be functional (accessible and with a satisfactory response code), this is what is validated. Another custom validation is done for the access and download URLs of the distributions of the data sets. Finally, it is verified that there are no taxonomically repeated catalogs in the files, this is done through the topic ID. It is validated that the distributions have a specific format.
- Through a proposed own development:** Initial prototype uses a processing script in Python language. This one has like functionality, receiving a list of CKAN – or Andino – portal nodes, or raw metadata in JSON or XLS format, usually called “data.json” or “data.xls” correspondingly. The developed script will generate two reports with a .CSV file extension, as a result of the validation generated on the analyzed file of the open data catalog.

3.3. Analysis of the results

Based on the sample indicated in the previous section, the self-developed prototype was used, to carry out a general study of the structures of the public data files that are available in the government portals of Latin America.

⁴³GitHub. <https://github.com/datosgobar/pydatajson>

It is worth mentioning that the prototype was developed with technical aspects that are used in Argentine government portals, that is, a guide to good practices for metadata⁴⁴, oriented to CKAN, so that the analysis of results could focus on aspects of noncompliance for not taking the same criteria of said country.

Table 2

Classification of aspects detected in Latin America

Case	Country	Type of Aspect	Description	Recommendation
1	Brazil; Chile; Colombia; Mexico; Uruguay	'superTheme' is a required property.	For these countries, no LATAM dataset has the "superTheme" property.	It should have a taxonomy of those specified in ⁴⁵
2	Brazil; Chile; Colombia; Mexico; Uruguay	'accrualPeriodicity' is a required property or value is invalid'	Some LATAM datasets do not have the "accrualPeriodicity" property or if they do, it does not have a valid value.	This field is the periodicity with which the data set should be updated.
3	Brazil; Chile; Colombia; Mexico; Uruguay	'**' is too short	730 of the 19921 LATAM datasets have a very short field title.	Minimum 4 characters.
4	Brazil; Chile; Colombia; México; Uruguay	'**' is too long	71 of the 19921 LATAM datasets have a very long field title.	Maximum 60 characters.
5	Uruguay	'**' is not a uri	79 of the 2408 Uruguay datasets do not have valid values in some of their fields, such as "downloadURL".	It must be a URI. This is validated by regex.
6	Brazil	'**' is not a uri	56 of the 169 datasets from Brazil do not have valid values in some of their fields, such as "url".	It must be a URI. This is validated by regex.
7	Colombia	'accessURL' is a required property	1058 of the 3365 datasets for Colombia do not have the "accessURL" property.	The property should contain the access URL of the dataset distribution.
8	Colombia	'issued' is a required property	In the sample dataset from Colombia, none does not have the "issued" property.	The property should contain the publication date of the dataset.
9	Colombia	'downloadURL' is a required property	2308 dataset from Colombia does not have the "downloadURL" property.	URL that allows direct download of the dataset distribution, links directly to a downloadable file in a given format.

Table 2 is shown below, with the classification of aspects found grouped by type, some of the errors are expected because the repositories of these countries do not comply with the open data package of Argentina⁴⁶. It should be noted that the types of errors indicated in Table 2 are the validations used by the software prototype proposed by the recommendations suggested by

⁴⁴Perfil de Aplicación Nacional de Metadatos para Datos Abiertos. <https://datosgobar.github.io/paquete-apertura-datos/perfil-metadatos/>

⁴⁶Perfil de Aplicación Nacional de Metadatos para Datos Abiertos. <https://datosgobar.github.io/paquete-apertura-datos/perfil-metadatos/>

the good practice guides of Argentina⁴⁷. This country was taken as the standardized criterion since it was one of the countries with the most detailed validations.

It is important to mention that compared to the results obtained in Table 2, these are aspects that were identified through the good practice guide of the Argentine Government⁴⁸. This means that, although it is detected as an aspect to be considered by the prototype made, it might not be a criterion that does not take you into account in the countries mentioned and identified as a sample. Also shows a column of recommendations for the aspects detected in the structure of public open data catalogues.

4. Conclusions and Future Lines

Based on the work carried out on the study of the interoperability treatment of the different Latin American countries, it can be observed that the concept of open government is increasingly booming, which allows transparency in government management for all citizens. Being able to share data in an open format together with your open data catalog will help different companies consume this data and work on it to analyze possible improvement scenarios in different citizens and their relevant contexts.

Collaborating with an adequate structure in open data catalogs will help improve the quality of interoperability management between the different software that share data so that they can be processed and thus obtain information with added value in citizen contributions. This work not only proposes a survey of the Latin American countries in aspects of the current state of interoperability but also proposes a validation prototype of the structure of these metadata of the data sets, that is, of the data catalogs. public open. These validations in these data structures are necessary to provide a higher quality in the data content that is available to all people.

In future lines, it is expected to increase the survey of countries in the context of open data interoperability, and on the other hand, concerning the programmed prototype, work will continue, adding new schemes that allow the management of the metadata of the different sets of data that are available on the government sites of the different governments at an international level. As another future proposal, possible aspects that could be standardized and that are common criteria between the different open data catalogs will be analyzed, this will allow a better reading, analysis, and interoperability between the different software, which will lead to openness of data and transparency in the management of public governments.

References

- [1] A. Maldonado Lozano, L. Paredes Aguilar, G. Alvarado, Open government management in public institutions: a review of the literature, *Sapienza: International Journal of Interdisciplinary Studies* (2021).

⁴⁷Perfil de Aplicación Nacional de Metadatos para Datos Abiertos. <https://datosgobar.github.io/paquete-apertura-datos/perfil-metadatos/>

⁴⁸Perfil de Aplicación Nacional de Metadatos para Datos Abiertos. <https://datosgobar.github.io/paquete-apertura-datos/perfil-metadatos/>

- [2] J. D. Twizeyimana, A. Andersson, The public value of e-government—a literature review, *Government information quarterly* 36 (2019) 167–178.
- [3] A. Naser, A. Williner, C. Sandoval, Participación ciudadana en los asuntos públicos: un elemento estratégico para la agenda 2030 y el gobierno abierto, 2021.
- [4] L. F. Luna-Reyes, M. M. Najafabadi, The us open data initiative: The road ahead, *Information Polity* 24 (2019) 163–182.
- [5] M. R. Martínez, Métricas de calidad para validar los conjuntos de datos abiertos públicos gubernamentales, Ph.D. thesis, Universidad Nacional de La Plata, 2022.
- [6] S. Cadena-Vela, Marco de referencia para la publicación de datos abiertos comprensibles basado en estándares de calidad, 2019.
- [7] G. Eletrônico, Ministério do planejamento, orçamento e gestão, secretaria de logística e tecnologia da informação; ministério da educação, secretaria de educação profissional e tecnológica: e-mag versão 3.0-modelo de acessibilidade em governo eletrônico, 2011.
- [8] A. Moser, L. Ribeiro, M. Silva, J. Olivera, Centralización y democratización de datos en américa latina: Un estudio de caso de brasil, 2023.
- [9] A. Lopez, C. J.I, Sistema nacional de interoperabilidade: o caso brasileiro., 2019.
- [10] M. Janssen, Y. Charalabidis, A. Zuiderwijk, Benefits, adoption barriers and myths of open data and open government, *Information systems management* 29 (2012) 258–268.
- [11] E. Styryn, L. F. Luna-Reyes, T. M. Harrison, Open data ecosystems: an international comparison, *Transforming Government: People, Process and Policy* 11 (2017) 132–156.
- [12] A. Perez, M. Rodriguez, Transformación digital en la administración pública: El caso de méxico, 2014.
- [13] J. Ramirez, P. Lopez, La plataforma digital nacional: Un paso hacia la modernización., 2022.
- [14] L. Garcia, R. Lopez, Interoperabilidad en sistemas gubernamentales: El reto del sni., 2016.
- [15] H. Rodriguez, L. Mendoza, Datos abiertos en méxico: Retos y oportunidades, 2020.
- [16] M. A. O. Sanabria, F. O. A. Fernández, M. P. G. Zabala, Colombian case study for the analysis of open data government: A data quality approach, in: *Proceedings of the 11th international conference on theory and practice of electronic governance*, 2018, pp. 389–394.
- [17] R. Martínez, R. Rodríguez, P. Vera, Analysis of datasets and catalogs in government open portals of the argentine republic, in: *2020 IEEE Congreso Bienal de Argentina (ARGENCON)*, IEEE, 2020, pp. 1–8.
- [18] R. e. a. Martinez, Analysis of interoperability in the argentine republic. proposal: Prototype for the study of open data catalogs, 2022.