

Criteria for the identification of ineffective open data portals: pretender open data portals

Alberto Abella; Marta Ortiz-de-Urbina-Criado; Carmen De-Pablos-Heredero

Nota: Este artículo se puede leer en español en:
<https://revista.profesionaldelainformacion.com/index.php/EPI/article/view/86553>

How to cite this article:

Abella, Alberto; Ortiz-de-Urbina-Criado, Marta; De-Pablos-Heredero, Carmen (2022). "Criteria for the identification of ineffective open data portals: pretender open data portals". *Profesional de la información*, v. 31, n. 1, e310111.

<https://doi.org/10.3145/epi.2022.ene.11>

Manuscript received on June 5th 2021
Accepted on January 6th 2022



Alberto Abella

<https://orcid.org/0000-0003-0243-4606>

Fiware Foundation
Franklinstrasse 13A
10587 Berlin Germany
alberto.abella@fiware.org



Marta Ortiz-de-Urbina-Criado

<https://orcid.org/0000-0001-7527-6798>

Universidad Rey Juan Carlos
Facultad de Ciencias Jurídicas y Sociales
Paseo de los Artilleros, s/n.
28032 Madrid, Spain
marta.ortizdeurbina@urjc.es ✉



Carmen De-Pablos-Heredero

<https://orcid.org/0000-0003-0457-3730>

Universidad Rey Juan Carlos
Facultad de Ciencias Jurídicas y Sociales
Paseo de los Artilleros, s/n.
28032 Madrid, Spain
carmen.depablos@urjc.es

Abstract

Open data are considered an essential resource for governments, businesses, and citizens. In that context, open data portals have potential for creating enormous economic growth. Open data portals should allow the reuse of open data, ensure the efficiency of data transmission, and enable professional initiatives based on data reuse. However, there are portals that are inefficient because they do not allow reuse of their data. The objective of this work is to define and identify open data portals that do not offer the possibility for professional reuse of their data. We refer to them as "pretender open data portals". The following research questions are considered herein: What minimum criteria must an open data portal satisfy to enable professional reuse of open data? How can portals that do not meet these criteria be identified? And, what problems might these portals present, and how they might be solved? The results of an analysis of two samples of open data portals in Spain reveal that 63.8% and 56.1% of the portals analyzed in 2019 and 2021, respectively, can be considered pretender open data portals. The existence of pretender open data portals can have negative economic and social impacts, such as wasting public resources and projecting a negative image of the government's open data policies. To find coordination mechanisms to develop open data portals that, through the professional reuse of their data can create economic and social value, is one important challenge. The analysis of best practices of open data portals can be also a way to go in deep in the understanding of open data reuse impact not only from a professional standpoint.

Keywords

Open data; Open government; Open data portals; Open data reuse; Pretender open data portals; Apparent reusability; Spain.

Funding

This paper has been supported by Project RTI2018-097447-B-I00 of the *Ministry of Science, Innovation and Universities* (Spain), and the *Bridge Project V946: The contribution of dialogic practices to the quality of teamwork: a proposal for evolving the relational coordination model*, of the “Openinnova” high performance research group (number 381, *University Rey Juan Carlos*), and by *Fireware Foundation*.

1. Introduction

Open data are an essential resource for governments, businesses, and citizens (*European Data Portal*, 2019). Governments at different levels have used open data to promote transparency, openness, and accountability (**Curto-Rodríguez**, 2020; **Veljković**; **Bogdanović-Dinić**; **Stoimenov**, 2014; **Zuiderwijk**; **Shinde**; **Janssen**, 2019). Some authors have stated that one way to increase awareness of open government is to collect datasets and their metadata and publish them on open data portals (**Máchová**; **Hub**; **Lnenicka**, 2018) without cost to potential users (**Dawes**; **Vidiasova**; **Parkhimovich**, 2016).

In that sense, open data portals have potential to deliver enormous economic growth (**Borzacchiello**; **Craglia**, 2012); but for that, they must enable data accessibility, discoverability, and reusability (**Máchová**; **Hub**; **Lnenicka**, 2018; **Royo-Montañés**; **Benítez-Gómez**, 2019). **García-García** and **Curto-Rodríguez** (2018) point out that, currently, open data portals are being systematically replicated by governments as a sign of modernity and demonstrate their legitimacy against political rivals. Sometimes, open data portals are built in the interest of presenting data in a particular fashion, underplaying the requirements of potential data reusers. In that line, **García-García** and **Curto-Rodríguez** (2019) define the concept of “apparent transparency,” but they do not consider in this concept another feature of open data portals, the data reuse.

The *Report on data reuse in Spain II* (**Abella et al.**, 2019) presents an analysis of the features of data portals in Spain and confirms the increase in the number of open data portals. A significant percentage of these new portals present limited features to allow reuse of data. These portals are inefficient because they do not allow a satisfying reuse of data, in special, the professional reuse. It thus becomes necessary to identify such portals and offer solutions to improve them and prevent the creation of new portals with these characteristics.

Professional reuse is one of the types of data reuse oriented to for-profit reusers aimed to enrich their products and services, provide new ones or develop innovative business models (**Abella**; **Ortiz-de-Urbina-Criado**; **De-Pablos-Heredero**, 2019a; **Jetzek**; **Avital**; **Bjorn-Andersen**, 2019). Focusing on professional data reuse, the first objective of this work is to define and identify open data portals that do not offer the possibility for professional reuse of their data. These will be referred to as “pretender open data portals” (PODPs). In that sense, we propose two research questions:

- 1) What minimum criteria should an open data portal meet to enable professional reuse of the data it publishes?
- 2) How can be identified portals that do not meet these criteria?

The second objective is to analyze the PODPs economic and social effects and propose some solutions. The following research question is proposed:

- 3) What problems could they cause, and how they might be solved?

This paper is organized as follows: After this introduction, the concept and context of an open data portal are described, followed by a discussion of the reuse of their data. Based on this analysis, PODPs are defined and a methodology to identify them is proposed. In the third section, the developed methodology is applied to a large sample of open data portals to identify PODPs in Spain in 2019 and 2021. The conclusions section reflects on the problems caused by these portals and offers solutions. Finally, the contributions and limitations of this study and future lines of research are described.

2. Open data portals**2.1. Concept and context**

Open data portals are digital platforms that are used to store, share, connect, and visualize open data (**Zuiderwijk**; **Shinde**; **Janssen**, 2019). **Máchová**, **Hub** and **Lnenicka** explain that

“open data portals provide information about datasets in form of a description with metadata and allow for direct access to datasets via download and additionally via application programming interface (API)” (**Máchová**; **Hub**; **Lnenicka**, (2018).

Some open data manager systems are based on open-source standards such as the *Comprehensive knowledge archive network* (CKAN) or *Drupal knowledge archive network* (DKAN). Other features can be provided by such portals for automatic and massive retrieval of data like APIs and semantic query resources to retrieve data (Sparql). They also can provide discussion and comment features, and data visualization tools (**Máchová**; **Hub**; **Lnenicka**, 2018; **Máchová**; **Lnenicka**, 2017).

Previous literature has analyzed various aspects of open data portals. The following is a summary of the issues that have been addressed to provide a picture of which aspects have been studied. Some authors have analyzed the features and content of open data portals (**Aful-Dadzie; Aful-Dadzie, 2017; Curto-Rodríguez, 2021; Thorsby et al., 2017; Sheffer-Correa; Mendes-de-Souza; Soares-Correa-da-Silva, 2019**), and their services and characteristics (**Chatfield; Reddick, 2017; Royo-Montañés; Benítez-Gómez, 2019**). Others have analyzed the quality of the published data and their metadata –completeness, consistency in data types, proper metadata available– (**Neumaier; Umbrich; Polleres, 2016; Máchová; Lnenicka, 2017; Kubler et al., 2018**). In addition, some papers have focused on the reusability of datasets published on open data portals at the level of Europe (**Ramos-Simón et al., 2012**), Spain (**Curto-Rodríguez, 2020; Máchová; Hub; Lnenicka, 2018**), or Spanish autonomous communities (**Vicente-Paños; Jordán-Alfonso, 2017**). However, more research is needed

Open data portals publish data to promote openness, transparency, and create industrial value

“to identify and define methodologies and best practices for open data publication” (**Máchová; Hub; Lnenicka, 2018**)

and to analyze the features that open data portals should include to enable data reuse (**Abella; Ortiz-de-Urbina-Criado; De-Pablos-Heredero, 2019b**).

Open data portals can enhance transparency (**Curto-Rodríguez, 2020**) and can enable the accountability and dissemination of open data (**García-García; Curto-Rodríguez, 2018; 2019; González-Limón; Rodríguez-Ramos, 2019**), as well as the reuse of open data of social or economic value (**Royo-Montañés; Benítez-Gómez, 2019**). However, Sheffer-Correa, Mendes-de-Souza and Soares-Correa-da-Silva state that

“many data portals, particularly those of local governments, appear to be poorly implemented and developed with the classic website model in mind, which provides access to data only through user interaction with web forms” (**Sheffer-Correa; Mendes-de-Souza; Soares-Correa-da-Silva, 2019**).

Sometimes, the government’s interest is in presenting data in a particular fashion to achieve political visibility, thereby limiting the use and provision of data to stakeholders who are interested in data reuse (**Beltrán-Orenes; Rodríguez-Mateos, 2020**).

The mere creation of open data portals does not guarantee successful promotion of transparency or greater government accountability (**Attard et al., 2015; Royo-Montañés; Benítez-Gómez, 2019**). **Yu and Robinson (2012)** refer to the concept of open government ambiguity, which suggests that transparency depends not only on the technical characteristics (licenses and formats) of open government rather than providing real content of value to different stakeholders. In that sense, relational coordination theory (**Gittel; Seidner; Wimbush, 2010**) explains that, whenever shared objectives, mutual respect, and efficient communication mechanisms exist among the different agents that benefit from a process, more value is created (**Haider; Fernández-Ortiz; De-Pablos-Heredero, 2017**). **Morozov (2013)** stresses the idea of technology solutionism as the incorrect belief that technology itself provides the power to promote transparency. **García-García and Curto-Rodríguez (2019)** identify two characteristics of open data portals: “sincere transparency” and “apparent transparency” (exhibited by portals that offer no or almost no data related to transparency and accountability). In addition to transparency and accountability, the reuse of data published on open data portals becomes another relevant aspect to be analyzed. Extending the proposal of **García-García and Curto-Rodríguez (2019)** that defines the concept of “apparent transparency,” we identify other attribute of open data portals: the professional data reuse. Then, our focus is on portals that exhibit professional “apparent reusability”.

The aim of an open data portal is thus to disseminate datasets in reusable formats because making the data available and even publishing data in reusable formats do not guarantee usage (**Gebre; Morales, 2020**). **Attard et al.** stated that

“the portal’s ‘success’ would not only be evaluated on the amount of data published, but also on the usability of this data” (**Attard et al. 2015**).

But, the reuse of open data can be carried out by internal and external reusers (**Abella; Ortiz-de-Urbina-Criado; De-Pablos-Heredero, 2019a**). The internal group includes open data publishers, usually the public administration, as well as other public administrations. The external reusers can be grouped into several categories:

- citizens;
- professional reusers (for-profit entities using the data to enrich or innovate products and services) and their clients;
- non-profit or third-sector organizations;
- academic reusers.

Then, an open data portal can thus be accessed by any people or organizations from around the world, who may have very different interests. Within this ecosystem, some actors will require professional reuse of open data, usually for the creation of new products and/or services or to improve existing ones (**Abella; Ortiz-de-Urbina-Criado; De-Pablos-Heredero, 2019a**). To develop such professional data reuse and enhance entrepreneurship, reusers must have adequate access to the data. If data are not published on the portal in a truly reusable way, inefficiencies will occur such as wastage

of resources because the creation of portals that do not enable one of their objectives, i.e., data reuse of its data, cannot be achieved.

“Pretender open data portals (PODP) do not allow the professional reuse of data”

However, the creation of open data portals does not mean that the data they publish are prepared for professional reuse. Organizations that provide open data portals should note that one of the values provided by data lies in its ability to be reused and should thus try to define and create open portals whose features enable adequate data reuse (Ferrer-Sapena; Peset; Alexandre-Benavent, 2011; Abella; Ortiz-de-Urbina-Criado; De-Pablos-Heredero, 2014; 2015; 2018; 2019b; Royo-Montañés; Benítez-Gómez, 2019). For reusing open data, some aspects have to be considered. One of them is the legal framework that allows its use (Abella; Ortiz-de-Urbina-Criado; De-Pablos-Heredero, 2014). For example, some regulations of the reuse of public-sector information are based on the EU *Directive 2019/1024 (European Union, 2019)* and Spanish *Law 37/2007 (España, 2007)*. However, there is not a legal framework for private open data unless they are providing public services. In addition, information is reusable when there are no technical barriers to reuse, the information can be accessed automatically, and the structure of the published information can be accessed (Abella; Ortiz-de-Urbina-Criado; De-Pablos-Heredero, 2014).

2.2. Pretender open data portals

Our focus here is on portals that exhibit professional “apparent reusability”. Professional use of an open data portal can enable the development of entrepreneurial initiatives oriented towards the creation of social and economic value (Abella; Ortiz-de-Urbina-Criado; De-Pablos-Heredero, 2019a). Considering professional use of an open data portal, two classes of portals can be found: those focused on data reuse, and those created with other interests. Therefore, an open data portal is pretender when it does not allow professional reuse of data.

So, how can we identify the criteria for qualifying an open data portal as pretender? For that, we consider the reports of Abella; Ortiz-de-Urbina-Criado; De-Pablos-Heredero (2017) and Abella *et al.* (2019). Various features are identified to assess the degree of maturity of a data portal:

- the availability of a catalog of resources (datasets) and whether the catalog was downloadable;
- a mechanism for announcing dataset updates;
- the use of a data management system (DMS) tool for data publication;
- the existence of mechanisms allowing automated data access (API or query endpoints);
- the availability of a section on the site that identifies and lists open data-driven services for reuse of the data published by the portal.

These five elements represent an adaptation of those presented in the methodology described by Carrara; Nieuwenhuis; Vollers (2016) to assess the maturity of open data in the European context based on three dimensions: data dissemination, portal usability, and data reusability.

Abella *et al.* (2019) concluded that some portals do not meet the minimum requirements for the data they publish to be reusable. An academic reflection process was then carried out to identify which of the five criteria a portal should meet as a minimum to not exhibit “apparent reusability.” Therefore, to be considered as an open data portal enabling professional reuse, a portal must meet at least three of the five criteria:

- provision of an updating mechanism (enabling the delivery of real-time information on data updates);
- a DMS (for providing automated access to data);
- API availability (for publishing data in a format that can be reused professionally).

The other two criteria -the number of datasets published and the availability of data-driven services- are not included because portals that have just been created may have not reached yet the minimum number of datasets or developed data-driven services. This might indicate that they are simply recently being introduced.

The first criterion is the provision of a mechanism to store information on data updates, to enable users to get real-time information without retrieving constantly the data to detect changes. The importance of updating was analyzed by Atz (2014), who found that many of the datasets were not updated regularly (thus lacking timeliness). In the same way, Smith, Ofe and Sandberg (2016) refer to support services, which include:

- documentation and code examples;
- a support forum with tier 1 support -the direct contact with the user-;
- the operational status of APIs;
- collection of usage statistics and user requests;
- data quality verification;
- information on upcoming updates.

The absence of such support hinders the use of the data by either users or publishers. This implies that, to obtain an updated version of the data, a user should visit a website, for example, every hour, to see whether the

“PODP lack of updating news, automated access to data and DMS use”

data have been updated, or possibly only once a day. In addition, this also represents a waste of resources for publishers, as it forces them to answer requests when they cannot provide updates on the publication of new data.

“ An analysis on a sample of 279 Spanish portals at national and local level is done ”

The second criterion that is necessary for the systematic publication of data is the provision of automated access to internal data sources. A DMS enables external users the automated access to publish data, but also the almost complete automation of internal data capture and external publication. Therefore, it allows the systematic publication of data (Sadiq; Indulska, 2017) in terms of content and data updating, and with a sustainable use of resources. Many PODPs are up to date on their launch but the information becomes obsolete a short time thereafter because of the high cost of manual updating. Failure to update can result in a loss of confidence in the publisher (Atz, 2014).

The third criterion is that data accessible to professional users should be automatically accessible through an API (or equivalent) in a format that allows integration with other digital services or applications that make use of published data (Groth *et al.*, 2014). Smith; Ofe and Sandberg (2016) refer to the importance of providing a management system that can handle requests in an automated way as well as collect statistics. They also explain that the website should include descriptions and documentation for the APIs, their operational status, and access to support services as well as a news section, project and member catalogs, and links to other API sources. Usually, these APIs also allow access to the fraction of the dataset required by the user rather than the full dataset.

In conclusion, we define “pretender open data portals” (PODPs) as those that do not meet the following minimum criteria for professional reuse of the information:

- the capability of providing real-time information through data updates;
- enabling automated publication of data via a DMS;
- providing automated access to data in a format such that they can be reused through an API.

3. Identifying pretender open data portals (PODP) in Spain

According to a 2019 report from the pan-European open data portal (Cecconi *et al.*, 2019), Spain went from being a leader in 2016 to the fifth position in 2019 in terms of open data policies, behind France, Denmark, Ireland, and Poland. Conversely, it remains a leader in terms of impact, in this case ahead of France and Ireland, whereas in terms of the maturity of data portals, it went up one position from third in 2016 to second in 2019, lying just behind France and ahead of Ireland. The *European Commission* (Van-Knippenberg, 2020) updated this report on open data maturity in 2020, revealing that Ireland, Spain, and France maintained their dominant position in open data management, being joined by Denmark, Estonia, Poland, and Austria. Despite the effort and investment made in the launch and operation of open data portals in Spain, we note that many have appeared in recent years, but some are barely used (Abella *et al.*, 2019).

3.1. Sample of open data portals

Two samples of open data portals in Spain were selected: one from 2019, the year before the start of the pandemic, and the other from 2021, the year after the lockdown in Spain due to the pandemic.

For the 2019 sample, the open data portals listed in the national Spanish open data portal were used. <http://datos.gob.es>

Of the 295 portals identified, 16 were not active or the data portal was not available (“404 error”). The remaining 279 were analyzed manually. In 2021, 330 open data portals were identified from the following sources: portals included in the report by Abella *et al.* (2019), data from datos.gob.es, *Consorti d'Administració Oberta de Catalunya*, and additional exploration by the authors. These were analyzed manually, revealing that 50 of them were not active or the portal was not available, leaving 280 open data portals for consideration. A description of the type of entity and the scope of the sampled open data portals is presented in Table 1.

Table 1. Entity type and scope of open data portals

Entity type	Scope	2019	2021
National organization	National	7	4
National agency	National	30	22
Regional government	Regional	15	15
Regional agency	Regional	5	7
Diputación (county)	Local	12	10
Cabildo (county)	Local	3	5
Municipality	Local	196	205
Local entity	Local	3	3
University	Others	8	9

3.2. Results

Table 2 presents a descriptive analysis of the final sample in terms of each criterion for professional reuse: updating mechanism, DMS use, and API availability.

Table 2. Criteria for professional reuse of open data

Criteria		2019		2021	
		Absolute frequency	Percentage	Absolute frequency	Percentage
Updates noticing mechanism	No	231	82.8	230	82.1
	Yes	48	17.2	50	17.9
Use of DMS	No	200	71.7	196	70.0
	Yes	79	28.3	84	30.0
Availability of an API	No	186	66.7	176	62.9
	Yes	93	33.3	104	37.1

The results presented in Table 2 show that, in 2019, only 17.2% of the open data portals had a source with data updates, 28.3% of them used a DMS, and 33.3% offered an API. In 2021, these figures remained similar (17.9%, 30%, and 37.1%).

We then identified from the sample those portals that did not meet the three proposed criteria and can thus be considered PODPs. In 2019, it was observed that 178 of the portals could be considered pretenders, representing 63.8% of the sample, whereas in 2021, 157 (56.1%) of the portals analyzed can be considered as pretenders (Table 3).

Table 3. Pretender open data portals 2019 and 2021

Spanish autonomous communities and cities	2019				2021			
	Reuse portals	PODP	Total	% PODP	Reuse portals	PODP	Total	% PODP
Andalusia	8	13	21	61.9	10	6	16	37.5
Aragon	5	0	5	0.0	5	0	5	0.0
Asturias	2	1	3	33.3	1	1	2	50.0
Balearic Islands	1	1	2	50.0	1	2	3	66.7
Basque Country	11	7	18	38.9	7	3	10	30.0
Canary Islands	9	4	13	30.8	13	3	16	18.8
Cantabria	2	1	3	33.3	2	1	3	33.3
Castilla and León	1	2	3	66.7	3	1	4	25.0
Castilla La Mancha	1	6	7	85.7	0	4	4	100
Catalonia	35	4	39	10.3	47	2	49	4.1
Ceuta	0	1	1	100	0	0	0	0.0
Community of Madrid	7	3	10	30.0	9	3	12	25.0
Community of Navarre	1	1	2	50.0	2	0	2	0.0
Extremadura	3	104	107	97.2	4	111	115	96.5
Galicia	1	2	3	66.7	3	0	3	0.0
Melilla	0	1	1	100	0	0	0	0.0
Region of Murcia	1	1	2	50.0	2	2	4	50.0
Rioja	0	1	1	100	1	0	1	0.0
Valencian Community	6	1	7	14.3	6	3	9	33.3
National portals ¹	7	24	31	77.4	7	15	22	68.2
Total	101	178	279	63.8	123	157	280	56.1

The results presented in Table 3 show that, in some cases, the percentage of PODPs was 100%, as in Ceuta, Melilla, and Rioja in 2019, while in 2021 only Castilla-La Mancha showed this percentage. Extremadura, with more than 100 portals, showed 97.2% PODPs in 2019 and 96.5% in 2021.

Other regions presenting percentages above 50% in 2019 were Castilla-La Mancha (85.7%), Castilla and León (66.7%), Galicia (66.7%), and Andalusia (61.9%), but only the Balearic Islands (66.7%) in 2021. Also, the vast majority of sites with national scope (77.4% in 2019, 68.2% in 2021) can be classified as PODPs.

The Community of Madrid (30%), the Canary Islands (30.8%), Asturias (33.3%), Cantabria (33.3%), and the Basque Country (38.9%) present percentages between 30% and 40% in 2019. Aragon (0%), Catalonia (10.3%), the Valencian Community (14.3%), the Community of Madrid (30%), the Canary Islands (30.8%) and are the communities exhibiting the lower

percentages of PODPs in 2019. However, in 2021, five autonomous communities and cities showed no PODPs: Aragon, Navarre, Rioja, Ceuta, and Melilla, while the others showed between 4% and 50% PODPs.

“ In Spain, a 63.8% of national and local open data portals are PODP ”

3.3. Discussion

The results of this analysis of open data portals in Spain reveals that 63.8% and 56.1% of the portals analyzed were PODPs in 2019 and 2021, respectively. The results of this research can help identify the features that an open data portal should include to allow data reuse and raise awareness among developers about the importance of redesigning such portals to orient them toward value creation. Moreover, one should note that a DMS can provide services not only to external users by publishing open data through open data portals but also within the organization by enabling sharing of data through a shared data portal. These shared data portals need not provide open data to the public, although they usually do it. The managers of these portals should invest effort in creating value-added services to make the data more understandable and accessible (**Ferrer-Sapena; Peset; Alexandre-Benavent**, 2011).

Shared data portals enable the publication of data that are open to external users as well as for internal use. Then, data can be reused by users who have the appropriate access permission. In fact, **Abella et al.** (2019) found that the widest reuse of open data is by users from the organization itself. Such shared data portals may open up further possibilities for data reuse by agents both internal and external to the organization.

In addition, we reached similar conclusions as **Thorsby et al.** (2017), **Royo-Montañés** and **Benítez-Gómez** (2019) and **Wang, Chen** and **Richards** (2018) in that the development of open data portals at the local government level remains at an early stage. Also, **Janssen** (2011) and **Conradie** and **Choenni** (2014) support that, in comparison with national governments, many local governments are currently facing unknown tasks and are not fully ready to adopt open data.

During the last two years, little improvement has been seen, despite the great digital revolution resulting from the pandemic. The Covid-19 crisis has changed the routines of citizens and organizations. The imposition of lockdowns, avoidance of face-to-face meetings, and closure of the public offices responsible for receiving and processing documents due to the pandemic have revealed all too clearly the importance of having up-to-date data available in real time for decision-making (**Cetina**, 2021).

4. Conclusions

4.1. PODPs: problems and solutions

Open government data offer important advantages such as political, social, economic, operational, and technical benefits to government itself through data reuse (**Dawes; Vidiasova; Parkhimovich**, 2016). When data can be accessed and reused, portals are perceived as more useful and are better valued. However, some open data portals do not enable such gains from reusability to be achieved. This work identified open data portals that are not created to allow professional reuse of the data published in them. Such portals can have negative economic and social impacts.

An important element for measuring the success of reusability is to compare results regarding the access and use of data, and the vitality of the ecosystem responsible for generating applications and the resulting services. Users can only develop sustainable initiatives when they add value to the published data by including new elements or performing processing on the data retrieved. The cost of data retrieval and update therefore impacts directly on the sustainability of such initiatives. However, this does not result in fees being charged to access the information, which generally remains free. Several reports have been published on this aspect, stressing the beneficial aspects of price suppression (**Ramos-Simón et al.**, 2012). Most countries have moved towards a marginal cost scheme in which only the costs of reproducing the information are charged, which has substantially increased the number of reusers (**Ramos-Simón et al.**, 2012). However, this could extend to costs related to obtaining information through the API of the portal, understanding the data models through which the information is presented, transformation of this information into standardized data models (**Abella et al.**, 2019), exploration to determine when new data are published, or cleaning of incorrect data, for example, that is out of range, of the wrong type, etc. (**Gudivada; Apon; Ding**, 2017). Considering the public investments required to create open data portals, PODPs do not achieve the objectives of reuse and do not allow optimal use of public resources, resulting in inefficiencies.

Diez-Rico and **Millán-Calenti** (2011) explained that an open data portal is a showcase for public transparency, having a low initial cost and being relatively easy to build. However, behind this lies poorly structured data and a supply of information that is inconsistent in terms of both quality and quantity, in relation to the available supply of public information. In this sense, in many cases, the deployment of organizational routines aimed at developing an ability to extract full value from the data, especially open data, has not been established in the public sector or is not the aim of such activity. The data available via open data portals and their format are mostly not useful to agents who need to reuse such data to create value. In this sense, PODPs can have a negative impact on other agents further than professional reusers, as for example, policymakers and citizens.

The development of PODPs also can reveal a lack of relational coordination (**Haider; Fernández-Ortiz; De-Pablos-Heredero**, 2017), which requires a capacity of coordination among organizations that maintain data portals and the professional reusers. As modeled by **Gittel; Seidner** and **Wimbush** (2010), and **Lcayo-Mendoza** and **De-Pablos-Heredero**

(2016), it would be desirable to develop organizational mechanisms to establish frequent, timely and relevant communication oriented to solve problems. It would

also be useful to share goals, knowledge, and practice mutual respect among government and professional reusers. Regarding open data portals, the potential of taking care of relational and communication ties would become a promising future line of research, under the assumption that an increase in relational coordination will favor an improvement in professional data reuse. The implementation of good practices by open data portal creators is another future interesting area of research that will help to ensure the efficient use of open data portals.

Results claim for coordination efforts to build PODP oriented to create value

4.2. Theoretical and practical contributions

This paper makes various theoretical and practical contributions:

1. The concept of pretender data portals (PODPs) is defined.
2. A process to identify PODPs is constructed.
3. The process is validated using a large sample of open data portals from 2019 and 2021.
4. A proposal for governments and organizations oriented toward using a relational coordination model to improve the setup of open data portals is provided.

4.3. Limitations and future research lines

This study presents the concept and the characterization of PODPs, but this proposal suffers from several limitations. On the one hand, it would be interesting to develop this concept by using qualitative methodologies such as the Delphi method. On the other hand, this study is limited to open data portals in Spain and its focus on just professional reuse, so it would be useful to extend the scope of this concept to include other countries and compare the results.

In addition, it would be interesting to study whether portals that were pretenders in 2019 and 2021 have disappeared or ceased to be inefficient. Additionally, it would be interesting to identify which portals can be considered to have implemented best practices, and even rank the portals to promote a virtuous circle of improvement. Both activities could provide examples to avoid the creation and spread of more inefficient data portals.

For future studies, the analysis of best practices of open data portals can be a way to go in deep in the improvement of the understanding of open data reuse not only from a professional standpoint. Best practices could also inspire policymakers to build more effective open data portals. Other studies could find additional empirical evidence and options to provide open data portals that allow an efficient and sustainable use of resources. Moreover, it would be interesting to delve deeper into knowledge on the relational coordination model, focusing on the sustainability of open data portals and its relationship with reuser satisfaction (Martin *et al.*, 2019; Welle-Donker; Van-Loenen, 2017). Some interesting questions in this regard are:

- Do open data builders maintain effective communication with potential data users? How often?
- Do they try to understand the main reasons for the use of the open data portals that they create?
- Are the objectives of both parties, i.e., creators and users, aligned?
- Can open data portals be sustainable if they do not serve the main objectives of reusers?

5. Note

1. The ports are administrative national entities (counted in Table 1), but they have been assigned to their regional location in table 3 which explains the difference in the amount of national portals.

6. References

- Abella, Alberto; Ortiz-de-Urbina-Criado, Marta; De-Pablos-Heredero, Carmen** (2014). "Meloda, métrica para evaluar la reutilización de datos abiertos". *El profesional de la información*, v. 23, n. 6, pp. 582-588.
<https://doi.org/10.3145/epi.2014.nov.04>
- Abella, Alberto; Ortiz-de-Urbina-Criado, Marta; De-Pablos-Heredero, Carmen** (2015). "Information reuse in smart cities' ecosystems". *El profesional de la información*, v. 24, n. 6, pp. 831-837.
<https://doi.org/10.3145/epi.2015.nov.16>
- Abella, Alberto; Ortiz-de-Urbina-Criado, Marta; De-Pablos-Heredero, Carmen** (2017). *La reutilización de datos abiertos: una oportunidad para España*. Madrid: Cotec. ISBN: 978 84 929333 8 9
<https://datos.gob.es/es/noticia/la-reutilizacion-de-datos-abiertos-una-oportunidad-para-espana-informe-cotec>
- Abella, Alberto; Ortiz-de-Urbina-Criado, Marta; De-Pablos-Heredero, Carmen** (2018). "Indicadores de calidad de datos abiertos: el caso del portal de datos abiertos de Barcelona". *El profesional de la información*, v. 27, n. 2, pp. 1699-2407.
<https://doi.org/10.3145/epi.2018.mar.16>
- Abella, Alberto; Ortiz-de-Urbina-Criado, Marta; De-Pablos-Heredero, Carmen** (2019a). "The process of open data publication and reuse". *Journal of the Association for Information Science and Technology*, v. 70, n. 3, pp. 296-300.
<https://doi.org/10.1002/asi.24116>

- Abella, Alberto; Ortiz-de-Urbina-Criado, Marta; De-Pablos-Heredero, Carmen** (2019b). "Meloda 5: a metric to assess open data reusability". *El profesional de la información*, v. 28, n. 6, e280620.
<https://doi.org/10.3145/epi.2019.nov.20>
- Abella, Alberto; Ortiz-de-Urbina-Criado, Marta; De-Pablos-Heredero, Carmen; Vidal-Cabo, Christian; Ferrer-Sapena, Antonia** (2019). *La reutilización de datos abiertos en España II*. Barcelona: DesideDatum.
https://www.desidedatum.com/wp-content/uploads/2019/11/La_reutilizacio%CC%81n_datos_abiertos_espan%CC%83a_2019.pdf
- Afful-Dadzie, Eric; Afful-Dadzie, Anthony** (2017). "Open government data in Africa: a preference elicitation analysis of media practitioners". *Government information quarterly*, v. 34, n. 2, pp. 244-255.
<https://doi.org/10.1016/j.giq.2017.02.005>
- Attard, Judie; Orlandi, Fabrizio; Scerri, Simon; Auer, Sören** (2015). "A systematic review of open government data initiatives". *Government information quarterly*, v. 32, n. 4, pp. 399-418.
<https://doi.org/10.1016/j.giq.2015.07.006>
- Atz, Ulrich** (2014). "The tau of data: a new metric to assess the timeliness of data in catalogues". In: Parycek, Peter; Edelmann, Noella (eds.). *Proceedings of CeDEM14 conference for e-democracy and open government*, pp. 257-267. Austria: Edition Donau-Universität Krems. ISBN: 978 3 902505 35 4
- Beltrán-Orenes, Pilar; Rodríguez-Mateos, David** (2020). "Presencia y ubicación de los portales de transparencia municipales españoles". *El profesional de la información*, v. 29, n. 6, e290635.
<https://doi.org/10.3145/epi.2020.nov.35>
- Borzacchiello, María-Teresa; Craglia, Max** (2012). "The impact on innovation of open access to spatial environmental information: a research strategy". *International journal of technology management*, v. 60, n. 1/2, pp. 114-129.
<https://doi.org/10.1504/IJTM.2012.049109>
- Carrara, Wendy; Nieuwenhuis, Margriet; Vollers, Heleen** (2016). *Open data maturity in Europe 2016. Insights into the European state of play*. European Commission. Directorate General for Communications Networks, Content and Technology. European Data Portal. ISBN: 978 92 78 42054 3
https://www.europeandataportal.eu/sites/default/files/edp_landscaping_insight_report_n2_2016.pdf
- Cecconi, Gianfranco; Blank, Marit; Radu, Cosmina; Linklaen-Arriëns, Eline; Huyer, Esther** (2019). *Open data maturity. Report 2019*. European Commission. Directorate General for Communications Networks, Content and Technology. European Data Portal. ISBN: 978 92 78 42052 9
https://www.europeandataportal.eu/sites/default/files/open_data_maturity_report_2019.pdf
- Cetina, Camilo** (2021). *Gobernanza de datos y capacidades estatales para la pos-pandemia*. Policy Brief #28. Caracas: Corporación Andina de Fomento.
<https://scioteca.caf.com/handle/123456789/1765>
- Chatfield, Akemi-Takeoka; Reddick, Christopher G.** (2017). "A longitudinal cross-sector analysis of open data portal service capability: the case of Australian local governments". *Government information quarterly*, v. 34, n. 2, pp. 231-243.
<https://doi.org/10.1016/j.giq.2017.02.004>
- Conradie, Peter; Choenni, Sunil** (2014). "On the barriers for local government releasing open data". *Government information quarterly*, v. 31, n. supplement 1, pp. S10-S17.
<https://doi.org/10.1016/j.giq.2014.01.003>
- Curto-Rodríguez, Ricardo** (2020). "Transparencia operativa de las comunidades autónomas españolas mediante sus portales de datos abiertos". *El profesional de la información*, v. 29, n. 1, e290115.
<https://doi.org/10.3145/epi.2020.ene.15>
- Curto-Rodríguez, Ricardo** (2021). "Análisis multidimensional de los portales de datos abiertos autonómicos españoles". *Revista española de documentación científica*, v. 44, n. 1, e284.
<https://doi.org/10.3989/redc.2021.1.1745>
- Dawes, Sharon S.; Vidasova, Lyudmila; Parkhimovich, Olga** (2016). "Planning and designing open government data programs: An ecosystem approach". *Government information quarterly*, v. 33, n. 1, pp. 15-27.
<https://doi.org/10.1016/j.giq.2016.01.003>
- Diez-Rico, Emilio-José; Millán-Calenti, Rafael-Álvaro** (2011). *Open data y RISP: generando valor social y económico. Transparencia e innovación en la eadministración*. Santiago de Compostela: Colexio Profesional de Enxeñaría en Informática de Galicia. ISBN: 978 84 615 5395 2
https://xornalgalicia.com/attachments/article/15241/expertoenprotecciondedato-millancalenti--2Open_Data_+_y_RISP.pdf

- España* (2007). “Ley 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público”. *Boletín oficial del estado*, n. 276, 17 noviembre, pp. 47160-47165.
<https://www.boe.es/eli/es/l/2007/11/16/37>
- European Data Portal* (2019).
<https://www.europeandataportal.eu>
- European Union* (2019). “Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the re-use of public sector information”. *Official journal of the European Union*, n. L 172/57 of 26/06/2019.
<http://data.europa.eu/eli/dir/2019/1024/oj>
- Ferrer-Sapena, Antonia; Peset, Fernanda; Aleixandre-Benavent, Rafael** (2011). “Acceso a los datos públicos y su reutilización: open data y open government”. *El profesional de la información*, v. 20, n. 3, pp. 260-269.
<https://doi.org/10.3145/epi.2011.may.03>
- García-García, Jesús; Curto-Rodríguez, Ricardo** (2018). “Divulgación de información pública de las comunidades autónomas españolas (2013-2017): portal de datos abiertos, portal de transparencia y web institucional”. *El profesional de la información*, v. 27, n. 5, pp. 1051-1060.
<https://doi.org/10.3145/epi.2018.sep.09>
- García-García, Jesús; Curto-Rodríguez, Ricardo** (2019). “El ejercicio de la rendición de cuentas mediante portales de datos abiertos en las comunidades autónomas españolas”. *IDP. Revista de internet, derecho y política*, n. 29.
<https://raco.cat/index.php/IDP/article/view/373562>
- Gebre, Engida H.; Morales, Esteban** (2020). “How ‘accessible’ is open data? Analysis of context-related information and users’ comments in open datasets”. *Information and learning sciences*, v. 121, n. 1/2, pp. 19-36.
<https://doi.org/10.1108/ILS-08-2019-0086>
- Gittel, Jody-Hoffer; Seidner, Rob; Wimbush, Julian** (2010). “A relational model of how high-performance work systems work”. *Organization science*, v. 21, n 2, pp. 490-506.
<https://doi.org/10.1287/orsc.1090.0446>
- González-Limón, Myriam; Rodríguez-Ramos, Asunción** (2019). “Revisión del cumplimiento de los datos abiertos por los ayuntamientos españoles en la iniciativa Aporta”. *Revista española de documentación científica*, v. 42, n. 4, e247.
<https://doi.org/10.3989/redc.2019.4.1605>
- Groth, Paul; Loizou, Antonis; Grayd, Alasdair J. G.; Goble, Carole; Harland, Lee; Pettifer, Steve** (2014). “API-centric linked data integration: the open phacts discovery platform case study”. *Journal of web semantics*, v. 29, pp. 12-18.
<https://doi.org/10.1016/j.websem.2014.03.003>
- Gudivada, Venkat; Apon, Amy; Ding, Junhua** (2017). “Data quality considerations for big data and machine learning: going beyond data cleaning and transformations”. *International journal on advances in software*, v. 10, n. 1, pp. 1-20.
https://www.ariajournals.org/software/soft_v10_n12_2017_paged.pdf
- Haider, Sajid; Fernández-Ortiz, Antonio; De-Pablos-Heredero, Carmen** (2017). “Organizational citizenship behavior and implementation of evidence-based practice: moderating role of senior management’s support”. *Health systems*, v. 6, pp. 226-241.
<https://doi.org/10.1057/s41306-017-0026-7>
- Janssen, Katleen** (2011). “The influence of the PSI directive on open government data: an overview of recent developments”. *Government information quarterly*, v. 28, n. 4, pp. 446-456.
<https://doi.org/10.1016/j.giq.2011.01.004>
- Jetzek, Thorhildur; Avital, Michel; Bjorn-Andersen, Niels** (2019). “The sustainable value of open government data”. *Journal of the Association for Information Systems*, v. 20, n. 6.
<https://doi.org/10.17705/1jais.00549>
- Kubler, Sylvain; Robert, Jérémy; Neumaier, Sebastian; Umbrich, Jürgen; Le-Traon, Yves** (2018). “Comparison of metadata quality in open data portals using the Analytic Hierarchy Process”. *Government information quarterly*, v. 35, n. 1, pp. 13-29.
<https://doi.org/10.1016/j.giq.2017.11.003>
- Lacayo-Mendoza, Alexander; De-Pablos-Heredero, Carmen** (2016). “Managing relationships and communications efficiently through digital social networks in higher education: the importance of the relational coordination model”. *Dyna-Colombia*, v. 83, n. 195, pp. 138-146.
<https://doi.org/10.15446/dyna.v83n195.49296>
- Máchová, Renáta; Hub, Miloslav; Lnenicka, Martin** (2018). “Usability evaluation of open data portals: evaluating data discoverability, accessibility, and reusability from a stakeholders’ perspective”. *Aslib journal of information management*, v. 70, n. 3, pp. 252-268.
<https://doi.org/10.1108/AJIM-02-2018-0026>

- Máchová, Renáta; Lnenicka, Martin** (2017). "Evaluating the quality of open data portals on the national level". *Journal of theoretical and applied electronic commerce research*, v. 12, n. 1, pp. 21-41.
<https://doi.org/10.4067/S0718-18762017000100003>
- Martin, Christopher; Evans, James; Karvonen, Andrew; Paskaleva, Krassimira; Yang, Dujuan; Linjordet, Trond** (2019). "Smart-sustainability: a new urban fix?". *Sustainable cities and society*, v. 45, pp. 640-648.
<https://doi.org/10.1016/j.scs.2018.11.028>
- Morozov, Evgeny** (2013). *To save everything, click here: Technology, solutionism and the urge to fix problems that don't exist*. London: Penguin books. ISBN: 978 0 24195770 7
- Neumaier, Sebastian; Umbrich, Jürgen; Polleres, Axel** (2016). "Automated quality assessment of metadata across open data portals". *Journal of data and information quality*, v. 8, n. 1, Article 2.
<https://doi.org/10.1145/2964909>
- Ramos-Simón, Luis-Fernando; Arquero-Avilés, Rosario; Botezán, Iuliana; Cobo-Serrano, Silvia; Sala-Jiménez, Andrea; Sánchez-Jiménez, Rodrigo; Del-Valle-Gastaminza, Félix** (2012). "De la reutilización de información del sector público a los portales de datos abiertos en Europa". *BiD: textos universitarios de biblioteconomía i documentació*, n. 29.
<https://doi.org/10.1344/BiD2012.29.5>
- Royo-Montañés, Sonia; Benítez-Gómez, Alberto** (2019). "Portales de datos abiertos. Metodología de análisis y aplicación a municipios españoles". *El profesional de la información*, v. 28, n. 6, e280609.
<https://doi.org/10.3145/epi.2019.nov.09>
- Sadiq, Shazia; Indulska, Marta** (2017). "Open data: quality over quantity". *International journal of information management*, v. 37, n. 3, pp. 150-154.
<https://doi.org/10.1016/j.ijinfomgt.2017.01.003>
- Sheffer-Correa, Andreiwid; Mendes-de-Souza, Raul; Soares-Correa-da-Silva, Flavio** (2019). "Towards an automated method to assess data portals in the deep web". *Government information quarterly*, v. 36, n. 3, pp. 412-426.
<https://doi.org/10.1016/j.giq.2019.03.004>
- Smith, Göran; Ofe, Hosea-Ayaba; Sandberg, Johan** (2016). "Digital service innovation from open data: exploring the value proposition of an open data marketplace". In: *2016 49th Hawaii international conference on system sciences (HICSS)*, pp. 1277-1286. Koloa, HI: IEEE.
<https://doi.org/10.1109/HICSS.2016.162>
- Thorsby, Jeffrey; Stowers, Genie N. L.; Wolslegel, Kristen; Tumbuan, Ellie** (2017). "Understanding the content and features of open data portals in American cities". *Government information quarterly*, v. 34, n. 1, pp. 53-61.
<https://doi.org/10.1016/j.giq.2016.07.001>
- Van-Knippenberg, Laura** (ed.) (2020). *Open data maturity. Report 2020*. European Commission. Directorate General for Communications Networks, Content and Technology. European Data Portal. ISBN: 978 92 78 42406 0
https://data.europa.eu/sites/default/files/edp_landscaping_insight_report_n6_2020.pdf
- Veljković, Nataša; Bogdanović-Dinić, Sanja; Stoimenov, Leonid** (2014). "Benchmarking open government: An open data perspective". *Government information quarterly*, v. 31, n. 2, pp. 278-290.
<https://doi.org/10.1016/j.giq.2013.10.011>
- Vicente-Paños, Adrián; Jordán-Alfonso, Aurea** (2017). "Acceso a la información pública y su reutilización en las comunidades autónomas: evaluación de la reutilización de datos abiertos". *El profesional de la información*, v. 26, n. 3, pp. 381-392.
<https://doi.org/10.3145/epi.2017.may.04>
- Wang, Di; Chen, Chuanfu; Richards, Deborah** (2018). "A prioritization-based analysis of local open government data portals: a case study of Chinese province-level governments". *Government information quarterly*, v. 35, n. 4, pp. 644-656.
<https://doi.org/10.1016/j.giq.2018.10.006>
- Welle-Donker, Frederika; Van-Loenen, Bastiaan** (2017). "How to assess the success of the open data ecosystem?". *International journal of digital earth*, v. 10, n. 3, pp. 284-306.
<https://doi.org/10.1080/17538947.2016.1224938>
- Yu, Harlan; Robinson, David G.** (2012). "The new ambiguity of "open government". *UCLA law review discourse*, n. 59, pp. 178-208.
<https://doi.org/10.2139/ssrn.2012489>
- Zuiderwijk, Anneke; Shinde, Rhythima; Janssen, Marijn** (2019). "Investigating the attainment of open government data objectives: is there a mismatch between objectives and results?". *International review of administrative sciences*, v. 85, n. 4, pp. 645-672.
<https://doi.org/10.1177/0020852317739115>