DGABr:
Metric for evaluating Brazilian open government data

Patricia Nascimento Silva*
Marta Macedo Kerr Pinheiro**

ABSTRACT   Open Government Data (OGD) is the term used since 2007 to refer to open government data, that is, public data produced or commissioned by the public sector, available on the internet for free use. The availability of OGD was encouraged by initiatives for disclosure of data carried out in 2009 by the United States of America and these initiatives were developed in several countries. However, with the amount of public information disclosed by governments, the idea of open content has been increasingly associated to access to information. In Brazil, plans that regard information policies do not establish a standard with guidelines or models for OGD publication. Thus, in the absence of metrics to evaluate the initiatives of data openness by the Brazilian government and its reuse, this article presents the Brazilian metric DGABr, one of the results of a doctoral research. DGABr is a proposed model to evaluate the OGD in the Federal Public Administration of Brazil, based on metrics and international indicators. The experience from other countries regarding OGD evaluation directed the creation of the DGABr; however, Brazil's legislation and information policies had a direct influence on OGD measurements. DGABr is an initial proposal to evaluate Brazil's OGD and their reuse and should evolve with the country's information policy. The proof of concept performed with the metric identified a positive result for the OGD evaluation, allowing validation of the perspectives and dimensions proposed in the metric and its potential reuse indicator.

* Doutora em Gestão & Organização do Conhecimento pela Universidade Federal de Minas Gerais, Brasil. Analista da Universidade Federal de Minas Gerais, Brasil.
E-mail: patricia.inf@gmail.com.

** Doutora em Ciência da Informação pelo convênio entre o Instituto Brasileiro de Infomação em Ciência e Tecnoogia e a Universidade Federal do Rio de Janeiro, Brasil. Professora e pesquisadora permanente da Universidade FUMEC, Brasil.
E-mail: martakerr@gmail.com.

Keywords: Open Government Data. Reuse of public information. Information Policy. Open Government Data Metrics.

1 INTRODUCTION

The public sector provides a large volume of data from different domains, representing a potential resource for organizations and individuals that may use them to improve their own data sets, to develop new products or to innovate in their services with efficiency gain. It may also promote transparency, collaboration, participation and social and economic values. As a result, the number of countries that have realized the value of government information

as a source for strategic actions has increased, along with the composition of future scenarios and market changes (BRAMAN, 2006; TARAPANOFF, 2006; ALBANO, 2014; KUCERA, 2014).

With open government initiatives, led by the United States of America (USA) in 2009, several countries have developed open data actions, plans and policies to stimulate the re-use of public sector information. Openness of data enables other uses of that information; that reuse of data has a greater impact on citizens'

ability to monitor efficacy and efficiency in governments. It may also enable public companies, universities, research institutes and the private sector to use open data to develop new products and services, stimulating innovation and improving quality in the provision of public services (SILVA, 2018; VAN DEN BROEK *et al*., 2012).

However, it is necessary to consider the amount of public information that ends up being processed and passed on to society as a form of access to information. The challenge of using public information and not merely releasing it by means of technological tools becomes increasingly relevant (KERR PINHEIRO, 2014). Therefore, in this context, it is necessary to establish guidelines with criteria, concepts and metrics so as to favor the simplification and regulation of access to public sector information.

In Brazil, the Law on Access to Information (LAI), in force since May 2012, with the various plans that are responsible for information policies in Brazil, in addition to the transparency initiatives implemented by data portals and the availability of data sets, can be considered the legal framework for the governmental duty to disclose data in the country. However, simply rendering open government data (OGD) available is an "incomplete" measure that requires further action for its evaluation. Metrics and indicators to evaluate these by products for government activities can be as important as their disclosure to society. In that sense, to assess OGD is important for transparency and for an open government, since an evaluation will allow adequacy improvement in the data available for effective use by government and society. An analysis beyond the information being released is necessary, concerned with what is being measured and how that measurement occurs, in order to develop a clearer conception of who uses the data and how data infrastructures operate in the world (GRAY, 2015; SILVA, 2018).

Given the importance of evaluating the initiatives of data opening by the Brazilian government and the lack of metrics or official indicators related to this subject in the country, this paper presents the creation of DGABr, model metric of evaluation for OGD in the Federal Public Administration of Brazil (APF, from the acronym in Brazilian Portuguese), with international metrics and indicators as its basis. DGABr is an initial proposal for Brazilian metrics based on context and the existing legislation in the country in order to evaluate the data disclosed by governments and their potential for reuse.

2 METHODOLOGY

The DGABr metric is a by product of a doctoral research [1]conducted between 2015 and 2018 focused on the reuse of Brazilian OGD. The study merged qualitative and quantitative elements with exploratory traces at the beginning of research and descriptive ones towards the end, using different techniques that allowed the identification of trends in OGD metrics in other countries. It was also possible to propose metrics and indicators to evaluate the reuse of OGD disclosed by the Brazilian government. The research was divided into three methodological stages: Stages: 1: Bibliographic Review; Stage 2: International Models; and Stage 3: Construction of a model metric for OGD reuse based on Brazilian legislation and standards. Chart 1 shows the methodological stages and the respective techniques and instruments used.

---

[1] The article is a result of the dissertation titled *Open Government Data: metrics and reuse indicators* defended on 19/03/2018 by Patrícia Nascimento Silva, directed by professor Dr. Marta Macedo Kerr Pinheiro in the Postgraduate Program in Management and Knowledge Organization of the Federal University of Minas Gerais.

Chart 1: Methodological stages for data collection

Methodological Stage	Research Technique	Instruments / Procedures
Stage 1	Bibliographic research Documentary research	1) Bibliographic research on databases and the Web on information policies. 2) Analysis of Applications for the Federal Government Application Guide (aplicativos.gov.br)
Stage 2	Document research Ground Theory	1) Data collection on indexed databases of Capes portal (IEEE, LISA, Science Direct, Web of Science, ACM, Springer Linker, SCOPUS), Google Scholar, institution and agency websites on open data between 2009 and 2016 in the following languages: Portuguese, English and Spanish, considering the following expressions in the title, abstract and keywords: "open government data", "datos de gobierno abierto", "gobierno abierto", "dados governamentais abertos" and "dados abertos governamentais". After removing the duplicates, 451 documents were found. 2) Analysis of the documents in the software Atlas ti. 2.1) Coding by year of publication, by type of evaluation and topic, and, finally, the documents on evaluation of open data or OGD were coded. 3) Analysis of the selected documents. 4) Identification of methodologies, models, metrics and international indicators for OGD evaluation.
Stage 3	Analytical Research	Construction of a model metric for OGD reuse based on Brazilian legislation and standards.

Source: authors.

An exploratory research involved a bibliographical review of the information policies and methods for disclosing OGD in Brazil, in Stage 1. Information policies in Brazil, identified by OGD actions and plans, were marked in the form of applications developed by public agencies in the APF and released in the Application Guide of the Federal Government (http://aplicativos.gov.br/). The website was the official government page for releasing the applications which were developed and maintained by public agencies. Due to the unavailability of information related to the form of access, an analysis was made in each application to evaluate information categories, access criteria, data sources and, especially, whether OGD was used.

A document-based research identified methodologies, models, metrics and indicators used in other countries to evaluate the use / reuse of OGD in Stage 2. Using document analysis as a reference, coding techniques were used for the 451 found publications which were, then, analyzed to identify different types of evaluation. Those involved OGD portals, evaluation rankings among countries and evaluations on the publication and reuse of OGD, the latter being the focus of the research and used as a basis for developing a metric for the Brazilian context.

Chart 2 presents a summary of the evaluation methodologies found during data collection, after the coding process, with their respective characteristics, year and place where they were created or used. Methodologies / models with different evaluation methods were found, some of them based on legislation, such as UNE 178301 (Spain) and MELODA (Spain); others, on the other hand, are very general, such as OpenGovB (USA) and OD-MM (Chile, Colombia and El Salvador). The UNE 178301 Standard is one of the most complex metrics; it includes multiple domains and dimensions with innovative content regarding the reuse of open data. The MELODA metric adopts the concept of levels and dimensions and allows the evaluation of data sets and applications. The MELODA, OD-MM and UNE 178301 metrics display structural similarity, using variables with their respective weights. Additionally, MePOD-VS (Czech Republic) and OPEN DATA MATURITY MODEL (United Kingdom) evaluate good publishing practices and maturity levels of the evaluated organization, respectively. Introducing a different focus, OpenGovB provides an analysis that involves the identification of data sources and the forms of involvement with its users (SILVA; KERR PINHEIRO, 2018).

Chart 2: Evaluation methodologies elicited from research

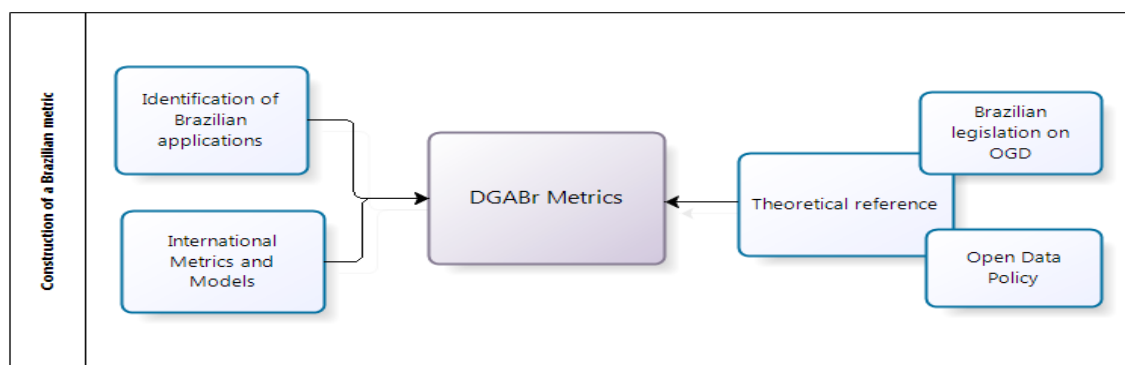| Methodology	Year / Country	Description / aims	Type of data	Author
MePOD-VS	2015 Czech Republic	The methodology is based on better practices for OGD publication. It consists of a set of defined principles with functions and processes organized into groups / domains. For each of the inputs and outputs, specific processes are proposed, and, in some processes, the relevant standards are identified.	OGD	Kucera (2015)
MELODA	2013 Spain	It proposes metrics to evaluate the reuse of open data based on the Spanish legislation. The items evaluated in each dimension are directly related to sharing methods and standards for their publishing. The metric designed for aiding in the reuse of the available data suggests five dimensions - Legal framework, Technical standards, Information accessibility, Data model sharing, geolocalized information and real-time information – with five levels of reuse (1 to 5), level 1 being the lowest and 5 the highest.	OGD Underlying scientific data Crowdsourced data	Rey Juan Carlos University of Madrid MELODA (2016)
UNE 178301 Standard	2015 Spain	It helps define, document and deploy open data projects. It establishes a collection of metrics that take into account aspects that need to be addressed in open data initiatives. Five domains and ten dimensions with established weights are proposed. There are four levels of evaluation (0 to 3), 0 being non-existing result and 3 being advanced result.	Open data	AENOR UNE 178301 (2015)
Open Data Maturity Model (OD-MM)	2014 Chile, Colombia and El Salvador	Diagnosis of open data in public services. It establishes references for design, management, monitoring and control of processes in the context of public administration. It establishes three domains, with nine subdomains and 33 critical variables. For each variable, a weight is assigned and a scale level of 1 to 4 is established: 1 corresponds to non-existent capacity and 4 to advanced capacity.	Open data	Solar *et al.*, (2014)
OPEN DATA MATURITY MODEL	2015 United Kingdom	It evaluates the variety so that the practice of open data can have an impact on an organization. Five themes are proposed - Data management processes, Knowledge & skills, Customer support & engagement, Investment & financial performance and Strategic oversight -, based on the Balanced Scorecard, related to the organization's performance. Activities can also be grouped according to their relationship to the publication and reuse of data or both areas. It has five levels of maturity - Initial, Repeatable, Defined, Managed and Optimising - that represent the different states by which the organization will go through.	Open data	OPEN DATA INSTITU-TE (2015)
OpenGovB	2014 USA	A reference proposal for Open Government and its application from the perspective of open data, using data available in the USA Government Open Data Portal (data.gov). It has two metrics: the open e-government index (e-GovOI) and a benchmark that indicates the government's progress over time.	Open data	Veljković *et al.*, (2014)

Source: Adapted from NASCIMENTO SILVA and KERR PINHEIRO (2017).

Experiences from other countries have provided a diverse view regarding OGD evaluation, especially with Spanish metrics, which directly evaluate the reuse of OGD. The metrics reflect results from information policies and their subdomain of open data and, even if they are tested and used in other countries, the intersection of their variables, principles and rules could contribute to the elaboration of a metric for the Brazilian context.

The third methodological stage consisted of the elaboration of a model with metrics and indicators to evaluate the reuse of OGD in Brazil, to be used both by the branchesthat publish OGD, the producers, as well as the users and citizens that can reuse OGD. The construction of the metric was based on a review of the literature on OGD policies in Brazil, identifying the main laws, decrees, standards and good practices of publication officially suggested in Brazil, along with the results obtained in Stages 1 and 2 of this research. Figure 1 presents an overview of the methodological framework for the construction of the DGABr metric.

Figure 1: Construction of the Brazilian model metric

Source: authors.

All elements presented in Figure 1 were used to develop a proposition for a Brazilian metric to evaluate the potential reuse of Brazilian OGD. The metric was titled DGABr, an acronym that includes its main theme, that is, OGD, and Brazil (from the terms in Brazilian Portuguese). In view of the Brazilian context, the initial version of the metrics also covers the evaluation of the published data, since there is no single standard for OGD publication in Brazil and, therefore, elements inherent to the publication of the data need to be evaluated to identify the potential reuse of the data provided. The metric is focused on the evaluation of data sets, since that is the predominant granularity in the Brazilian Open Data Portal.

3 RESULTS AND DISCUSSIONS

DGABr is a model metric created to evaluate Brazilian OGD. It has five perspectives: Open Data, Legal, Technical, Managerial and Reuse. Each perspective has a set of dimensions and a respective weight. For each dimension, a level has also been assigned.

The number of levels was based on the experience with other standards, metrics and methodologies with emphasis on: UNE 178301:2015 from Spain (UNE, 2015), MELODA from Spain (MELODA, 2016), OD-MM, used in Chile, Colombia and El Salvador (SOLAR *et al.*, 2014) and the Open Data Maturity Model of the United Kingdom (ODI, 2015). Six levels were created within the DGABr. The first three levels (0,1 and 2)

portray the non-existence or non-fulfillment of the analyzed dimension and the last three (3,4 and 5) depict the fulfillment of the dimension analyzed. Level 3 was considered the minimum for OGD to be reused and this level was assigned an average value, that is, 50% of the assessed dimension. Level 4 contemplates great assistance in the evaluated item, that is, 80%. Finally, level 5 was created to include full service of the item evaluated, that is, 100%. It should be noted that the values defined for level assessment were used for most dimensions, but for specific dimensions such as D16, D24, D25 and D26, other criteria were considered, duly justified and defined in the metric. Chart 3 presents the DGABr metric summary with the perspectives and their respective weights and levels.

Chart 3: Construction of the Brazilian model metric.

DGABr Metric				
Perspectives	**Weights**	**Levels**	**Definition**	**Description**
Open Data	1	Level 0	Nonexistent	The institution does not publish OGD or has no definition for publication.
Legal	3	Level 1	Under construction	The institution does not publish or has no definition for publication, but there is a study or the construction of a standard.
Technical	2	Level 2	Not executed	The institution has a standard or there is a recommendation for the publication of the OGD, but the publication does not follow what has been defined.
Managerial	2	Level 3	Partially Performed	The institution complies with the standards and recommendations defined in a partial way by up to 50%.
Reuse	3	Level 4	Existing Results	The institution complies with the standards and recommendations defined in part, by at least 80%.
		Level 5	Advanced Results	The institution meets all the standards and recommendations defined.

Source: authors.

3.1 Elements of DGABR metrics

The Open Data perspective has ten dimensions related to the basic characteristics related to the OGD principles and the five-star model, which were the first recommendations and models for good practices in the publication of open data. Since these were fundamental criteria for OGD, weight 1 was assigned from a scale of 1 to 3, because they are basic elements previously required in open data publications. As assessed from the literature, the evaluation of these criteria is present in the main international models and metrics. Chart 4 details the dimensions of the Open Data perspective.

Chart 4: Dimensions and evaluation levels in the Open Data Perspective

D1 - Complete data assesses whether the data provided are complete	Level 0: The data is not complete; the institution only publishes OGD without validating its completeness
	Level 1: The institution is building criteria and strategies to verify that the data is complete, but currently only publishes OGD
	Level 2: The institution has criteria and strategies to verify that the data is complete, but these were not used for the analyzed release
	Level 3: At least 50% of OGD are complete
	Level 4: At least 80% of OGD are complete
	Level 5: All OGD are complete and validated prior to publication
D2 - Primary data evaluates whether data is presented as it was collected at source, in its crudest form, with the highest possible level of granularity, without aggregation or modification, except for the protection of personal data	Level 0: The institution does not present data as collected at the source; aggregations and modifications are not justified or documented
	Level 1: The institution does not present the data as it was collected at the source, but it has a document under construction to justify the aggregations and modifications made in the OGD
	Level 2: The institution does not present the data as it was collected at the source; it has a document to record the aggregations and modifications, but the information was not recorded
	Level 3: At least 50% of the OGD are presented as collected and, when they present aggregations and /or modifications, they are documented
	Level 4: At least 80% of the OGD are presented as collected and, when they present aggregations and /or modifications, they are documented
	Level 5: All data are presented as they were collected and when they present aggregations and /or modifications, they are documented
D3 - Timely data evaluates whether the data is up to date. To be up-to-date, they must be disclosed as quickly as it is necessary to preserve their value.	Level 0: OGD are not updated and no update frequency has been set
	Level 1: OGD are not updated, but an update frequency is under construction
	Level 2: OGD are not updated, but an update frequency has been set
	Level 3: At least 50% of the OGD are updated according to the defined update frequency
	Level 4: At least 80% of the OGD are updated according to the defined update frequency
	Level 5: All OGD are updated according to the defined update frequency
D4 - Accessible data evaluates whether data is accessible, whether it is disclosed to the widest possible range of users and to the widest possible set of purposes.	Level 0: OGD are not available via web or manual request
	Level 1: OGD are made available by manual request
	Level 2: OGD are made available via web access with some form of identification, registration or login
	Level 3: At least 50% of OGD are made available via web access without some form of identification
	Level 4: At least 80% of OGD are made available via web access without some form of identification
	Level 5: All OGD are made available via web access without some form of identification
D5 - Machine processable evaluates whether the data can be processed by a machine if they are reasonably structured, so as to enable automated processing in order to allow the analysis of large numbers of records without manual intervention.	Level 0: OGD are not structured
	Level 1: OGD are not structured, but there is a plan under construction to structure them
	Level 2: OGD are not structured, but the institution has set the standards for OGD structuring
	Level 3: At least 50% of OGD are structured and machine processable
	Level 4: At least 80% of OGD are structured and machine processable
	Level 5: All OGD are structured and machine processable

D6 - Non-discriminatory Access assesses whether access is non-discriminatory, if the data is available to everyone without requiring identification, registration, application or submission.	Level 0: OGD are not available and access is made only upon request
	Level 1: OGD are available via identification of the applicant (personal data or data from the requesting institution), but there are studies in the institution to turn access freely available
	Level 2: OGD are available to any person through some type of identification, not necessarily with the validation of personal data, such as creating a login
	Level 3: At least 50% of OGD are available without any type of identification requirement
	Level 4: At least 80% of OGD are available without any type of identification requirement
	Level 5: All OGD are available without any type of identification requirement
D7 - Non-proprietary data format evaluates whether data formats are non-proprietary, if the data is available in a format over which no entity has exclusive control.	Level 0: OGD are only published in proprietary formats
	Level 1: OGD are available in proprietary formats, but the definition of a standard that uses non-proprietary formats is under construction
	Level 2: OGD are published in proprietary formats, although there is a standard that uses non-proprietary formats defined for OGD publication in the institution
	Level 3: At least 50% of published OGD use non-proprietary standards
	Level 4: At least 80% of published OGD use non-proprietary standards
	Level 5: All published OGD use non-proprietary standards
D8 - License-free data assesses whether the data is free of license, if it is not subject to any restriction of copyright, patent, intellectual property or industrial secret. Restrictions related to privacy, security and access privileges are allowed.	Level 0: Data has some type of license, restriction or access privilege
	Level 1: Data has some type of license, restriction or access privilege, but new conditions of use and licensing are under construction by the institution
	Level 2: Data has some type of license, restriction or access privilege, although the institution has defined conditions of use and licensing to allow the reuse of the OGD
	Level 3: At least 50% of published OGD are free of licenses, restrictions or access privileges
	Level 4: At least 80% of published OGD are free of licenses, restrictions or access privileges
	Level 5: All published OGD are free of license, restrictions or access privileges
D9 - Designed URIs evaluates whether well-designed URIs are used to identify the data, so that people can reference and identify them easily and simply.	Level 0: Does not use URIs or URLs that identify OGD
	Level 1: Does not use URIs or URLs that identify OGD, but it has a standard or recommendations under construction
	Level 2: Does not use URIs or URLs that identify OGD, but there is a standard or recommendation in the institution
	Level 3: At least 50% of published OGD use URIs or URLs that identify OGD
	Level 4: At least 80% of published OGD use URIs or URLs that identify OGD
	Level 5: All published OGD use URIs or URLs that identify them
D10 - Linked data assesses whether the data is linked to data from other people or other sources to provide context.	Level 0: Published OGD are not linked to other sources
	Level 1: The published OGD are not linked too other sources, but the institution studies ways of connecting with other sources
	Level 2: The published OGD are not linked too ther sources, but there are recommendations and publication standards in the institution that encourage and allow that connection
	Level 3: At least 50% of published OGD have connections to other sources
	Level 4: At least 80% of published OGD have connections to other sources
	Level 5: All published OGD have links to other sources

Source: authors.

The Legal perspective has three dimensions that evaluate specific elements of Brazilian legislation, based mainly on the LAI and the documents from Decree n. 8.777 / 2016, which established the Open Data Policy of the Executive Power. The perspective has three weights because it is a requirement of great relevance, since compliance with the legislation is essential in the evaluation of the metric. Chart 5 details the dimensions of the Legal perspective.

Chart 5: Dimensions and evaluation levels in the Open Data Perspective

D11 - Type of information evaluates origin and type of document, according to the standing legislation. In order to do so, it is necessary to check whether that type has been foreseen in legislation and whether it really is OGD, or just a common information that does not concern the government	Level 0: The type of published information is not foreseen in the legislation and the institution does not evaluate whether that information is ODG
	Level1: The type of published information is not foreseen in the legislation, but the institution studies rules and criteria to evaluate whether that information is ODG
	Level2: The type of document is not foreseen in the legislation and it is not OGD, despite the institution having rules and criteria to evalute the type and origin of the information
	Level3: At least 50% of information / published data is foreseen in the legislation and it is considered OGD
	Level4: At least 80% of information / published data is foreseen in the legislation and it is considered OGD
	Level5: All information / published data is foreseen in the legislation and it is OGD
D12 - Sensitive Data Management evaluates conditions of use and approach standards in the de-sensibilization of the institution. If information is marked as classified, it evaluates whether deadlines are met and whether information is disclosed after that period prescribes.	Level0: The institution does not have rules or internal standards for desensitizing data and it does not evaluate deadlines for periods of secrecy
	Level1: Internal norms or patterns for desensitizing sensitive data and deadlines for secrecy periods are under construction
	Level2: Internal norms or patterns for desensitizing sensitive data and deadlines for secrecy periods exist, but they are not used for OGD publication
	Level3: At least 50% of published OGD have de-sensitized data according to norms and standards defined by the institution and follow deadlines for secrecy periods
	Level4: At least 80% of published OGD have de-sensitized data according to norms and standards defined by the institution and follow deadlines for secrecy periods
	Level5: All published OGD have de-sensitized data according to norms and standards defined by the institution and follow deadlines for secrecy periods
D13 - Open Data Plan evaluates the instituition's ODP. According to Decree Nº. 8.777/2016, Art. 5º, ODP publication is mandatory for sectors directly related to foundational, autarchic APF.	Level0: Institution has no ODP
	Level1: The institution's ODP is under construction
	Level2: Institution has ODP, but its commitments and deadlines have not been fulfilled
	Level3: At least 50% of ODP commitments and deadlines have been fulfilled
	Level4: At least 80% of ODP commitments and deadlines have been fulfilled
	Level5: All ODP commitments and deadlines have been fulfilled

Source: authors.

The Technical perspective has seven dimensions that evaluate the use of the technical standards for the publication of OGD in Brazil, with recommendations and good practices related to the technical elements for the publication of data around the world. This perspective has a weight of 2 because it relates elements necessary to the OGD, associated to the form of publication and data management. The great difficulty found in this perspective was to define a standard, since in Brazil there are several documents on the publication of OGD with different recommendations based on different legislations and periods. The Electronic Government Interoperability Standards (e-PING) and the Electronic Government Metadata Standard (e-PMG) were key to building this perspective. Chart 6 details the dimensions of the Technical perspective.

Chart 6: Dimensions and levels of evaluation for the Technical Perspective

D14 - Source of information assesses whether information on the source of the OGD (raw data) is available in OGD catalogs, delivery platforms and similar elements.	Level 0: OGD source is not informed
	Level 1: The source of the OGD is not informed, but the institution studies ways to disclose that information
	Level 2: The source of the OGD is not informed, although there is a recommendation that this information be disclosed by the institution
	Level 3: At least 50% of OGD sources are available
	Level 4: At least 80% of OGD sources are available
	Level 5: All OGD sources are available
D15 - Data format evaluates whether the recommended formats for OGD publication are used. The recommended formats include the suggestions from e-PING, with the recommendations of the Technical Book of Data Publishing.	Level 0: The institution publishes OGD in non-recommended file formats
	Level 1: The institution publishes OGD in file formats not recommended, but studies using the recommended standards
	Level 2: The institution publishes OGD in non-recommended file formats, although it recommends standards already in place for publication
	Level 3: At least 50% of published OGD use the formats recommended by existing standards [JSON, XML, CSV, ODS, ODF, RDF, HTML (W3C)], API or SPARQL service. For special cases, they use compressed files in the recommended formats (7Z, TAR / GZIP or ZIP, GML)
	Level 4: At least 80% of published OGD use the formats recommended by existing standards [JSON, XML, CSV, ODS, ODF, RDF, HTML (W3C)], API or SPARQL service. For special cases, they use compressed files in the recommended formats (7Z, TAR / GZIP or ZIP, GML)
	Level 5: All published OGD use the formats recommended by existing standards [JSON, XML, CSV, ODS, ODF, RDF, HTML (W3C)], API or SPARQL service. For special cases, they use compressed files in the recommended formats (7Z, TAR / GZIP or ZIP, GML)
D16 – Quality evaluates the quality of OGD published by the institution. This indicator is calculated from the average of the levels reached in the ten dimensions of the Open Data perspective. Thus, since the weight for these dimensions equals one, the calculation to obtain the dimension level D16 is the sum of the levels divided by ten: D16 = (D1 + D2 + D3 + D4 + D5 + D6 + D7 + D8 + D9 + D10) / 10.	Level 0: In the Open Data perspective, the average of the obtained levels was equal to zero
	Level 1: In the Open Data perspective, the average of the levels obtained was equal to one
	Level 2: In the Open Data perspective, the average of the levels obtained was equal to two
	Level 3: In the Open Data perspective, the average of the levels obtained was equal to three
	Level 4: In the Open Data perspective, the average of the obtained levels was equal to four
	Level 5: In the Open Data perspective, the average of the levels obtained was equal to five.

D17 - Metadata evaluates whether the minimum metadata has been published for the OGD available.	Level 0: The institution does not provide the minimum metadata suggested by the ODP
	Level 1: The institution does not provide the minimum metadata suggested by an ODP, because an ODP is under construction
	Level 2: The institution does not provide the minimum metadata suggested by the ODP, despite having a published ODP
	Level 3: The institution provides the suggested minimum metadata for at least 50% of the published OGD
	Level 4: The institution provides the suggested minimum metadata for at least 80% of the published OGD
	Level 5: The institution provides the suggested minimum metadata for all published OGD
D18 - Vocabulary assesses whether a specific controlled vocabulary, a vocabulary used in the market or the Controlled Vocabulary of Electronic Government (VCGE, in Brazilian Portuguese) is used.	Level 0: Published OGD do not use a controlled vocabulary
	Level 1: Published OGD do not use a controlled vocabulary, but the use or creation of a vocabulary is being considered
	Level 2: The published OGD do not use a controlled vocabulary, but the institution recommends using a vocabulary
	Level 3: At least 50% of published OGD use a controlled vocabulary
	Level 4: At least 80% of published OGD use a controlled vocabulary
	Level 5: All published OGD use a controlled vocabulary
D19 - Geographic dimension assesses whether the data are in accordance with the geographic region stipulated in the ODP.	Level 0: The published OGD lack a geographical coverage defined in the ODP
	Level 1: The published OGD lack geographic coverage, but the institution's ODP is under construction
	Level 2: OGD have geographical coverage, but not the same as defined in the ODP
	Level 3: At least 50% of published OGD conform to the geographical coverage defined in the ODP
	Level 4: At least 80% of published OGD conform to the geographical coverage defined in the ODP
	Level 5: All published OGD conform to the geographical coverage defined in the ODP
D20 - Data cataloging evaluates if the OGD are cataloged in the Brazilian Open Data Portal. If the institution has its own catalog, it evaluates whether the data is also cataloged in the Brazilian Open Data Portal.	Level 0: The institution does not have a catalog of its own and has not catalogued the data in the Brazilian Open Data Portal
	Level 1: The institution does not have its own catalog and has not catalogued the data in the Brazilian Open Data Portal, but the catalog is under construction
	Level 2: The institution has its own catalog and has not catalogued the data in the Brazilian Open Data Portal
	Level 3: The institution catalogs at least 50% of the OGD in the Brazilian Open Data Portal
	Level 4: The institution catalogs at least 80% of the OGD in the Brazilian Open Data Portal
	Level 5: The institution catalog all OGD in the Brazilian Open Data Portal

Source: authors.

The Managerial has three dimensions and it was created based on institutional planning, broadcast by documents, action plans and its ODP. It involves management elements related to the planning of OGD publications and the monitoring and control of published OGD. To this perspective, a weight of 2was assigned. The dimensions of the Managerial perspective are detailed in Chart 7.

Chart 7: Dimensions and levels of evaluation in the Managerial Perspective

D21 - Priorities and Open Strategies assesses whether the institution has defined criteria for opening data and databases. As foreseen in the ODP, the institution should prioritize the thematic areas most requested by Citizen Information Service (SIC, in Brazilian Portuguese) and other channels of communication with civil society.	Level 0: The institution does not have a criterion for prioritizing data opening
	Level 1: The institution does not have a criterion for prioritizing data opening, but it is under construction
	Level 2: The institution has a criterion for prioritizing data opening, but it is not fulfilled
	Level 3: At least 50% of OGD are published according to the opening priority criterion
	Level 4: At least 80% of OGD are published according to the opening priority criterion
	Level 5: All OGD are published according to the opening priority criterion
D22 - Investment strategies and timelines assesses whether the institution has a schedule for data opening investments and a timeline, and whether the strategies are met.	Level 0: The institution does not have a schedule for opening databases, updating and improving them
	Level 1: The institution does not have a schedule for opening databases, updating and improving, but is under construction
	Level 2: The institution has a OGD publication schedule, but it is not fulfilled
	Level 3: At least 50% of the published OGD comply with the publication, update and improvement schedule defined in the ODP
	Level 4: At least 80% of the published OGD comply with the publication, update and improvement schedule in the ODP
	Level 5: All published OGD conform to the publication, update and improvement schedule defined in the ODP
D23 - Maintenance and correction of problems evaluates if the catalogs have some type of maintenance and if problems detected in the available data communicated by the citizens are promptly corrected.	Level 0: The institution is not maintained and it does not correct the problems identified in the catalog or clarifies doubts in interpretation
	Level 1: The institution is not maintained and it does not correct the problems identified in the catalog or clarifies doubts in interpretation, but it plans to perform that task
	Level 2: The institution is aware of the problems of the catalog, but does not maintain or correct or clarify doubts in interpretation
	Level 3: At least 50% of OGD have maintenance and correction and some document or communication channel to clarify interpretation doubts
	Level 4: At least 80% of OGD have maintenance and correction and some document or communication channel to clarify interpretation doubts
	Level 5: All OGD have maintenance and correction and some document or communication channel to clarify interpretation doubts

Source: authors.

The Reuse perspective, with five dimensions, was created based mainly on international experiences, since Brazil does not have specific legislation on the reuse of open data or OGD. With five dimensions, it evaluates whether the OGD has been reused in other applications. Because it is of great importance, weight 3 has been assigned. Chart 8 details the dimensions of the Reuse perspective.

Chart 8: Dimensions and levels of evaluation in the Reuse Perspective

D24 - Number of published data sets evaluates data sets published by the institution. The objective of this dimension is to evaluate only the number of publications for future comparisons and to establish a control for published set growth APF branch.	Level 0: The institution has not published any data set
	Level 1: The institution has not published any data set, but its publication is already under implementation
	Level 2: The institution publishes at least 20% of the OGD sets proposed in the ODP
	Level 3: The institution publishes between 21 and 40% of the OGD sets proposed in the ODP
	Level 4: The institution publishes between 41 and 60% of the OGD sets proposed in the ODP
	Level 5: The institution publishes more than 80% of the OGD sets proposed in the ODP
D25 - Number of developed applications evaluates quantitatively the reuse of a given set of data in applications developed by public agencies.	Level 0: Published OGD sets were not used in applications
	Level 1: Published OGD sets were not used in applications, but applications are under development
	Level 2: The institution's OGD sets generated at least one application
	Level 3: The OGD sets of the institution generated between two and five applications
	Level 4: The OGD sets of the institution generated between six and ten applications
	Level 5: The institution's OGD sets have generated more than ten applications
D26 - Download rate of applications evaluates the number of people who downloaded the application developed from the published OGD. Although the download does not guarantee the use of the application, this information may reflect, at least in the disclosure of the application, that it has been downloaded and installed. Because apps are available on application platforms / stores, such as Google Play, this amount can easily be viewed by any user.	Level 0: Applications developed with the institution's OGD sets are not available in app stores (example: Google play, Apple store)
	Level 1: Applications developed with the institution's OGD sets are not available in application stores, but they are being implemented
	Level 2: Applications developed with the institution's OGD sets have less than 1000 downloads
	Level 3: Applications developed with the institution's OGD sets have between 1000 and 3000 downloads
	Level 4: Applications developed with the institution's OGD sets have between 3001 and 5000 downloads
	Level 5: Applications developed with the institution's OGD sets have more than 5000 downloads

D27 - Civil society engagement assesses whether the institution has any service or system / application directed to civil society, whether it releases a product as a database or promotes events on the OGD of the organization as engagement actions. The actions promoted may be disclosed on the institution's website or in its ODP.	Level 0: The institution has no engagement with society
	Level 1: The institution has no engagement with society, but engagement actions are under development
	Level 2: The institution has no engagement with society, but it has planned engagement actions
	Level 3: At least 50% of planned engagement actions were performed
	Level 4: At least 80% of planned engagement actions were performed
	Level 5: All planned engagement actions were performed
D28 - Information on OGD reuse evaluates the information on the OGD used, the data sources and the frequency of updates in the products and services generated from the OGD. This data is important to the users generated products and services.	Level 0: The products and services developed by the institution do not have information about the sources and updates of the OGD used
	Level 1: The products and services developed by the institution do not have information about the sources and updates of the OGD used, but the it studies ways of disseminating that information
	Level 2: Despite recommendations, the products and services developed by the institution do not have information about the sources and updates of the OGD used;
	Level 3: At least 50% of the products and services developed by the institution have information about the sources and updates of the OGD used
	Level 4: At least 80% of the products and services developed by the institution have information about the sources and updates of the OGD used
	Level 5: All the products and services developed by the institution have information about the sources and updates of the OGD used

Source: authors.

It should be emphasized that the Reuse perspective is very complex to be measured and the proposed dimensions mainly evaluate quantitative aspects that suggest, but do not ensure, the reuse of the OGD. It is an initial evaluation, which should evolve with metric use.

The value obtained by the sum of the 28 dimensions and their respective weights corresponds to the reuse indicator [Reuse Indicator = Σ (level$_n$ x weight$_n$)]. The highest possible score is 270. Based on the maximum value of each level for the 28 dimensions, the figure obtained was divided into 5 indicators (1 to 5) (Chart 9), which represent, in ascending order, the OGD reuse rate in the organization or the data set evaluated.

Chart 9: Final value obtained by metric

Result value	Indicator
Total <54	1
108> Total ≥ 54	2
162> Total ≥ 108	3
216> Total ≥ 162	4
Total ≥ 216	5

Source: authors.

A summary of the DGABr Metric with its perspectives and dimensions is presented in Chart 10.

Chart 10: Summary of the DGABr Metric

DGABr Metric					
Perspective	**ID**	**Dimension**	**Weight**	**Acquired level**	**Sum (weight x level)**
Open Data	D1	Complete data	1		
	D2	Primary data	1		
	D3	Timely data	1		
	D4	Accessible data	1		
	D5	Machine processable	1		
	D6	Non-discriminatory access	1		
	D7	Non-proprietary data format	1		
	D8	License-free data	1		
	D9	Designed URIs	1		
	D10	Linked data	1		
Legal	D11	Type of information	3		
	D12	Sensitive data management	3		
	D13	Open Data Plan	3		
Technical	D14	Source of information	2		
	D15	Data format	2		
	D16	Quality	2		
	D17	Metadata	2		
	D18	Vocabulary	2		
	D19	Geographical dimension	2		
	D20	Data cataloging	2		
Managerial	D21	Priorities and open strategies	2		
	D22	Investment strategies and timelines	2		
	D23	Maintenance and correction of problems	2		
Reuse	D24	Number of published data sets	3		
	D25	Number of developed applications	3		
	D26	Download rate of applications	3		
	D27	Civil society engagement	3		
	D28	Information on OGD reuse	3		

Source: authors.

The degree of maturity of the organization is reflected by the indicator of potential reuse in relation to the reuse of OGD, since the dimensions evaluate essential elements for the OGD to be reused by other actors. Indicators 1 and 2 correspond to levels 0, 1 and 2 in most dimensions, that is, it corresponds to levels "nonexistent", "under construction" and "not executed" in relation to OGD publication. Indicator 3 is considered the minimum level for the institution's OGD to be reused. Indicators 4 and 5 show that OGD were published in a planned and orderly manner by the institution, resulting in a higher degree of maturity in relation to OGD reuse.

The metric under went a proof of concept by analyzing the OGD from a branch of the Brazilian APF, the Ministry of Education (MEC, in Brazilian Portuguese). This evaluation enabled a test for the proposed dimensions and the potential reuse indicator of the OGD. With the tests, it was possible to adjust elements that were theoretically correct, but which, when evaluated in practice, could be visualized in another way, indicating a need for minor modifications. The DGABr focuses on the evaluation of OGD, but it can be adapted for the evaluation of open data; in fact, it pervades several elements, such as data catalogs, OGD publication portals and documents produced by the institutions in order to be able to measure OGD. The different perspectives that were created enabled a convergence of similar elements, which, in turn, allowed us to organize the evaluation of specific items of data sets and dimensions that evaluate the institution. In this way, the metric is comprehensive in relation to the OGD context, since it can be used to evaluate a single data set or several data sets released by an institution.

4 FINAL REMARKS

DGABr is one of the results of a doctoral research developed between 2015 and 2018 by the authors, based on the reuse of open government data. The metric was built according to international experiences about the evaluation of open data, based mainly on the ODP and other existing legislation and standards in Brazil for the publication of open government data. The analyzes of the Brazilian OGD and their products, during one of the stages of research, also subsidized the construction of the metric. Composed of qualitative and quantitative dimensions, DGABr evaluates elements of the OGD life cycle and suggests an indicator, based on the scores obtained in the dimensions, to measure the potential reuse of the evaluated OGD.

The proof of concept carried out with data sets from a Brazilian APF branchallowed the adaptation of the metric, mainly in the form of measuring some dimension and validating it within the Brazilian context. Thus, the metric meets its objective to evaluate OGD published in Brazil.

The initial version of DGABr provides an evaluation of the OGD published by Brazil's APF and its potential reuse, in a Brazilian context, in accordance with international standards and experiences. It is an important instrument for measuring results obtained from the efforts and investments made in the publication of open data. However, for the metric to remain in accordance to the Brazilian government's open data policy, still under construction, new versions should be published, adapting the metric to each change in legislation. The definition of a single standard for publication of OGD in Brazil will allow an evolution of the metric and an evaluation directed to the results of the availability and reuse of OGD.

DGABr: Métrica para avaliação de dados governamentais abertos brasileiros

RESUMO   *Dados Governamentais Abertos (DGA) é o termo utilizado desde 2007 para se referir aos dados abertos de domínio governamental, ou seja, dados públicos produzidos ou encomendados pelos órgãos do setor público, disponibilizados através da internet para a livre utilização. A disponibilização dos DGA foi incentivada pelas iniciativas de abertura de dados protagonizadas em 2009 pelos Estados Unidos da América (EUA) e desenvolvidas em vários países. Porém, com o quantitativo de informações públicas disponibilizadas pelos governos, tem se evidenciado cada vez mais o conteúdo aberto como acesso à informação. No Brasil, os planos que respondem por políticas de informação não estabelecem um padrão com diretrizes ou modelos de publicação dos DGA. Assim, diante da ausência de métricas para avaliar as iniciativas de abertura de dados pelo governo brasileiro e seu reúso, este artigo apresenta a métrica brasileira DGABr, um dos resultados de uma pesquisa de doutorado. A DGABr é um modelo proposto para avaliar os DGA na Administração Pública Federal do Brasil, tendo como base métricas e indicadores internacionais. A experiência de outros países com relação à avaliação dos DGA direcionou a criação da DGABr, contudo a legislação e as políticas de informação do Brasil influenciaram diretamente na forma como os DGA são mensurados. A DGABr é uma proposta inicial para avaliar os DGA brasileiros e seu reúso e deverá evoluir juntamente com a política de informação do Brasil. A prova de conceito realizada com a métrica identificou um resultado positivo na avaliação dos DGA, permitindo validar as perspectivas e dimensões propostas na métrica e seu indicador de potencial reúso.*

Palavras-chave: *Dados Governamentais Abertos. Reúso de informações públicas. Política de Informação. Métricas de Dados Governamentais Abertos.*

REFERENCES

ALBANO, C. S. **Dados governamentais abertos:** proposta de um modelo de produção e utilização de informações sob a ótica conceitual da cadeia de valor. 2014. Tese (Doutorado em Administração) - Faculdade de Economia, Administração e Contabilidade, Universidade de São Paulo, São Paulo, 2014. Disponível em: <http://www.teses.usp.br/teses/disponiveis/12/12139/tde-03062014-170642/pt-br.php>. Acesso em: 14 jul. 2017.

BRAMAN, S. **Change of State**: information, policy and power. Cambridge: MIT Press, 2006.

GRAY, J. **A data revolution for whom**?**Open Democracv UK**. 2015. Disponível em: <https://www.opendemocracy.net/ ourkingdom/ jonathan-grey/data-revolution-for-whom>. Acesso em: 14 jun. 2018.

KERR PINHEIRO, M. M. **O processo de construção de políticas de informação**. In: MOURA, Maria Aparecida. (Org.). A construção social do acesso público à informação no Brasil: contexto, historicidade e repercussões. Belo Horizonte: UFMG, 2014, v.1, p.27-46.

KUCERA, J.**Methodologies for publication of open government data.** Research report: VSE 2014. Disponível em: <http:// nb.vse.cz/ ~xkucj30/dissertation/Kucera_OGD_methodologies_EN_v1.pdf> Acesso em: 14 jun. 2018.

KUCERA, J. Open Government Data Publication Methodology. **Journal of Systems Integration**, v.6, n.2, p.52, 2015. Disponível em: <https://web.a.ebscohost.com/abstract? direct=true&profile=ehost&scope= site&authtype=crawler&jrnl= 18042724&AN=116295181&h= sBJFmzXA4VJo%2bNvEJJ2Ue% 2ba6sugXe3ZH7

u V K i V i p % 2 f O Z % 2 f g A l F 8 8 c A 7 I I % 2 f p 4 1 A X l g L I 2 Z d K l 4 z U m c 1 3 % 2 b a J I O p D H A % 3 d % 3 d & c r l = f & r e s u l t N s = A d m i n W e b A u t h & r e s u l t L o c a l =ErrCrlNotAuth&crlhashurl= login. aspx%3fdirect% 3dtrue%26profile%3dehost% 2 6 s c o p e % 3 d s i t e % 2 6 authtype%3dcrawler%26jrnl% 3d18042724% 26AN%3d116295181>. Acesso em: 10 jul. 2018.

METRIC FOR ASSESSING OPEN DATA SOURCES (MELODA). **Metric for data reusability**. Full description of Meloda. 2016. Disponível em: <http://www.meloda.org/full-description-of-meloda/>. Acesso em: 14 jun. 2018.

SILVA, P. N. ; KERR PINHEIRO, M. M. MÉTRICAS PARA DADOS GOVERNAMENTAIS ABERTOS: ANÁLISE DE MODELOS INTERNACIONAIS. In: XVIII Encontro Nacional de Pesquisa em Ciência da Informação (XVIII ENANCIB), 2017, Marília - SP. **ANAIS XVIII ENANCIB**, 2017. Disponível em: <http://enancib.marilia.unesp.br/ index.php/xviiienancib/ENANCIB/paper/ viewFile/90/525>. Acesso em: 14 jun. 2018.

_____. Métricas para Dados Governamentais Abertos. **RICI: Revista Ibero-Americana de Ciência da Informação**, v. 11, n.1, p. 314-332, 2018. Disponível em: <http://periodicos.unb.br/ index.php/RICI/article/view/25726>. Acesso em: 14 jun. 2018.

SILVA, P. N. **Dados governamentais abertos:** métricas e indicadores de reúso. 2018. 322f. Tese (Doutorado em Gestão e Organização do Conhecimento) - Escola de Ciência da Informação, Universidade Federal de Minas Gerais, Belo Horizonte, 2018. Disponível em: < http://www.bibliotecadigital.ufmg.br/dspace/ handle/1843/BUBD-AYNG4U>. Acesso em: 10 jul. 2018

OPEN DATA INSTITUTE (ODI). A guide to the Open Data Maturity Model Assessing your open data publishing and use. 2015. Disponível em: <https://pt.scribd.com/document/260481608/ O D I - M a t u r i t y - M o d e l - G u i d e - A s s e s s i n g-youropen-data-publishing-and-use>. Acesso em: 14 jun. 2018.

SOLAR, M. *et al*. A Model to Guide the Open Government Data Implementation in Public Agencies. **Journal Of Universal Computer Science**, v. 20, n.11, p.1564-1582, 2014. Disponível em: <https://pdfs.semanticscholar.org/361a/ aec3cf949217469abc9a4795be84abf56a08.pdf >. Acesso em: 14 jun. 2018.

TARAPANOFF, K. (Org.). **Inteligência, informação e conhecimento em corporações**. Brasília: IBICT; UNESCO, 2006. p.73-98.

UNE 178301. **Ciudades Inteligentes**. Datos Abiertos (Open Data). AENOR, 2015. Disponível em: < https://www.une.org/encuentra-tu-norma/busca-tu-norma/norma/?c=N0054318 >. Acesso em: 10 jul. 2018.

VAN DEN BROEK, T.; RIJKEN, M.; VAN OORT, S. **Towards Open Development Data**: A review of open development data from a NGO perspectives. Holanda, julho 2012. Disponível em: <https://repository.tudelft.nl/view/tno/ uuid:c1ef3a5a-155d-4139-bb47-360a401ca339/>. Acesso em: 14 jun. 2018.

VELJKOVIĆ, N.; BOGDANOVIĆ-DINIĆ, S.; STOIMENOV, L. Benchmarking open government: An open data perspective. **Government Information Quarterly**, v.31, n 2, p.278-290, 2014. Disponível em: <https:// www.sciencedirect.com/science/article/pii/ S0740624X14000434>. Acesso em: 14 jun. 2018.