RDBCI | Revista Digital de Biblioteconomia e Ciência da Informação
Digital Journal of Library and Information Science

ISSN 1678-765X
DOI 10.20396/rdbci.v18i00.8661528/25795

RESEARCH ARTICLE

Frictionless Data-based model applied to open government data

Melissa Figueira Fagundes [1] https://orcid.org/0000-0003-1082-9245
Divino Ignácio Ribeiro Junior [2] https://orcid.org/0000-0002-9705-6507

[1] Santa Catarina State University, Florianópolis, SC, Brazil / e-mail: melffagundes@gmail.com
[2] Santa Catarina State University, Florianópolis, SC, Brazil / e-mail: divino@udesc.br

ABSTRACT

This study discusses how a model based on Fricitionless Data (FD) can facilitate the publication of open government data (DAGs). FD is an initiative of the Open Knowledge Foundation, which aims to eliminate "friction" in working with data, that is, when much time and resources are lost to understand and work with the data. Within public institutions, the lack of standards for publication and processing of DAGs is a common problem when it comes to opening DAGs. The study tried to contribute by applying the model to the set of data with information on good practices executed by the Judiciary related to the Objectives of Sustainable Development (ODS). Through a bibliographic survey, the main characteristics found in the models for opening DOGS and barriers inherent to the publication process, were identified in the literature. The laboratory research technique was used to describe and analyze the application of the proposed model in a controlled environment for specific purposes of this work. The model was based on Data Publication Workflow, a data publication flow available at the Frictionless Data initiative website that addresses steps such as data packaging, treatment and publication. The implementation of the model also used open source tools based on Frictionless Data. The results showed the feasibility of the model for opening up a set of open government data, as well as demonstrated that the tools available in the Frictionless Data initiative contributed to the verification of the steps in the model. It was concluded that the model needs validation in other contexts, such as integration and opening of different public databases related to ODS.

KEYWORDS: Open data government. Government. Frictionless data. Open data.

Modelo baseado em *Frictionless Data* aplicado aos dados abertos governamentais

RESUMO

O presente trabalho discute como um modelo baseado em *Fricitionless Data* (FD) pode auxiliar na publicação de dados abertos governamentais (DAGs). FD é uma iniciativa da *Open Knowledge Foundation*, que pretende remover o "atrito" no trabalho com os dados, ou seja, quando se perde muito tempo e recursos para entender e trabalhar com o dado. No âmbito das instituições públicas, a ausência de padrões para publicação e processamento dos DAGs é problema comum quando se trata de abertura dos DAGs. O estudo procurou contribuir ao aplicar o modelo ao conjunto de dados com informações sobre as boas práticas executadas pelo Judiciário relacionadas aos Objetivos de Desenvolvimento Sustentável (ODS). Por meio de levantamento bibliográfico se identificou na literatura as principais características presentes nos modelos para abertura de DAGs e barreiras inerentes ao processo de publicação. Utilizou-se a técnica de pesquisa em laboratório para descrever e analisar a aplicação do

modelo proposto em um ambiente controlado para fins específicos do presente trabalho. O modelo se baseou no *Data Publication Workflow*, um fluxo para publicação de dados disponível no site da iniciativa *Frictionless Data* e que aborda etapas como empacotamento, tratamento e publicação dos dados. A aplicação do modelo também utilizou ferramentas de código aberto baseados em *Frictionless Data*. Os resultados mostraram a viabilidade do modelo para a abertura de um conjunto de dados abertos governamentais, assim como demonstrou que as ferramentas disponíveis na iniciativa *Frictionless Data* contribuíram para a verificação das etapas do modelo. Concluiu-se que o modelo precisa de validação em outros contextos como, por exemplo, integração e abertura de diferentes bases públicas relacionadas aos ODS.

PALAVRAS-CHAVE: Dados abertos governamentais. Governo. *Frictionless data*. *Open data*..

JITA: IM. Open data

1 INTRODUCTION

Public organizations, especially governments, are subject to new demands from society, such as: increased transparency and participation in the management of public resources, greater control over the quality of services provided, greater accountability of public managers, among others. From these new political, economic and social demands, new forms of management and bringing governments closer to society are necessary. In the search for these new forms of management, and especially aiming at a greater approximation with the new demands of society, one of the possibilities is the use of technological tools. With the advent of current technologies, especially linked to information and communication technology (ICT) - with emphasis on the Internet, public organizations are enabling the offer of a wide range of new products and services to society (ALBANO; ARAUJO; REINHARD, 2017). In the last decade, a movement known as Government Open Data has emerged and gained importance, establishing a set of requirements to be followed by public institutions to publish public data, performing this task mainly through the use of ICT (CORREA; SOUZA; SILVA, 2019).

Access to information, as we know it today, is directly linked to the adoption of UN Resolution 59 of 1946, which presents this access as a fundamental human right and cornerstone of the freedoms on which the UN is founded. Moreover, access to information implies the right to collect, transmit and publish information (CARDOSO, 2019). In September 2011, Brazil became a member of the Open Government Partnership (OGP), a multinational initiative to promote worldwide wide adoption of Open Government Data (OGD). The commitment includes political and technical milestones and the launch of the Brazilian Open Data Portal. The initial goal was to gather large amounts of aggregated government data for digital publication and create a central catalog of information on public activity, with the intention of improving governance and monitoring government activity (BREITMAN et al., 2012).

In 2017, the United Nations made available the Government Open Data Planning Guide for Sustainable Development, inspired by the experience of Brazil, established by Decree 8.777/2016, which deals with the Open Data Policy, where it is defined that each agency and entity of the federal public administration must make its own Open Data Plan. The document contains recommendations and a step-by-step for organizations to plan actions related to open data in order to promote the Sustainable Development Goals (SDOs) of Agenda 2030.

The Global Agenda 2030 is a commitment made by leaders of 193 countries, including Brazil, and coordinated by the United Nations through the United Nations Development Programme (UNDP), pursuant to UN General Assembly Resolution A/RES/72/279.OP32 of 2018. There are 17 Sustainable Development Goals and 169 goals to be achieved from 2016 to 2030, related to the realization of human rights and promotion of development, which incorporate and give continuity to the 8 Millennium Development Goals, based on subsidies built at the Rio +20 conference (NATIONAL COUNCIL OF JUSTICE, 2018). Rio +20 was a United Nations conference on Sustainable Development held in Rio de Janeiro in 2012, 20 years after the United Nations Conference on Environment and Development, known as Rio-92 or Eco-92 (GALLOPÍN; VESSURI, 2019).

ODS is like a list of tasks to be accomplished by governments, civil society, the private sector and all citizens in the collective journey to a sustainable 2030. In the coming years of implementation of Agenda 2030, ODS and its goals will stimulate and support actions in areas of critical importance to humanity: People, Planet, Prosperity, Peace and Partnership (NATIONAL COUNCIL OF JUSTICE, 2018). Therefore, the successful implementation of ODS depends on partnerships between society, public and private organizations. To help meet the goals, UN member states can use open government data as a means to achieve ODS,

specifically ODS 16, which describes: "Promoting peaceful and inclusive societies for sustainable development, providing access to justice for all, and building effective, accountable and inclusive institutions at all levels" (Applied Economic Research Institute, 2018). An example of this use is public data as subsidies to increase transparency, public accountability and citizen participation, exposing and preventing the running and mismanagement of public resources (WORLD BANK, 2015).

Through CNJ Ordinance n. 133/2018, an Interinstitutional Committee was established to evaluate the integration of the goals of the Judiciary with the goals and indicators of the ODS, Agenda 2030 and to prepare a work report with the support of all the courts of the country, whose composition is included in CNJ Ordinance n. 148/2018. According to CNJ Ordinance n.133 /2018, the alignment of the Judiciary with the UN Agenda 2030 may represent progress in the field of the realization of the fundamental rights of citizens, because the Judiciary may provide relevant and necessary information - whose database is produced and maintained by the Judiciary itself - for the fulfillment of the goals of the ODS, which binds all people, from all countries, by translating into a Global Pact to eradicate poverty and promote decent life for all, within the limits of the planet (NATIONAL COUNCIL OF JUSTICE, 2018)

To show that it is possible this correlation of the Judiciary Power to the ODS, the Committee mapped all the CNJ Resolutions, as well as the recommendations, provisions and orientations of the National Justice Corps, in addition to the normative acts and good practices of the Superior Courts, Federal Regional Courts, Judicial Sections, State and Federal District Courts of Justice, Regional Labor Courts, Regional Electoral Courts, Military Courts, the latter set of data being used in this article to validate the model based on Frictionless Data.

Public agencies in general create, store and disseminate a wide variety of information, from demographic and economic data to family income. These data can take different formats, disseminated as primary or processed data. They may also be related to public services or internal processes (REN; GLISSMANN, 2012). For example, the judicial information made available by the sites of the Courts of Justice, accessed daily by ordinary citizens and all operators of the Law, due to its great volume and importance, needs to have good quality (ALBUQUERQUE; BASTOS; LINO, 2009). Several initiatives have arisen to solve the problem of the quality of information regarding the technical barriers that involve the publication of DAG. In 2016 an initiative called Frictionless Data, created by the Open Knowledge Foundation, appears as an alternative to remove "friction" in working with data, focusing on simplicity and a basic structure that allows anyone to publish and use data sets more easily, without going through an intermediary. In view of the above, this article proposes to answer the following question: "how can a model based on the Frictionless Data specification contribute to the opening of open government data?

The following sections of this article are divided as follows: the context of the research, related works and the theoretical foundations with main concepts about the problem of working with open government data from the perspective of Quality of Information; a brief survey on Open Data and Government Open Data; an introduction to the FAIR principles (Findable, Accessible, Interoperable and Reusable); an introduction to the fundamental concepts of this research on Frictionless Data and Data Package; a presentation of the methodology, structure of the model and its application in the context of the Brazilian judiciary; final considerations and future works.

1.1 The Problem of Quality of Information in the context of Government Open Data

In the context of public institutions, the focus of this article is on public agencies creating, storing and disseminating a wide variety of information, from demographic and economic data to family income. These data assume different formats and can be disseminated as primary data or processed. They can also be related to public services or internal processes (REN; GLISSMANN, 2012). For example, the judicial information made available by the sites of the Courts of Justice, accessed daily by ordinary citizens and all operators of the Law, due to its great volume and importance, needs to have good quality (ALBUQUERQUE; BASTOS; LINO, 2009).

According to Quarati and De Martino (2019), the dissemination of open government data is considered the driving force of economic and social growth, in addition to being an essential factor for advertising government actions. However, the high availability of data does not guarantee the quality of the information present in these public databases. The dissemination of data can show that the quality of the data on which important decisions are made is poor. Success in opening government data requires more than simply providing access to data. It also requires improving the quality of government information, creating and institutionalizing an open government culture, and providing the tools and instruments with which the data will be used. This broader perspective needs to be adopted by governments that are now just opening a portal to make the data accessible. (JANSSEN; CHARALABIDIS; ZUIDERWIJK, 2012).

However, the task of objectively defining what information quality is is not trivial. In the area of Information Science, for example, it is observed that the definition of information quality is not yet a consensus. Nehmy and Paim (1998) present an analysis on the concept of information quality, in order to "focus the discussion on the ways of approaching information quality in literature, in the effort to unveil limitations and challenges for the theoretical construction of the concept. According to the study, there is no generally accepted definition of quality of information, being a vague and subjective concept (NEHMY; PAIM, 1998 apud SCHWUCHOW 1990, P. 55).

For Paim (1996), the subject of quality of information still encounters the inherent difficulty of conceptualizing the term quality, and there is no consensus in the literature on theoretical and operational definitions of quality of information. The authors make an analysis about the existing difficulties around the quality of information and the tendency in the literature to study it under three aspects: transcendent (or philosophical, or metaphysical) quality of information; one aspect that is based on intrinsic aspects, which some authors understand to be synonymous with the transcendent, and another on contingential attributes.

Works such as Trindade and Oliveira (2007) have sought to define attributes for evaluating the quality of information in systems that make available large amounts of information that can be shared and transformed into knowledge. In the understanding of these authors, some attributes that can help evaluate the quality of information are: accessibility, clarity, completeness, conciseness, consistency, credibility, availability, packaging, ease of use, history, interface, interpretability, objectivity, accuracy, readiness, quantity, relevance, reputation, security, temporality and value.

For Bobrowski, Marré and Yankelevich (1999), the problems related to the quality of information derive from mistaken practices in software engineering. Seven attributes were raised to measure information quality: completeness, relevance, reliability, consistency, temporality, precision and concision. Other authors show the evolution on the subject of information quality and its application in the organizational context. According to Calazans (2008), the lack of information quality in an organization can have social and business impacts.

Factors such as information with multiple origins, use of subjective judgments, systematic errors in the production of information, and the large amount of information from different sources influence the quality of information (CALAZANS, 2008, apud STRONG; LEE; WANG, 1997).

In a study that carried out a survey on frameworks for information quality evaluation, twenty information quality structures were identified that define and categorize information quality criteria in different contexts (EPPLER; WITTIG, 2000). It is worth mentioning that the study adopts the framework definition assumed by (PORTER, 1991), which is considered a way to help identify relevant variables and the questions that the user must answer to develop personalized conclusions for a certain sector and company (PORTER, 1991, apud EPPLER; WITTIG, 2000).

In the next sections the discussion on quality of information will be resumed, this time from the perspective of open government data, but first concepts will be presented about what makes a data open. According to Davenport (2001), information sharing is defined as the voluntary act of making information available to others, and any information provider can add value to the information by making it accessible. With the paradigm of Information and Communication Technologies (ICT), it has become a challenge for organizations to manage large information production, that is, information needs to be managed and shared, aiming at building organizational knowledge and, consequently, promoting survival in a highly competitive market (VALENTIM, 2010).

The Open Knowledge Foundation (OKF), in the Open Data Handbook, defines open data as data that can be freely used, reused and redistributed by any person - subjects, having as its only requirement the allocation of the source and compartment by the same rules (OPEN KNOWLEDGE FOUNDATION, 2011). Some points are important related to the full definition of open are:

a) Availability and Access: the data must be available as a whole and at no greater cost than a reasonable cost of reproduction, preferably possible to be downloaded over the Internet; they must also be available in a convenient and modifiable manner;
b) Reuse and Redistribution: data must be provided under terms that allow reuse and redistribution, including combination with other data sets;
c) Universal Participation: everyone should be able to use, reuse and redistribute - there should be no discrimination against areas of activity or against persons or groups; for example, 'non-commercial' use restrictions that would prevent 'commercial' use, or use restrictions for certain purposes (e.g. educational only) exclude certain data from the concept of 'open'.

Two important qualities make data "open": data must be "legally" open, meaning that it is unrestricted by restrictive copyright terms and can be legally shared and used by individuals, companies, academia, non-profit organizations and others; data must also be open in the "technical" sense of the word, which refers to the delivery of data in a standard and well-defined format, usually via the Internet (AYRE; CRANER, 2017). Another relevant concept related to open data is interoperability, which means "the ability of several systems and organizations to work together (interoperate). That is, the ability to interoperate - or combine - different data sets. According to the Open Data Manual (2011), data is defined as open when it allows information to be published and disseminated in a way that third parties can reuse it. It is also considered that there are three laws of open data, as listed below:

1) The data needs to be found and indexed on the Web
2) They must be in open and machine-readable format
3) No restrictions on use

Open data refers to data collected and shared with others to use as desired, without restrictions on copyright or use. Traditional examples of open data include data collected by the government (e.g., weather reports, criminal incident reports, zip codes), as well as some academic sources (open access journals, research and research data, results of scientific experiments). Other examples of open data application are data made available by state and national governments, development agencies and non-governmental organizations, which are used to reduce costs and improve the services they offer to citizens (AYRE; CRANER, 2017).

From an information quality perspective, W3C has created Data on the Web Best Practices, a document that provides best practices related to the publication and use of data on the Web, created to help and support a self-sustaining ecosystem for open data. According to the guide, consisting of 35 best practices, data needs to be discovered and understandable by humans and machines, and the goal of these recommended practices is to facilitate interaction between the consumer and the publisher of open data. Also according to Data on the Web Best Practices, not all data and metadata should be shared openly, because issues such as security, commercial sensitivity and, above all, privacy of individuals need to be considered and it is up to data producers to determine which policies are appropriate and which data should be shared and under which circumstances (MANUAL..., 2011).

In view of the above, it is necessary to differentiate open data from other concepts in the literature. According to Albano and Reinhard (2014), open government, open data and open government data have different meanings:

a) **Open government:** is the provision of information in any format by governments and other actions aimed at promoting greater transparency;
b) **Open data:** is the availability of information, in some formats (pre-established), by governments, private organizations, with or without profit purposes, or by other actors of a society;
c) **Open government data:** is the provision of data by governments, in the same formats established for open data

Open government data is also characterized from three perspectives (KUMAR, 2016 apud GONZALEZ-ZAPATA; HEEKS, 2015):

a) **Government data**: Governments collect and generate a large amount of data and data sets; these data sets contain data related to, for example, social assistance payments, impact of the social assistance structure, uniquely identifiable citizen data and public service related data;
b) **Open data**: concerns about overcoming accessibility barriers, bringing uniformity to the way data are captured, stored, processed and generated; open data is any data that can be freely used, reused and redistributed by anyone;
c) **Open Government:** a government may be called Open Government if it shows a governance tendency, prioritizing the opening of its data; the opening also involves the creation of acts of right to information, more civil liberties and more transparency.

The opening of public data is intrinsically related to the creation of the Freedom of Information Act (FOIA), which can be defined as the right to access information held by public bodies, being an integral part of the fundamental right to freedom of expression, as recognized in Resolution 59 of the UN General Assembly, adopted in 1946 and by Article 19 of the

Universal Declaration of Human Rights (1948). The laws associated with FOIA reflect the fundamental premise that all information held by governments and government institutions is, in principle, public and can only be retained for reasons such as privacy and security. As for FOIA membership, in 1990, only 13 countries adopted national FOIA laws, while today there are over 90 laws adopted worldwide.

FOIA laws determine who can access government public data, what information is available, how long governments must respond, and what penalties result when the government fails to comply. In addition, implementation of FOIA laws often follows democratization and anti-corruption movements around the world, and it began with what was called the "Opening Decade," which lasted from the end of the Cold War to the terrorist attacks in the United States (FINK, 2018). Open public data is a powerful tool and also a derivative of FOIA. Without the FOIA, open data cannot deliver what they promise: true freedom of information (IZDEBSKI, 2015).

In 2009, the Obama administration outlined a set of open principles of the U.S. government that encompass three fundamental objectives: collaboration, participation, and transparency. In order to disseminate U.S. government data to the taxpayer so that it can be used to generate social and economic value, the Open Government Directive (OGD) was instituted, a set of directives that require all U.S. government agencies to publish data sets on a publicly available website, providing important context and metadata through reports, public service briefs, links to major programs, and downloadable data sets (DAWES; HELBIG, 2010).

In Brazil, according to Oliveira and Lóscio (2014), the Law on Access to Public Information (BRAZIL, 2014), enacted on November 18, 2011 and effective from May 16, 2012, allowed anyone to have access to documents and information that are under the custody of public agencies, at all powers and levels of government. The launch of the Brazilian Open Data Portal made it possible for society and government actors to interact, making it possible to locate data in open format from a single source and place through a search tool.

Several initiatives have arisen with the creation of the Access to Information Law and many are the possibilities regarding the use of open public data, such as promoting transparency regarding government actions. According to Barbalho (2018), some functions of transparency for the public sector are: trust in government and institutions; citizen participation; strategic planning; directing products and services; and exercising citizenship. Open data platforms can help create public policy, and it is a powerful approach to making data locatable, interoperable and machine-readable, dramatically improving the efficiency of analysis and insights (KLEIN; KLEIN; LUCIANO, 2018). However, in Brazil, despite the existence of quite advanced programs of public transparency - such as budget transparency - there are still very few agencies or secretariats that make open data available. At best, there is data available for visualization, but there are numerous technical and even political barriers to be reused by society in the creation of new projects and services (MANUAL..., 2011).

It is well known that open data implemented and maintained by governments and their agencies, civil society organizations and business sector offer several benefits, such as the creation of new products and services and generation of new economic opportunities (ALBANO; REINHARD, 2014). When it comes to the data market, its value is estimated to be around 40 to 140 billion euros annually (IZDEBSKI, 2015). In view of this, some examples of the use of open data for economic purposes have emerged, such as the initiative of the World Bank to institute a competition to indicate the best solution (developed application) created from its data (data that the bank publishes in open format) (ALBANO; REINHARD, 2014).

Regarding the quality of open government data, some works address the evaluation of the quality of information in the Brazilian public sector, such as Albuquerque, Bastos and Lino (2009), which presents an evaluation model for measuring the quality of information published

in portals of the courts of justice, which have vital importance for the community that accesses the judicial content on a daily basis and needs mechanisms and methodologies to assess the quality of information available on their websites.

The definitions of quality in the context of open data vary considerably. The European Data Portal, for example, considers high quality open data if "human beings can understand it and machines can manipulate it" and are aligned with the 5-Star Open Data classification system. Others, such as the G8 Open Data Charter and the Open Data Institute Certification Badges, focus on providing metadata, data scheme descriptions, use of shared data dictionaries, license used, file format and editor support to interact with data users (CORSAR; EDWARDS, 2017).

Improving the quality of access to open data can also be an essential enabler for science, technology and innovation (OECD, 2018). An example are initiatives to share research data, increasing the pace of knowledge discovery and scientific progress. Reusing research data has the potential to avoid duplication of data sets and bring new perspectives from multiple analyses of the same data set. One example is studies that use data from various sources to analyze genomic variations associated with cancer in order to select the most appropriate therapy for a specific patient (FIGUEIREDO, 2017).

In 2016, Scientific Data, a data focused publication, published the FAIR Guiding Principles for Scientific Data Management and Stewardship, which aimed to provide guidelines for improving the location, accessibility, interoperability and reuse of digital assets. The principles emphasize the ability of computer systems to find, access, interoperate and reuse data with little or no human intervention, as greater computational support is required to cope with the growing volume of data.

FAIR principles can be defined as:

> FAIR's Guiding Principles precede implementation choices and do not suggest any specific implementation technology, standard or solution; moreover, the Principles are not themselves a standard or a specification. They act as a guide for data editors and administrators to help them assess whether their specific implementation options are making their digital search artifacts locatable, accessible, interoperable, and reusable (WILKINSON *et al*., 2016).

FAIR principles advocate increased localization, accessibility, interoperability and reuse of research data and academic digital objects in general. Under the acronym FAIR, an acronym for Findable, Accessible, Interoperable and Reusable, 15 principles have been formulated to guide the actions of publishers, data administrators and other interested parties.

The central point of the FAIR concept is its application to human guided and machine-oriented activities. Furthermore, FAIR is not binary (i.e., FAIR/non FAIR), but a spectrum along which "FAIRness" variables are possible, i.e., as much as possible its application (HIGMAN; BANGERT; JONES, 2019).

It should be noted that the FAIR principles apply to data and its metadata, i.e., records on data sets. Therefore, the term "metadata" is declared in the principles. Another important aspect about the FAIR principles is that they do not only refer to open data, it is possible to work with private data, for example. The FAIR principles do not represent a quality standard for evaluation of tools, data, policies, among others, but their implementation can be a gradual and systematic adaptation of new working routines (HANSEN, 2018).

In 2010, in order to classify the degree of openness and connection of data, Tim Berners-Lee proposed five principles known as "5-star system" or 5-Star OPEN DATA principles. The principles present a scheme, which determines that the greater the number of stars, the greater the degree of openness and connection of the data (CAMPOS; VILELA, 2018).

The 5-Star and FAIR approaches have convergent points, however it is worth noting that the FAIR principles cover a wider scope and include non-open data. The 5-Star scheme is oriented to open data. Another point of distinction is related to the accessibility requirements as an essential element of reuse, being this aspect directly linked to the IT infrastructures that meet the data. In this sense, the FAIR principles are more implementation oriented than the 5-Star principles (HASNAIN; REBHOLZ-SCHUHMAN, 2018).

The next section will address the Frictionless Data initiative, which has been used to address open data publication issues.

1.2 Frictionless Data

Frictionless Data is an Open Knowledge Foundation initiative that aims to remove "friction" in working with data. It is understood as "friction" or "friction" when data consumers waste too much time and resources just to understand and work with the information. On the other hand, information producers find it difficult to describe the types of data allowed in each column of a spreadsheet, for example (OPEN KNOWLEDGE FOUNDATION, 2018).

Also according to the official website of the initiative, Frictionless Data is a progressive framework for building data infrastructure - management, integration, flow, among others, that is, unlike other frameworks, FD was designed to be adopted in an incremental and "progressive" way. The goal of the framework is to work with, develop and enhance existing data and tools used in working with data, rather than replacing them. Therefore, it is not a large database or data management system, but an approach that assists in automating repetitive tasks over data sets, such as obtaining, cleaning and using the data.

A good analogy for understanding the friction in working with data is the process of making a cake, as Rufus Polock, creator and founder of the Open Knowledge Foundantion, exemplifies: when you decide to bake, you have the ingredients readily available on the market or in your own kitchen. No need to travel to a farm, collect eggs, grind the corn, make the bacon. Instead, thanks to logistics, transportation (containers) and payment, the ingredients are available at a supermarket, without the need to pick them up at the source. The problem is that when it comes to data, there is still a great cost to obtain and work with the data. When performing an analysis or building an application, for example, depending on the set of data needed, you need to discover the data source, extract it, clean it and prepare it before you can start the work.

Frictionless Data also offers a range of tools, standards and best practices for publishing data. Frictionless Data is based on principles such as:

a) focus: concentrates on specific types of data;
b) Web oriented: uses formats that work well on the Web, such as JSON (Javascript Object Notation), a format commonly used for exchanging data by APIs (Application Programming Interface);
c) distribution: project with a view of an ecosystem without data centralization;
d) openness: anyone can freely and openly use and reuse the data;
e) use of existing tools: integration with existing technological tools;
f) simplicity: maintenance of formats and metadata in a simple and light way, easy to interpret and use (OPEN KNOWLEDGE FOUNDATION, 2019).

The central idea of the Frictionless Data initiative is the concept of "data containerization", i.e. the way the data is shared follows the concept of container as understood

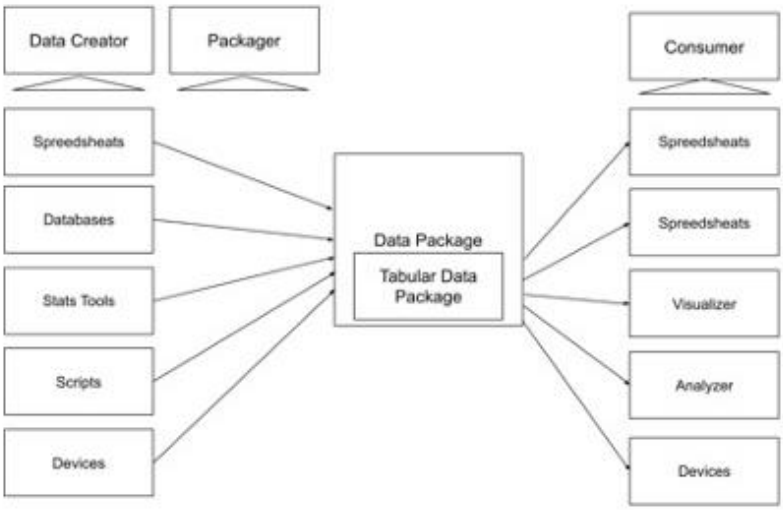
in the transportation of goods on a cargo ship, for example. The objective is to develop standards specifying the containers in order to reduce costs and time for data transportation (OPEN KNOWLEDGE FOUNDATION, 2019).

A key concept in the Frictionless Data initiative is the Data Package or data package. It is a way to simplify the grouping of data and its descriptions in a single place, in order to make data more easily shared and used. The format of the data package is very simple, friendly and extensible (OPEN KNOWLEDGE FOUNDATION, 2019).

There are several data "packaging" initiatives that share a "disk file" model to compress data with metadata. These include BagIt (Kunze et al., 2016) and CSV on the W3C Web (CSVW) (W3C, 2016). BagIt is a popular data set package format among libraries, including the Library of Congress in the U.S., designed to support disk-based storage and arbitrary digital content network transfer. BagIt focuses on archiving digital assets (primarily content), while the Data Package focuses on describing the structure of packaged data. In addition, a key element of the Data Package approach is integration with existing tools and extensibility for many data types, both of which are not priorities for BagIt. CSVW is the result of a W3C working group completed in 2016, is more directly comparable. CSVW started as an effort to standardize the specification of the then prototype Tabular Data Package with several modifications to ensure Web compatibility. (FOWLER; BARRATT; WALSH, 2018).

A Data Package, as shown in figure 1, consists of metadata that describes the structure and content of the package, as well as resources such as data files that form the package content. The metadata of the package is stored in a "descriptor", which in turn transforms a data collection into a data package. The structure of this descriptor is the main content of the data package specification. The Data Package specification does not impose any requirement on its form or structure and can therefore be used to "package" any type of data, contributing to the publication of data, for example (OPEN KNOWLEDGE FOUNDATION, 2019).

Figure 1. Data sharing process using Data Package



Source: Adapted from Open Knowledge Foundation (2019)

The process of data sharing using Data Package has, in only one layer, a basic data transport structure that considerably reduces the "friction" in data integration and distribution, supporting automation, without imposing drastic changes in the data to be packaged. The

structure, by its simple and lightweight nature, makes it easy for content publishers, data users and software tool developers to adopt (OPEN KNOWLEDGE FOUNDATION, 2019).

The next section will present the Frictionless Data based model to collaborate with the publication and sharing of open government data, as well as a contextualization on the details of the model and its application.

1.3 Related works

According to Geraldo and Dalenogare (2019), there is a consensus in the literature on the need to define a standard framework or model for government data availability in order to provide greater uniqueness and respect for certain internationally accepted criteria. According to Janssen, Charalabidis and Zuiderwijk (2012), public managers simply prefer to make the data available, without worrying about other activities involving the publication of data. However, the data can generally not be used in their raw state, and an assessment of the quality, modification and processing of this data may be required first. The above authors also state that standardization of methods and development of robust metadata can increase access to data. Another problem is the lack of feedback mechanisms that show what is done with open data. Governments that publish data sets commonly face criticism such as poor usability, lack of data feedback and inadequate metadata and improvement mechanisms (DAWES & HELBIG, 2010).

On the other hand, it is worth noting that standard setting can provide several benefits because, according to Kassen (2013), DAGs can create a favorable environment for proactive civic engagement, providing a real opportunity for independent developers to create applications using DAG sets, along with citizen participation, creating a new environment for cooperation between local government and citizens, resulting in a real transformation in the form of communication between government and society.

In order to carry out the survey of the work related to this survey, the search engines on the CAPES (Coordination for the Improvement of Higher Level Personnel) periodical portal were used. Through the use of keywords such as model, open government data, open data and Frictiotnless Data, we identified studies with proposals of models for publication of open government data. The literature survey on open government data publication showed works, such as that of Ávila (2015), which proposed a model of data publication process considering the level of maturity of the publishing institution, contributing to the strategies of production and availability of open data and connected open data of the public sector. However, despite the work considering the maturity level of the institutions and contemplating several steps in the process of publishing open data, "there was no greater technical detail (involving tools, for example) of how the publication of open data should be developed. Consequently, it does not contemplate any recommendation regarding good practices for the publication of open data of a more technical nature" (ÁVILA, 2015).

Works, such as Dzikrullah and Rinjani (2017), have developed a model called SmartGov for integration of data systems based on Big Data with focus on e-gov. The proposal is to facilitate the exchange of open government data between government agencies, creating an integrated service environment. The focus of the model was to solve data interoperability problems and, although the model contemplates several steps inherent to the publication of open data, such as a layer called Open Data Layer, which uses CKAN (Comprehensive Knowledge Archive Network), a platform developed by the Open Knowledge Foundation, it does not have a feedback step to report quality problems found in the data.

Another work presents a heuristic model to understand the relationship between open governmental data and the activities of citizens and governments from the point of view of

democratic monitoring, deliberative and participatory processes. The model adapts open data according to the type of democratic process and helps identify challenges of using open data for democracy. The work has as its strong point to consider the complexity of democratic processes, but it does not present a stage for data quality treatment, despite mentioning that the importance of information provision and data quality are crucial (RUIJER; GRIMMELIKHUIJSEN; MEIJER, 2017).

Saxena (2016) proposed a model to evaluate the behavior intention for the use of DAGs. A quantitative analysis using multiple regression was performed to analyze 341 catalogs with government data. The strong point of this work was to verify the popularity of an open data set among end users. The model proposed by Zhu and Freeman (2019) provided a framework for evaluating the performance of portals that publish DAGs, focusing on possible interactions and user experience. The strong point of the study was the evaluation of 34 portals, based on comprehensive criteria, which aimed to assess users' sense of confidence in the portals, in addition to the ability to understand the content and integrate data and forms of citizen participation. However, both models mentioned with a focus on citizen participation do not contemplate other important steps for the publication of open data, such as the treatment of published information.

In the article by Tambouris (2016), a process for publication, expansion and consumption of open government data was outlined using a multidimensional model, such as the data cube approach together with Linked Data tools. As a strong point of the work, following the proposed process, data producers can transform raw data into RDF format data cubes. This allows data consumers to find, view, merge and analyze data more easily. However, the work does not cover other key processes for publishing open government data, such as data handling and user feedback.

Another work presents a connected open data model (Linked Open Data - LOD) for a legislative open data set of the House of Representatives. The model developed in the survey was based on the actual data set provided by the House of Representatives using RDF to structure the data. As a strong point of the work, there was the contribution to areas of Information Science, such as the representation of information and knowledge, with the use of vocabularies and ontologies to represent and bring meaning to the relationships among the modeled entities. However, the research does not contemplate issues such as alignment with best practices for data publication and use of metadata (BRANDT et al., 2018).

The model proposed by Kalampokis, Tambouris and Tarabanis (2011), has two main dimensions: organizational complexity, technological and added value for data consumers. The idea was to complement existing e-government models for reuse of DAGs, with the aim of providing a roadmap for reuse and DAGs, and to allow the evaluation of relevant initiatives. The strong point of the model was the integration of data which, in addition to the two dimensions mentioned, has four phases: aggregation of government data, integration of government data with social and formal non-governmental data. However, there was no implementation of a prototype for each stage of the model, as well as the treatment of the data.

The authors Wiese et al. (2019) used Frictionless Data to solve the problem of the large amount of data provided by companies belonging to the European Energy System with insufficient documentation, besides data in inconsistent formats. The work listed some challenges that, with the use of Frictionless Data, obtained satisfactory results to solve related problems: the identification of appropriate sources, different data sources, lack of standardized nomenclature and classification of energy sources, lack of metadata, data quality (inconsistencies, errors, data gaps).

Frictionless Data has already been used in international governmental initiatives such as the Workforce Data Initiative, which aims to modernize the United States workforce through

the use of data. The goal of this initiative is to help state and local workforce councils collect, aggregate and distribute statistics on the effectiveness of training providers and programs they offer. The U.S. Department of Labor requires eligible training providers (ETP) to work with state council Workforces to track their students' results in order to receive federal funding. For this task, they are building a set of open source tools based on the Frictionless Data specification called etp-uploader, a tool that allows training providers to upload data on a single platform. There are many hundreds or thousands of training providers under each board and each must securely upload their data and, at the same time, it must be easy for the board itself to process and automatically validate the data sets. A tool like etp-uploader could be developed based on the proposed model and used to send the good practice data related to the Sustainable Development Goals of the various courts of law to the CNJ in a Data Package format (FRICTIONLESS DATA, 2017).

2 METHODOLOGY, MODEL PRESENTATION, APPLICATION AND RESULTS

According to Marconi and Lakatos (2002) a problem, to be considered relevant, needs to be analyzed from the perspective of its value. According to the authors, a research problem must be feasible, relevant, new, feasible and timely. The present work was based on a direct documentation of the approaches to the publication of open government data. The use of keywords such as model, open government data, open data and Frictiotnless Data was searched through the search mechanisms present in the CAPES (Coordination for the Improvement of Higher Education Personnel) Journals portal. Through the bibliographic survey, the main characteristics present in the models for open government data and the inherent barriers to the publication process were identified in the literature.

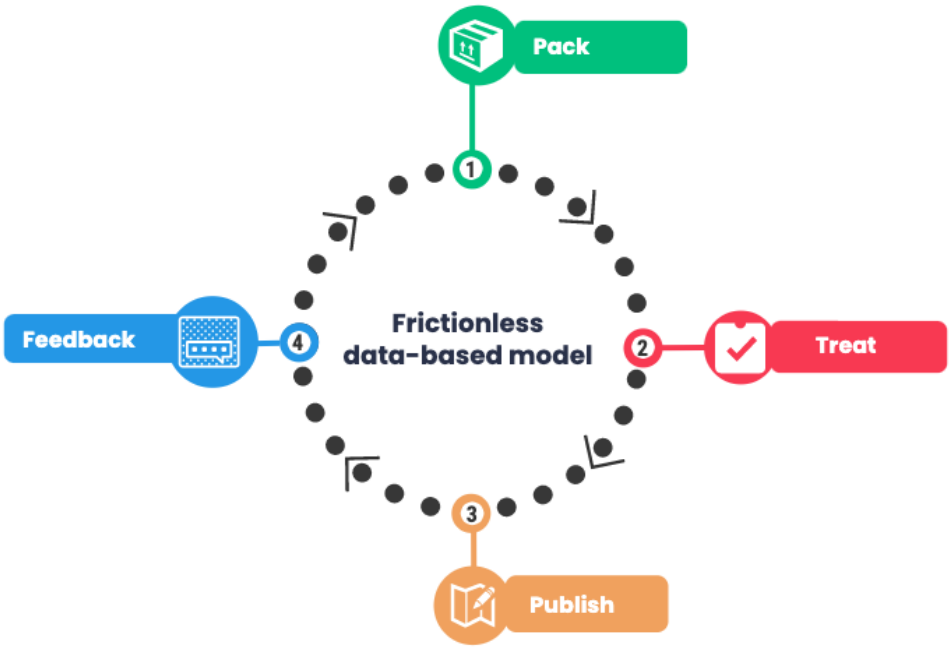
The laboratory research technique offers a more accurate result, although it is considered a more difficult research procedure. Furthermore, it requires specific, precise instruments and adequate environments (MARCONI E LAKATOS, 2003, pg. 190), which is in line with the objectives outlined in this research. For these reasons, the laboratory research technique was selected to describe and analyze the application of the proposed model in a controlled environment for specific purposes of this work, using the online tools available in the Frictionless Data initiative to conduct the experiments using the CNJ database related to the Sustainable Development Objectives and verify all the steps proposed in this model. The Frictionless Data initiative provides a set of tools, specifications and best practices to describe, publish and validate data.

The model is based, in addition to studies of related works, on the FAIR principles mentioned earlier in the article to support the main interoperability features; the Frictionless Data Spec, a set of specifications for use and publication of data using Fricionless Data; the requirements for open data process proposed by Zuiderwijk, Janssen and Jeffery; the key elements for creating an open data ecosystem (ZUIDERWIJK; JANSSEN; DAVIS, 2014) and adherence to best practices for open data publishing.

The steps of the model were outlined in order to respond to the problem outlined in this article. In other words, the model aims to assist the publication of open government data in order to reduce the time spent working with open government data, as well as to promote the quality of information published and provide a basic structure that allows any citizen to publish and use data sets more easily and without going through an intermediary.

Figure 2 shows the steps of the Frictionless Data based model for publishing open government data:

Figure 2. Frictionless Data based model for publishing government data



Source: Prepared by the authors

In the proposed model, the Pack step refers to the transformation of a data set into a data package in order to allow interoperability and data reuse. This step is based on the Frictionless Data specification, which contains the concept called Data Package, explained in previous sections with more details. A Data Package consists of metadata, which describes the structure and content of the data package, and resources such as data files, which form the contents of the package. The Data Package metadata is stored in a descriptor. This descriptor is the main content of the Frictionless Data specification. In addition to the descriptor, a Data Package will include other resources such as data files. One of the strengths of the specification is that the Data Package does not impose any requirement on its form or structure and can therefore be used to package any type of data.

As public agencies make a large amount of information available and these databases are usually in different formats or in formats that make it impossible for stakeholders to access the data (ARAUJO et al., 2012), this step helps in making the data available regardless of format. In addition, the step meets the requirement of Zuiderwijk, Janssen and Jeffery (2014), "binding and combining data". Another advantage of this step is the alignment with the 15 FAIR principles, which advocate increased localization, accessibility, interoperability and data reuse. Although the initial focus of FAIR principles is research data and academic digital objects in general, they can be used for any type of data. Therefore, it is a relevant contribution of this research to the unprecedented application of FAIR principles to open government data through the Frictionless Data model.

The results resulting from the verification of the model showed that the "Packaging" step allowed the transformation of the good practice base related to the Sustainable Development Objectives into a data package, a concept called Data Package, within the Frictionless Data specification, through the use of the Data Package Creator tool. Data Package

Creator is an online service that facilitates the creation and editing of a data package. The service automatically generates a datapackage.json file, as it is possible to add and edit data that is part of the database (FRICTIONLESS DATA, 2017). The Packing step of this research made the data available regardless of the format, since Frictionless Data does not impose any requirement on its form or structure, and can be used to package any type of data. Therefore, the "Packaging" step is considered to be the heart of the proposed model and the results showed that the goal of the step was achieved in a satisfactory manner.

Another relevant issue in the context of public agencies is the quality of information present in open government databases. The dissemination of DAGs by public institutions has continued at a very fast pace, however, the dissemination of data without adequate quality control hinders their reuse and negatively affects civic participation (VETRÒ et al., 2016).

Research shows that there are several interpretations and different ways to measure data quality, as previously commented in this work. Generally, approaches include data interpretability, accuracy, publication opportunity, reliability, accessibility, discoverability and processing capacity, or integrity. However, since people use data for different purposes, certain data qualities are more important to one group of users than to others. Many data quality metrics are user-focused, but it is critical that the government, as a data producer, understands, monitors and improves the inherent quality of the data it produces. Open Knowledge International targets data producers and data quality issues, primarily through the Frictionless Data initiative. In this research, proposing the model based on the specifications contained in the Frictionless Data initiative, suggests that the "Treat" step contemplates essential quality aspects for tabular data files contained in the Data Quality Spec, a project supported by Open Knowledge International. Finally, for this step it is recommended to meet the data quality requirement - contemplating dimensions of information quality such as concision, consistency and accuracy (TRINDADE; OLIVEIRA, 2007).

The results presented in this step suggest that the "Treat" step contemplated essential dimensions of information quality such as concision, consistency and precision, when performing the validation of the data set, verifying information quality problems such as blank headers, duplicate lines and data with wrong type.

The "Publish" stage suggests that fundamental aspects for publication and availability of DAGs should be considered. It is intended that the following requirements proposed by Zuiderwijk, Janssen and Jeffery be met regarding the processes required to publish DAGs: access, search, navigation, upload, download and visualization. For this purpose, the DataHub tool, a data management platform created by Open Knowledge International, based on the CKAN data management system, was used in this survey. Datahub provides free access to many of CKAN's main features, allowing data search, registration of published data sets, creation and management of data set groups and monitoring of data set and interest group updates. These features are available through a Web interface or API, the CKAN API.

The results of the "Publish" step of this model sought to meet the main challenge: open government data has no value in itself but when it allows users to use it effectively. Therefore, it is necessary to make the data available in open format and with quality. During the present research, after checking the "Pack" and "Treat" steps, exposed in previous sections, it was possible to publish the data through the Data.io tool, belonging to the Fricionless Data initiative. The "Publish" stage fulfilled its objective by contemplating fundamental aspects for publication and availability, by allowing access, upload and download of the CNJ's database of good practices related to the Sustainable Development Objectives. However, it is worth noting that requirements such as visualization were not met with the use of the Data.io tool, and it will be necessary in the future to create such functionalities in order to verify such important requirements for the publication of open government data.

The "Feedback" step, proposed in the present model, was established with the objective of filling the gap found in the literature regarding the existing models for publishing and opening open data, by adding participation and feedback on the publications. The intention of the stage was a possible contribution through the creation of feedback cycles, from which the government can learn from the public and promote improvements in the quality of open data. During the application of the model, it was not possible to verify the "Feedback" step using tools from the Frictionless Data initiative, since the specification does not have a tool that pays attention to the data feedback cycle. However, the "Feedback" step in this model was designed without the intention of determining a specific tool or fixed process, in order to make the choice of technology or tool more flexible as required by the public body. Therefore, the absence of a tool in the Frictionless Data initiative that allows monitoring the feedback cycle does not compromise the use of the model, since the presence of the stage meets the gap identified in the literature, and becomes an opportunity for future development of a tool by the community.

3 FINAL CONSIDERATIONS AND FUTURE WORK

Given the growth in recent years in the amount of data produced by public organizations and the divergence present in the format of this data available for consultation by citizens, which makes it difficult to work with the data and their sharing, in addition to the low quality of open government data described in the literature, with the possible consequences of undermining popular participation and weakening democracy, that this article proposed to answer the question "how a model based on the Frictionless Data specification can contribute to the opening of open government data? In the course of this research, a survey was carried out on national and international scientific production bases, as well as a search for good practices for the opening of open governmental data. Concepts such as quality of information from the perspective of Information Science, open governmental data, FAIR and Frictionless Data principles were addressed in the present work, with the objective of supporting the proposition of the model to help the publication and sharing of open governmental data.

The work related to the context of this research showed a consensus on the need for a framework or model for government data availability, with the objective of bringing greater uniqueness and respect to certain internationally accepted criteria. In addition, in the Packaging stage of this model, with the use of the Data Package Creator tool, by standardizing methods and developing metadata, for example, one can increase access to data. Another problem presents in the data opening scenario is the lack of feedback that shows what is done with the open data. A gap was noticed related to the possibility of feedback to report problems in data quality, as well as making it possible for citizens to send suggestions for improvements in government open data sets.

The results of this work were considered positive and led to the presentation of the model and respective verification of the steps. One of the points of difference regarding the models for publication of open governmental data found in the literature is the use of the Frictionless Data specification as the basis for an open governmental data model in the context of the Brazilian judiciary, as well as the inclusion of the "Feedback" stage to fill the gap found in the other models.

Another possibility as future works would be the use of Data Package Pipelines, a flow-based declarative structure for the construction of tabular data processing pipelines. Pipelines are defined as a form of task parallelization in a system that is executed in a decomposed way, that is, it allows a task flow to divide its work steps into units or stages and

execute them simultaneously. Each stage of the pipeline is normally the input of its queue, which is the output queue of the previous stage, and after processing the data, serves input of the next stage. In the case of Data Package Pipelines, it could be used for all extraction, transformation and loading (ETL) tasks as a framework to facilitate the implementation of the proposed model, as it is particularly suitable for working with diverse and heterogeneous data sources of variable and unknown quality. There are many tools and structures to do ETL work with data, however, the Data Package Pipelines focus here to organize and control confusing data from various sources and use the frictionless data tools to transform this data into a clean and consistent flow of data (FRICTIONLESS DATA, s.d.).

Still on other possible contributions, this work can collaborate, for example, in a similar way with another initiative that uses ODS for monitoring results, the Citizen PPA Portal, which was designed to allow citizens to monitor the implementation of goals and objectives of the Multiyear Plan of the Federal Government and the Objectives of Sustainable Development (ODS), being a partnership of the Federal Government with the Open Government Partneship (OGP) and coordinated by the General Warden of the Union (BRAZIL, 2018). Therefore, the applicability and generalization of this research has great potential for scientific and social contribution.

The proposed model remains open to new applications and suggestions for improvements, especially with regard to its applicability in other contexts such as, for example, integration and opening of public bases related to the Objectives of Sustainable Development. The improvement of the "Feedback" stage, with the use of other tools and the addition of methodologies such as Net Promoter Score, could contribute to measuring citizen satisfaction in relation to published open government data, and could allow governments to improve data quality. NPS was presented by Reichheld (2003), a simple way to measure customer loyalty. NPS is based on the following question: How likely are you to recommend a brand or company to a friend or colleague? Rated on a scale from 0 to 10. Customers are grouped into 'promoters' (rating from 9-10), 'passively satisfied' (rating from 7-8) and 'detractors' (rating from 0-6). According to Reichheld (2003), promoters are considered loyal customers who probably continued to use the service.

This research proposal may also, by contributing to the reuse and sharing of data, allow the construction of models for recognition of DAG standards and, in the case of the use of the judiciary's good practice bases, allow the automatic classification of good practices and sustainable development issues, being its application possible in several areas, contributing to the implementation of Agenda 2030.

REFERENCES

ALBANO, Cláudio Sonaglio. **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. Available on: https://teses.usp.br/teses/disponíveis/12/12139/tde-03062014-170642/publico/ClaudioSnaglioAlbanoVC.pdf. Access at: 20 set. 2020.

ALBANO, Claudio Sonaglio; ARAUJO, Marcelo Henrique de; REINHARD, Nicolau. Fatores motivadores e facilitadores dos relacionamentos em redes: como os gestores públicos reconhecem esses fatores em dados governamentais abertos. **Navus - Revista de Gestão e Tecnologia**, Florianópolis, v. 7, n. 1, p. 73–92, jan. 2017.

ALBUQUERQUE, Almir dos Santos; BASTOS, Rogério Cid; LINO, Manuel Rosa de Oliveira. Qualidade da informação em portais do judiciário. **Encontros Bibli**: revista eletrônica de biblioteconomia e ciência da informação, Florianópolis, v. 14, n. 27, p. 115-134, maio 2009. Available on: https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2009v14n27p115. Access at: 03 set. 2020.

ARAUJO, Marcelo Henrique de *et al*. Dados governamentais abertos. *In*: ENCONTRO DA ASSOCIAÇÃO NACIONAL DE PÓS-GRADUAÇÃO E PESQUISA EM ADMINISTRAÇÃO, 36., 2012, Rio de Janeiro. **Anais do EnANPAD.** Rio de Janeiro: Enanpad, 2012. p. 1-16.

ÁVILA, Thiago. Uma proposta de modelo de processo para publicação de dados abertos conectados governamentais. **Comunidade Áreas de Integração,** nov. 2015. Available on: http://areasdeintegracao.blogspot.com.br/2017/09/preparacao-de-partes-interessadas-para.html. Access at: 25 ago. 2020.

AYRE, Lori Bowen; CRANER, Jim. Open data: what it is and why you should care. **Public Library Quarterly**, v. 36, n. 2, p. 173–184, 2017. Available on: https://doi.org/10.1080/01616846.2017.1313045. Acesso em 10 set. 2020.

BARBALHO, Fernando Almeida. A emergência do campo de políticas públicas de dados abertos governamentais no Brasil. **Conhecer:** debate entre o público e o privado, Fortaleza, v. 8, n. 20, p. 118–137, 2018.

BOBROWSKI, Monica; MARRÉ, Martina; YANKELEVICH, Daniel. A homogeneous framework to measure data quality. *In*: CONFERENCE ON INFORMATION QUALITY, 4., 1999, Massachusetts. **Proceedings [...].** Massachusetts: Mit Sloan School of Management, 1999. p. 115-124.

BRANDT, Mariana Baptista *et al.* Modelo de dados abertos conectados para informação legislativa. **Informação & Sociedade**: estudos, João Pessoa, v. 28, n. 2, p. 149–161, 2018.

BRASIL. **Lei nº 12.965**, de 23 de abril de 2014. Publicada no DOU de 24.4.2014. Diário Oficial, Brasília, DF. 2014. Available on: https://www.planalto.gov.br/ccivil_03/_ato2011-2014/2014/lei/l12965.htm. Access at: 09 nov. 2020.

BREITMAN, Karin *et al*. Open government data in Brazil. **IEEE Intelligent Systems**, v. 27, n. 3, p. 45–49, 2012.

CALAZANS, Angélica Toffano Seidel. Qualidade da informação: conceitos e aplicações. **Transinformação**, Campinas, v. 20, n. 1, p. 29–45, 2008.

CARDOSO, Paulo Henrique. **Ciência de dados aplicada a dados governamentais abertos sob a ótica da Ciência da Informação**. 2019. 110 f. Dissertação (Mestrado) - Curso de Mestrado em Ciência da Informação, Escola de Comunicação, Universidade Federal do Rio de Janeiro; Instituto Brasileiro de Informação em Ciência e Tecnologia, Rio de Janeiro, 2019. Available on: http://ridi.ibict.br/handle/123456789/1039. Access at: 15 ago. 2020.

CONSELHO NACIONAL DE JUSTIÇA. **Agenda 2030 no Poder Judiciário**. 2018. Available on: http://novo.more.ufsc.br/homepage/inserir_homepage. Access at: 26 ago. 2020.

CORREA, Andreiwid Sheffer; SOUZA, Raul Mendes de; SILVA, Flavio Soares Correa da. Towards an automated method to assess data portals in the deep web. **Government Information Quarterly**, v. 36, n. 3, p. 412–426, 2019. Available on: https://doi.org/10.1016/j.giq.2019.03.004. Access at: 05 set. 2020.

CORSAR, David; EDWARDS, Peter. Challenges of open data quality: More than just license, format, and customer support. **Journal of Data and Information Quality**, v. 9, n. 1, p. 10–12, 2017.

DAVENPORT, Thomas H. **Ecologia da informação**. São Paulo: Campus, 2001.

DAWES, Sharon S.; HELBIG, Natalie. Information Strategies for Open Government: challenges and prospects for deriving public value from government transparency. *In*: INTERNATIONAL CONFERENCE ON ELECTRONIC GOVERNMENT, 9., 2010, Lausanne, Suíça. **Proceedings [...].** Lausanne, Suíça: Ifip Wg, 2010. p. 50-60. Available on: https://hal.inria.fr/hal-01056566/document. Access at: 05 set. 2020.

DZIKRULLAH, Fahmi; RINJANI, Muhammad Angga. A framework design to develop integrated data system for smart e-government based on big data technology. **Bulletin of Social Informatics Theory and Application**, v. 1, n. 2, p. 41–51, 2017.

EPPLER, Martin; WITTIG, Dörte. Conceptualizing information quality: a review of information quality frameworks from the last ten years. *In*: CONFERENCE ON INFORMATION QUALITY, 5., 2000, Cambridge. **Proceedings [...].** Massachusetts: Mit, 2000. p. 1-14.

FIGUEIREDO, Ana Sofia. Data Sharing : convert challenges into opportunities. **Front Public Health,** Londres, v. 5, p. 1–6, dec. 2017. Disponível em : https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5723929/. Access at: 15 set. 2020.

FINK, Katherine. Opening the government's black boxes: freedom of information and algorithmic accountability. **Information Communication and Society**, v. 21, n. 10, p. 1453–1471, 2018.

FOWLER, Dan; BARRATT, Jo; WALSH, Paul. Frictionless Data: making research data quality visible. **International Journal of Digital Curation**, v. 12, n. 2, p. 274–285, 2018.

FRICTIONLESS DATA (ed.). **Center for Data Science and Public Policy, Workforce Data Initiative**. 2017. Available on: https://frictionlessdata.io/blog/2017/08/15/center-for-data-science-and-public-policy-workforce-data-initiative/. Access at: 25 ago. 2020.

FRICTIONLESS DATA (ed.). **Data package pipelines**: a framework for processing data packages in pipelines of modular components. [s.d.]. Available on: https://frictionlessdata.io/tooling/data-package-pipelines/#check-it-out. Access at: 25 ago. 2020.

GALLOPÍN, Gilberto; VESSURI, Hebe. Science for sustainable development. **Interfaces between Science and Society**, v. 336, n. June 2012, p. 35–51, 2019.

GERALDO, Lucas; DALENOGARE, Campos. Abordagens teóricas em dados governamentais abertos. **Revista Gestão e Tecnologia**, v. 19, n. 5, p. 296-314, 2019.

HASNAIN, Ali; REBHOLZ-SCHUHMAN, Dietrich. Assessing FAIR data principles against the 5-star open data principles. *In*: EXTENDED SEMANTIC WEB CONFERENCE, 15., 2018, Grécia. **The Semantic Web**: ESWC 2018 Satellite Events. Grécia: Eswc, 2018. p. 469-477. Available on: https://www.researchgate.net/publication/326757298. Access at: 07 set. 2020.

HANSEN, Karsten Kryger; BUSS, Mareike; HAAHR, Lea Sztuk. **A fairy tale:** a fake story in a trustworthy guide to the fair principles for research data. [s.l.] FAIR Projetct, 2018. Available on: http://doi.org/10.5281/zenodo.2248200. Access at: 03 set. 2020.

HIGMAN, Rosie; BANGERT, Daniel; JONES, Sarah. Three camps , one destination : the intersections of research data management, FAIR and open. **Insights**, v. 32, n. 1, maio 2019. Available on: http://doi.org/10.1629/uksg.468. Access at: 23 ago. 2020.

INSTITUTO DE PESQUISA ECONÔMICA APLICADA. **Agenda 2030**: ODS - Metas Nacionais dos Objetivos de Desenvolvimento Sustentável. Brasília, DF: IPEA, 2018. Available on: http://www.ipea.gov.br/portal/index.php?option=com_content&view=article&id=33895&Itemid=433. Acesso em 02 set. 2020.

IZDEBSKI, Krzysztof. **Transparency and open data principles:** why they are important and how they increase public participation and tackle corruption.Varsóvia, Polônia: [s.n.], 2015. Available on: https://transparencee.org/wp-content/uploads/2015/12/open-data-principles-by-krzysztof-izdebski.pdf. Acesso em 15 set. 2020.

JANSSEN, Marijn; CHARALABIDIS, Yannis; ZUIDERWIJK, Anneke. Benefits, adoption barriers and myths of open data and open government. **Information Systems Management**, v. 29, n. 4, p. 258–268, 2012.

KALAMPOKIS, Evangelos; TAMBOURIS, Efthimios; TARABANIS, Konstantinos. Open Government Data: a stage model. *In*: INTERNATIONAL CONFERENCE, EGOV, 10., 2011, Delft, Holanda. **Electronic Government, Proceedings.** Berlim, Alemanha: Springer, 2011. p. 235-246. Available on: https://doi.org/10.1007/978-3-642-22878-0_20. Access at: 10 set. 2020.

KASSEN, Maxat. A promising phenomenon of open data: a case study of the Chicago open data project. **Government Information Quarterly**, v. 30, n. 4, p. 508–513, 2013. Available on: http://dx.doi.org/10.1016/j.giq.2013.05.012. Acesso em 25 ago. 2020.

KLEIN, Rodrigo Hickmann; KLEIN, Deisy Barbiero; LUCIANO, Edimara Mezzomo. Open Government Data: concepts, approaches and dimensions over time. **Revista Economia & Gestão**, Belo Horizonte, v. 18, n. 49, p. 4–24, 2018.

KUMAR, V. Open Government Data Initiatives: a comparison of India and China. **Research Gate**, n. March, 2016. Available on: https://www.researchgate.net/profile/Vinit_Kumar9/publication/305768855_Open_Government_Data_Initiatives_a_comparison_of_India_and_China/links/57a05a6408aece1c72171d69/Open-Government-Data-Initiatives-a-comparison-of-India-and-China.pdf.

MANUAL dos dados abertos: governo. [s.l.]: Laboratório Brasileiro de Cultura Digital; W3C; Cgi.Br, 2011. (Dados Abertos Governamentais). Traduzido e adaptado de opendatamanual.org. Available on: https://www.w3c.br/pub/Materiais/PublicacoesW3C/Manual_Dados_Abertos_WEB.pdf. Access at: 25 ago. 2020.

NEHMY, Rosa Maria Quadros; PAIM, Isis. A desconstrução do conceito de "qualidade da informação". **Ciência da Informação**, Brasília, DF, v. 27, n. 1, p. 36–45, 1998.

OLIVEIRA, Lairson E. Rodrigues de Alencar; LÓSCIO, Bernadette Farias. Uma abordagem para captura de informações sobre aplicações que fazem uso de dados abertos. **Revista Brasileira de Administração Científica**, Aracaju, v. 5, n. 2, p. 127–140, 2014.

OPEN KNOWLEDGE FOUNDATION. What is open data? **Opend data handbook,** [2011]**.** Available on: https://opendatahandbook.org/guide/en/what-is-open-data/. Access at: 25 ago. 2020.

PAIM, Isis. Problematização do conceito" Qualidade" da Informação. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 1, n. 1, p. 111–119, 1996. Available on: http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/8. Access at: 28 ago. 2020.

PAIM, Isis; NEHMY, Rosa Maria Quadros; GUIMARÃES, Cesar Geraldo. Problematização do conceito "Qualidade" da informação. **Perspectivas em Ciência da Informação**, [s.l.], v. 1, n. 1, nov. 2007. ISSN 19815344. Available on: http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/8/27. Access at: 22 set.. 2020.

PORTER, Michael Eugene. Towards a dynamic theory of strategy. **Strategic Management Journal,** v. 12, p. 95-117, 1991.

QUARATI, Alfonso; DE MARTINO, Monica. Open government data usage: a brief overview. *In*: INTERNATIONAL DATABASE APPLICATIONS & ENGINEERING SYMPOSIUM, 23., 2019, Atenas, Grécia. **Proceedings [...].** Atenas, Grécia: Ideas, 2019. p. 1-8. Available on: https://doi.org/10.1145/3331076.3331115. Access at: 28 ago. 2020.

REICHHELD, F. Frederick. The one number you need to grow. **Harvard Business Review**, v.81, n.12, 46–57, 2013.

REN, Guang-Jie; GLISSMANN, Susanne. Identifying information assets for open data: the role of business architecture and information quality. *In*: INTERNATIONAL CONFERENCE ON COMMERCE AND ENTERPRISE COMPUTING, 14., 2012, Washington, Dc. **Proceedings [...].** Washington, Dc: Ieee Computer Society, 2012. p. 94-100.

RUIJER, Erna; GRIMMELIKHUIJSEN, Stephan; MEIJER, Albert. Open data for democracy: developing a theoretical framework for open data use. **Government Information**

Quarterly, v. 34, n. 1, p. 45–52, 2017. Available on: http://dx.doi.org/10.1016/j.giq.2017.01.001. Access at: 25 ago. 2020.

SAXENA, Stuti. Open Government Data (OGD) usage in India: a conceptual framework using TOE & UTAUT frameworks. **Socrates**, Kanpur City, India, v. 4, n. 3, p. 124–144, 2016.

TAMBOURIS, Efthimios. Multidimensional Open Government Data. **JeDEM - eJournal of eDemocracy and Open Government**, Krems, Áustria, v. 8, n. 3, p. 1–11, 2016.

TRINDADE, Ana Lucia Batista; OLIVEIRA, Mírian. Atributos Para Avaliação da Qualidade da Informação em Sistemas de Gestão do Conhecimento. *In*: ENCONTRO DE ADMINISTRAÇÃO DA INFORMAÇÃO, 1., 2007, Florianópolis. **Anais [...].** Maringá, Pr: Anpad, 2007. p. 1-14. Available on: http://www.anpad.org.br/admin/pdf/enadi318.pdf. Access at: 15 set. 2020.

VALENTIM, Marta. **Gestão, mediação e uso da informação**. São Paulo: Unesp, 2010.

VETRÒ, Antonio et al. Open data quality measurement framework: Definition and application to Open Government Data. **Government Information Quarterly**, v. 33, n. 2, p. 325–337, 2016.

VILELA, Allyson Bruno Campos Barros. **OpenData Processor**: uma ferramenta para a automatização do processo de extração e publicação de dados abertos. 2018. 97 f. Dissertação (Mestrado) - Curso de Engenharia de Software, Programa de Pós-Graduação em Engenharia de Software, Universidade Federal do Rio Grande do Norte, Natal, 2018. Available on: https://repositorio.ufrn.br/jspui/bitstream/123456789/26014/1/OpenDataprocessorferramenta_Vilela_2018.pdf. Access at: 28 ago. 2020.

WIESE, Frauke *et al*. Open power system data – frictionless data for electricity system modelling. **Applied Energy**, v. 236, p. 401-409, fev. 2019. DOI:10.1016/j.apenergy.2018.11.097. Available on: https://www.sciencedirect.com/science/article/pii/S0306261918318130. Access at: 20 ago. 2020.

WILKINSON, Mark D. *et al*. Comment: The FAIR Guiding Principles for scientific data management and stewardship. **Scientific Data**, v. 3, 2016.

W3C. CSV on the Web: a primer. W3C Working Group Note, 25 feb. 2016. Available on: https://www.w3.org/TR/tabular-data-primer/. Access at: 25 ago. 2020.

ZHU, Xiaohua; FREEMAN, Mark Antony. An evaluation of U.S. municipal open data portals: A user interaction framework. **Journal of the Association for Information Science and Technology**, v. 70, n. 1, p. 27–37, 2019.

ZUIDERWIJK, Anneke et al. Socio-technical Impediments of Open Data. **Electronic Journal of e-Government**, v. 10, n. 2, p. 156–172, 2012. Available on: http://www.ejeg.com/issue/download.html?idArticle=255. Access at: 25 ago. 2020.

ZUIDERWIJK, Anneke; JANSSEN, Marijn; DAVIS, Chris. Innovation with open data: Essential elements of open data ecosystems. **Information Polity**, v. 19, n. 1–2, p. 17–33, 2014.

Article submitted in the similarity system

Submitted: 09/11/2020 – **Accepted**: 13/11/2020 – **Published**: 20/11/2020