

**UNIVERSIDADE DO ESTADO DE SANTA CATARINA – UDESC
CENTRO DE CIÊNCIAS HUMANAS E DA EDUCAÇÃO – FAED
MESTRADO PROFISSIONAL EM GESTÃO DE UNIDADES DE INFORMAÇÃO**

MELISSA FIGUEIRA FAGUNDES

**MODELO BASEADO EM *FRICITIONLESS DATA* APLICADO AOS DADOS
ABERTOS GOVERNAMENTAIS**

**FLORIANÓPOLIS
2020**

MELISSA FIGUEIRA FAGUNDES

**MODELO BASEADO EM *FRictionless DATA* APLICADO AOS DADOS
ABERTOS GOVERNAMENTAIS**

Dissertação apresentada ao Programa de Pós-Graduação em Gestão da Informação, da Universidade do Estado de Santa Catarina, como requisito parcial para obtenção do grau de Mestre em Gestão de Unidades de Informação.

Orientador: Prof. Dr. Divino Ignácio Ribeiro Junior.

**FLORIANÓPOLIS
2020**

**Ficha catalográfica elaborada pelo programa de geração automática da
Biblioteca Setorial do FAED/UDESC,
com os dados fornecidos pelo(a) autor(a)**

Fagundes, Melissa Figueira

Modelo baseado em frictionless data aplicado aos dados
abertos governamentais / Melissa Figueira Fagundes. -- 2020.
113 p.

Orientador: Divino Ignácio Ribeiro Junior
Dissertação (mestrado) -- Universidade do Estado de
Santa Catarina, Centro de Ciências Humanas e da Educação,
Programa de Pós-Graduação Profissional em Gestão de
Unidades de Informação, Florianópolis, 2020.

1. Framework. 2. Frictionless Data. 3. Dados Abertos
Governamentais. 4. Boas Práticas. 5. Judiciário. I. Ribeiro
Junior, Divino Ignácio. II. Universidade do Estado de Santa
Catarina, Centro de Ciências Humanas e da Educação,
Programa de Pós-Graduação Profissional em Gestão de
Unidades de Informação. III. Título.

MELISSA FIGUEIRA FAGUNDES

**MODELO BASEADO EM *FRictionless DATA* APLICADO AOS DADOS
ABERTOS GOVERNAMENTAIS**

Dissertação apresentada ao Programa de Pós-Graduação em Gestão da Informação, da Universidade do Estado de Santa Catarina, como requisito parcial para obtenção do grau de Mestre em Gestão de Unidades de Informação.

Orientador: Prof. Dr. Divino Ignácio Ribeiro Junior.

BANCA EXAMINADORA

Drº. Divino Ignácio Ribeiro Junior
Universidade do Estado de Santa Catarina

Membros:

Drª. Ana Maria Pereira
Universidade do Estado de Santa Catarina

Ms. Daniel Lacet de Faria Fireman
Instituto Federal do Alagoas

Florianópolis, ____ de _____ de 2020.

RESUMO

O presente trabalho discute como um modelo baseado em *Frictionless Data* (FD) pode auxiliar na publicação de dados abertos governamentais (DAGs). FD é uma iniciativa da *Open Knowledge Foundation*, que pretende remover o "atraito" 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, além do Modelo de Referência para publicação de dados abertos do governo brasileiro. 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: *Framework. Frictionless Data. Dados Abertos Governamentais. Boas Práticas. Judiciário.*

ABSTRACT

This work discusses how a model based on *Frictionless Data* (FD) can assist in the publication of open government data (DAGs). FD is an initiative of the *Open Knowledge Foundation*, which aims to remove the "friction" in working with the data, that is, when a lot of time and resources is lost to understand and work with the data. Within public institutions, the absence of standards for publishing and processing DAGs is a common problem when it comes to opening DAGs. The study sought to contribute by applying the model to the data set with information on good practices carried out by the Judiciary related to the *Sustainable Development Goals* (SDGs). Through a bibliographic survey, the main characteristics present in the models for opening DAGs 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 research. The model was based on the Data Publication Workflow, a flow for publishing data available on the website of the Frictionless Data initiative and which addresses steps such as packaging, processing and publication of data, in addition to the Reference Model for publishing open data from the Brazilian government. The application of the model also used open source tools based on *Frictionless Data*. The results showed the feasibility of the model for opening a set of open government data, as well as demonstrating that the tools available in the *Frictionless Data* initiative contributed to the verification of stages of the model. It was concluded that the model needs validation in other contexts, for example, integration and opening of different public bases related to the SDGs.

Keywords: *Framework. Frictionless Data. Open Government Data. Good Practices. Judiciary.*

LISTA DE QUADROS

Quadro 1 - Algumas manifestações conceituais de informação na trajetória da Ciência da Informação	211
Quadro 2 - Dimensões e atributos da qualidade da informação	24
Quadro 3 - Melhores práticas para abertura de dados	311
Quadro 4 - Formatos recomendados para publicação de DAGs	355
Quadro 5 - Iniciativas que fazem uso de DAGs	388
Quadro 6 - Barreiras e possíveis fatores inibidores	3939
Quadro 7 - Possíveis benefícios	411
Quadro 8 - Guia dos princípios FAIR	433
Quadro 9 - Definição de cada uma das cinco estrelas dos dados abertos	466
Quadro 10 - Mapeamento dos princípios FAIR com esquema 5-Star	4747
Quadro 11 - Comparativo entre os atributos para avaliar a QI, princípios FAIR e ..	488
Quadro 12 - Estudos de caso com aplicações de <i>Frictionless Data</i>	511

LISTA DE ILUSTRAÇÕES

Figura 1 - <i>Framework</i> conceitual de qualidade de dados	28
Figura 2 - Ecossistema de dados abertos.....	311
Figura 3 - Fundamentos dos DAGs.....	333
Figura 4 - Relacionamento entre FAIR e dados abertos.....	455
Figura 5 - 5-Star <i>Open Data</i>	466
Figura 6 - Modelo baseado em <i>Frictionless Data</i> para publicação de DAGs.....	611
Figura 7 - Modelo ideal de processo para publicação de DAGs	666
Figura 8 - Objetivos de desenvolvimento sustentável da ONU	70
Figura 9 - Assuntos mais demandados por ODS em 2018.....	72
Figura 10 - Base de dados com boas práticas dos tribunais relacionados aos ODS	75
Figura 11 - Ferramenta <i>Data Package Creator</i>	77
Figura 12 - Página inicial da ferramenta <i>Goodtables</i>	78
Figura 13 - Página inicial da ferramenta <i>DataHub</i>	79
Figura 14 - Ferramenta <i>Data Package Creator</i>	82
Figura 15 - <i>Preview</i> do arquivo datapackage.json	84
Figura 16 - Resultado da criação do arquivo datapackage.json	86
Figura 17 - Ferramenta <i>Goodtables</i> com resultado da validação do arquivo datapackage.json	88
Figura 18 - Instalação da ferramenta Datahub.io em ambiente de validação.....	89
Figura 19 - Publicação do arquivo datapackage.json na ferramenta Datahub.io.....	90
Figura 20 - Sítio para <i>download</i> do arquivo datapackage_json	91

LISTA DE ABREVIATURAS E SIGLAS

BBPODS	Bases de Boas Práticas do Judiciário relacionadas aos Objetivos de Desenvolvimento Sustentável
BI	<i>Business Intelligence</i>
CI	Ciência da Informação
CKAN	<i>Comprehensive Knowledge Archive Network</i>
CNJ	Conselho Nacional de Justiça
DAGs	Dados Abertos Governamentais
EUA	Estados Unidos da América
FAIR	<i>Findable, Accessible, Interoperable e Reusable</i>
FOIA	<i>Freedom of Information Act</i>
LIODS	Laboratório de Inovação, Inteligência e Objetivos de Desenvolvimento Sustentável
LOD	<i>Linked Open Data</i>
MARC	<i>Machine-Readable Cataloging</i>
NPS	<i>Net Promoter Score</i>
ODM	Objetivos de Desenvolvimento do Milênio
ODS	Objetivos de Desenvolvimento Sustentável
OAI-PMH	<i>Open Archives Iniciative Protocol for Metadata Harversting</i>
OGD	<i>Open Government Data</i>
OGP	<i>Open Government Partnership</i>
OKF	<i>Open Knowledge Foundation</i>
ONU	Organização das Nações Unidas
PNDU	Programa das Nações Unidas para o Desenvolvimento
PPGInfo	Programa de Pós-Graduação em Gestão da Informação
QI	Qualidade da Informação
RDF	<i>Resource Description Framework</i>
STF	Supremo Tribunal Federal
TIC	Tecnologia da Informação e Comunicação
TCU	Tribunal de Contas da União
W3C	<i>World Wide Web Consortium</i>

SUMÁRIO

1	INTRODUÇÃO	10
1.1	PROBLEMA DE PESQUISA	13
1.2	OBJETIVOS	14
1.2.1	Objetivo geral.....	14
1.2.2	Objetivos específicos.....	15
1.3	JUSTIFICATIVA	15
1.4	ESTRUTURA DO TRABALHO.....	18
2	FUNDAMENTAÇÃO TEÓRICA	19
2.1	O problema da Qualidade da Informação no âmbito dos Dados Abertos Governamentais	20
2.1.1	Qualidade da informação.....	23
2.2	Dados Abertos	29
2.2.1	Dados abertos governamentais	33
2.3	Princípios FAIR e dados abertos	42
2.4	Frictionless Data e Data Package	50
3	MODELO BASEADO EM <i>FRICITIONLESS DATA</i> PARA DAGS	55
3.1	VISÃO GERAL DO MODELO	55
3.1.1	Trabalhos relacionados	55
3.1.2	Apresentação do modelo e metodologia	57
3.1.2.1	<i>Etapa empacotar</i>	62
3.1.2.2	<i>Etapa tratar.....</i>	63
3.1.2.3	<i>Etapa publicar.....</i>	64
3.1.2.4	<i>Etapa feedback.....</i>	65
3.2	SUJEITOS DA PESQUISA, CONTEXTO E COLETA DE DADOS	68
3.3	FERRAMENTAS	76
3.3.1	Aplicação do modelo e resultados	80
3.3.1.1	<i>Verificação da etapa empacotar.....</i>	81
3.3.1.2	<i>Verificação da etapa tratar</i>	87
3.3.1.3	<i>Verificação da etapa publicar</i>	89
4	CONCLUSÕES E TRABALHOS FUTUROS.....	93
	REFERÊNCIAS.....	98

1 INTRODUÇÃO

As organizações públicas (em especial os governos) estão sujeitas a novas exigências da sociedade, tais como: aumento da transparência na utilização dos recursos públicos, maior participação na gestão desses recursos, maior controle sobre a qualidade dos serviços prestados, maior responsabilização dos gestores públicos, entre outras. A partir dessas recentes exigências políticas, econômicas e sociais, novas formas de gestão e de aproximação dos governos com a sociedade se fazem necessárias. Na busca dessas novas formas de gestão e, especialmente visando maior aproximação com as novas exigências da sociedade, uma das possibilidades é a utilização de ferramentas tecnológicas. Com o advento das tecnologias atuais, especialmente vinculadas à Tecnologia da Informação e Comunicação (TIC), com destaque para a Internet, as organizações públicas estão viabilizando a oferta de uma variada gama de novos produtos e serviços à sociedade (ALBANO; ARAUJO; REINHARD, 2017).

Na última década, um movimento conhecido como Dados Abertos Governamentais (DAGs) emergiu e ganhou importância, o qual estabelece um conjunto de requisitos a serem seguidos pelas instituições públicas para disponibilização de dados públicos, realizando principalmente essa tarefa por meio do uso de TIC (CORREA; SOUZA; SILVA, 2019). Os DAGs são uma metodologia para a publicação de dados do governo em formatos reutilizáveis, visando o aumento da transparência e uma maior participação política por parte do cidadão, além de gerar diversas aplicações desenvolvidas colaborativamente pela sociedade. Alguns efeitos dos DAGs sobre as políticas públicas são (GOVERNO DIGITAL, 2019):

- a) inclusão: fornecer dados em formatos padronizados abertos e acessíveis permite que todo cidadão utilize qualquer ferramenta de software para adaptá-los às suas necessidades;
- b) transparência: informações do setor público abertas e acessíveis melhoram a transparência, pois as partes interessadas podem usá-las da maneira mais adequada ao seu propósito;
- c) prestação de contas: os conjuntos apropriados de dados abertos, devidamente associados, podem oferecer vários pontos de vista sobre o

desempenho dos governos no cumprimento de suas metas em políticas públicas.

O acesso à informação, nos moldes que se conhece hoje, está diretamente vinculado à adoção da Resolução 59 da Organização das Nações Unidas (ONU) de 1946 (ASSEMBLY, 1946), que apresenta esse acesso como um direito humano e pedra fundamental das liberdades nas quais a ONU é fundamentada. Além disso, o acesso à informação implica no direito de coletar, transmitir e publicar informações (CARDOSO, 2019).

Em setembro de 2011, o Brasil tornou-se um membro da *Open Government Partnership* (OGP), uma iniciativa multinacional para promover mundialmente ampla adoção de *Open Government Data* (OGD). O compromisso compreende marcos políticos e técnicos e o lançamento do Portal Brasileiro de Dados Abertos. O Portal Brasileiro de Dados Abertos é a ferramenta disponibilizada pelo governo para que todos possam encontrar e utilizar os dados e as informações públicas, que disponibiliza dados relativos às mais variadas temáticas da administração pública, tais como dados da saúde suplementar, do sistema de transporte, segurança pública, indicadores de educação, gastos governamentais, processo eleitoral etc. (PORTAL BRASILEIRO DE DADOS ABERTOS, [2019]). O objetivo inicial era reunir grandes quantidades de dados agregados do governo para publicação digital e criar um catálogo central de informações da atividade pública, com intenção de melhorar a governança e monitorar a atividade do governo (BREITMAN et al., 2012).

Em 2017, a ONU disponibilizou o Guia de Planejamento de Dados Abertos Governamentais para o Desenvolvimento Sustentável, inspirado na experiência do Brasil, estabelecida pelo Decreto nº 8.777/2016, que trata da Política de Dados Abertos, no qual está definido que cada órgão e entidade da administração pública federal deve fazer o seu próprio Plano de Dados Abertos. O documento contém recomendações e um passo a passo para que organizações planejem ações relacionadas à abertura de dados, com vistas a promover os Objetivos de Desenvolvimento Sustentável (ODS) da Agenda Global 2030 (PORTAL BRASILEIRO DE DADOS ABERTOS, 2017).

A Agenda Global 2030 é um compromisso assumido por líderes de 193 países, inclusive o Brasil, e coordenada pelas Nações Unidas, por meio do Programa das Nações Unidas para o Desenvolvimento (PNUD), nos termos da Resolução

A/RES/72/279.OP32, de 2018, da Assembleia Geral da ONU. São 17 ODS e 169 metas a serem atingidas de 2016 a 2030, relacionadas com a efetivação dos direitos humanos e promoção do desenvolvimento, que incorporam e dão continuidade aos 8 Objetivos de Desenvolvimento do Milênio (ODM), a partir de subsídios construídos na Conferência Rio+20, conforme evidenciado na Portaria nº 133, de 28 de setembro de 2018 (BRASIL, 2018).

Os ODS são como uma lista de tarefas a ser cumprida pelos governos, pela sociedade civil, pelo setor privado e por todos cidadãos na jornada coletiva para um 2030 sustentável. Nos próximos anos de implementação da Agenda Global 2030, os ODS e suas metas irão estimular e apoiar ações em áreas de importância crucial para a humanidade: pessoas, planeta, prosperidade, paz e parceria (BRASIL, 2018). Logo, o sucesso da implementação dos ODS depende de parcerias entre sociedade, organizações públicas e privadas. Para ajudar a cumprir as metas, os estados membros da ONU podem usar DAGs, especificamente o objetivo 16, que tem como descrição: “promover sociedades pacíficas e inclusivas para o desenvolvimento sustentável, proporcionar o acesso à justiça para todos e construir instituições eficazes, responsáveis e inclusivas em todos os níveis” (INSTITUTO DE PESQUISA ECONÔMICA APLICADA, 2019). Um exemplo desse uso são os dados públicos como subsídios para aumentar transparência, prestação de contas públicas e participação do cidadão, expondo e prevenindo a corrupção e a má administração dos recursos públicos (WORLD WIDE WEB CONSORTIUM, 2009).

Os órgãos públicos em geral criam, armazenam e divulgam uma grande variedade de informações, desde dados demográficos e econômicos, até renda familiar. Esses dados podem assumir formatos diferentes, divulgados como dados primários ou processados. Podem também estar relacionados a serviços públicos ou processos internos (REN; GLISSLAMM, 2012). Por exemplo, a informação judicial disponibilizada pelos sites dos Tribunais de Justiça, acessada diariamente pelo cidadão comum e por todos os operadores do direito, devido ao seu grande volume e importância, necessita ter boa qualidade (ALBUQUERQUE; BASTOS; LINO, 2009).

A disseminação dos DAGs é considerada a força motriz do crescimento econômico e social, além de ser um fator essencial para publicidade das ações governamentais, porém a alta disponibilidade de dados não garante a qualidade das informações presentes nessas bases de dados públicos (QUARATI; MARTINO, 2019). Além disso, os consumidores de dados públicos, que podem ser cidadãos ou

instituições privadas, perdem muito tempo e recursos apenas para entender e trabalhar com as informações disponíveis, um problema conhecido como ‘atrito’ no trabalho com os dados (OPEN KNOWLEDGE FOUNDATION, 2019). Por outro lado, os produtores de dados públicos enfrentam dificuldades como falta de uniformidade nas políticas de publicação de dados institucionais, o que se constitui uma barreira para a abertura dos dados públicos (ALBANO, 2014).

De acordo com Araujo *et al.* (2012), o setor público em geral possui uma grande quantidade de informações publicadas em diferentes formatos ou em formatos que impossibilitam acesso pelas partes interessadas, ou seja, a inexistência de uma integração entre essas bases compromete a abertura dos dados, assim como interfere na qualidade das informações. Diversas iniciativas surgiram para resolver o problema da qualidade da informação referente às barreiras técnicas que envolvem a publicação de DAG. Recentemente, uma iniciativa chamada *Frictionless Data*, criada pela *Open Knowledge Foundation*, se apresenta como uma alternativa para remover o ‘atrito’ no trabalho com os dados, com foco na simplicidade e uma estrutura básica que permita qualquer pessoa publicar e usar conjuntos de dados com mais facilidade e sem passar por um intermediário. O *Frictionless Data* é um conjunto de ferramentas e estratégias focado em facilitar a interoperabilidade de dados por meio da definição de padrões e critérios técnicos e com o objetivo de otimizar o armazenamento e o uso de dados. Nesse sentido, a iniciativa oferece uma série de bibliotecas, aplicações e instruções em diferentes linguagens para aplicar os padrões *Frictionless*, facilitando a gestão de metadados, a limpeza, a organização, a extração, o compartilhamento, entre outros recursos (OLIVEIRA, 2020).

Para fins do presente trabalho, frente à complexidade do assunto e relevância social, optou-se por limitar o escopo da pesquisa às Bases de Boas Práticas do Judiciário relacionadas aos Objetivos de Desenvolvimento Sustentável (BBPODS), com o objetivo de propor um modelo baseado em *Frictionless Data* para compartilhamento e publicação de DAGs.

1.1 PROBLEMA DE PESQUISA

Nos últimos anos houve um crescimento na quantidade de dados gerados pelas organizações públicas a partir de suas diferentes fontes. Porém, esses dados nem sempre estão disponíveis ao público em formato adequado que permita o livre uso da

informação, já que essas fontes muitas vezes se encontram em formato proprietário. Segundo Janssen, Charalabidis e Zuiderwijk (2012), os gestores públicos simplesmente preferem disponibilizar os dados sem se preocupar com outras atividades que envolvam a publicação de dados. No entanto, os dados geralmente não podem ser usados em estado bruto, podendo ser necessário antes uma avaliação da qualidade, modificação e processamento desses dados. Já Dawes e Helbig (2010) identificaram que usuários de dados públicos almejam obter dados de melhor qualidade e gastam recursos consideráveis para obter, melhorar e padronizar os dados, sendo grande parte desse custo associado à necessidade de complementar, verificar, corrigir e integrar os dados coletados.

Quanto aos benefícios que os dados abertos proporcionam à sociedade, de acordo com Araujo *et al.* (2012), destacam-se: aumento da participação popular; *empowerment* dos cidadãos; melhores produtos e serviços privados, desenvolvimento de soluções inovadoras a partir dos dados divulgados; melhoria na eficiência dos serviços governamentais; medição de impacto das políticas públicas e geração de conhecimento a partir da combinação de fontes de dados e padrões.

Diante do exposto, por meio desta pesquisa, pretende-se responder à seguinte pergunta: **Como um modelo baseado na especificação *Frictionless Data* pode contribuir para a abertura dos DAGs?**

1.2 OBJETIVOS

No intuito de auxiliar a responder à questão de pesquisa, foram delineados os objetivos geral e específicos a seguir.

1.2.1 Objetivo geral

Desenvolver um modelo baseado no arcabouço *Frictionless Data* para auxiliar a publicação e o compartilhamento de DAGs, tendo como subsídios para a pesquisa as BBPODS no contexto do Judiciário Brasileiro.

1.2.2 Objetivos específicos

Para atingir o objetivo geral da pesquisa, foram definidos os objetivos específicos relacionados às etapas do processo de observação:

- a) identificar e avaliar abordagens recentemente empregadas para publicação de DAGs;
- b) identificar dificuldades e barreiras apresentadas no processo de publicação de DAGs;
- c) mapear os dados presentes nas BBPODS;
- d) demonstrar a aplicação prática do modelo baseado em *Frictionless Data* em uma BBPODS.

1.3 JUSTIFICATIVA

De acordo com Janssen, Charalabidis e Zuiderwijk (2012), a disponibilidade de dados abertos públicos aumentou significativamente, com pressão sobre todos os tipos de organizações públicas para liberar seus dados brutos. Algumas das motivações principais são que os dados abertos públicos geralmente são indispensáveis para o desenvolvimento de políticas públicas e a prestação de serviços, além de serem valiosos para terceiros, como as informações de tráfego.

No Brasil, com a criação da Lei de Acesso à Informação (LAI) para promover o acesso à informação pública, os órgãos e entidades públicas devem utilizar todos os meios e instrumentos legítimos de que dispuserem, sendo obrigatória a divulgação dos dados públicos em sítios oficiais na internet. Esses sítios devem atender alguns requisitos, tais como: conter ferramenta de pesquisa de conteúdo que permita o acesso à informação de forma objetiva, transparente, clara e em linguagem de fácil compreensão; possibilitar a gravação de relatórios em diversos formatos eletrônicos, inclusive abertos e não proprietários, tais como planilha de texto, de modo a facilitar a análise das informações; possibilitar o acesso automatizado por sistemas externos em formatos abertos, estruturados e legíveis por máquina; divulgar em detalhes os formatos utilizados para estrutura da informação e garantir a autenticidade e integridade das informações disponíveis para acesso (BRASIL, 2011).

Em 1º de outubro de 2020 o Governo Federal lançou o Modelo de Referência para Abertura de Dados. O documento tem como finalidade integrar, capacitar e sensibilizar a sociedade e as três esferas de governo no contexto da abertura de dados públicos, buscando, portanto, orientar gestores, agentes públicos e sociedade quanto à importância, ao uso, à publicação, à sustentação e ao monitoramento de dados abertos nos seus respectivos campos de atuação (BRASIL, 2020). O Modelo de Referência para Abertura de Dados contempla diversos aspectos relativos à publicação de dados abertos, abordando os parâmetros recomendados para catalogação, bem como critérios relevantes para propiciar maior interatividade aos consumidores de dados. O grande volume de dados governamentais disponíveis atinge o status de realmente abertos quando ocorre uma comunicação social bidirecional entre a administração pública e a população. Um dos postos-chave para tornar essa comunicação efetiva é a disponibilização de dados por meio de ferramentas que possibilitem compreender seu significado e sua reutilização (OLIVEIRA, 2020).

A abertura de dados governamentais deve ser estimulada pelo princípio de que, ao menos que haja argumentos para a não divulgação, os dados devem ser publicados. Porém, muitas instituições públicas estão divulgando dados sem uma política sólida de abertura. Essa ausência de política acarreta alguns problemas, como a falta de padronização de formatos e de um portal centralizado para publicação, com consequente pulverização das informações. Além disso, por conta dos dados terem um valor intrínseco baixo, apenas sua combinação é capaz de produzir valor (CARDOSO, 2019). Logo, apesar de todo conhecimento disponível nos portais que disponibilizam DAGs, tais informações estão em formatos digitais diferentes, os quais dificultam o livre uso, reuso e redistribuição por qualquer cidadão (ALDEBERT; AUGUSTI, 2020).

Outro problema é a falta de mecanismos de *feedback* que mostre o que é feito com os dados abertos. Comumente os governos que publicam conjuntos de dados enfrentam críticas como má usabilidade, falta de *feedback* de dados e mecanismos de melhoria e metadados inadequados (DAWES; HELBIG, 2010). Por outro lado, ressalta-se que a definição de padrões pode proporcionar diversos benefícios, pois de acordo com Kassen (2013), os DAGs podem criar um ambiente favorável para o engajamento cívico proativo, fornecendo uma oportunidade real para desenvolvedores independentes criar aplicativos usando conjuntos de DAGs,

juntamente com a participação dos cidadãos, criando um novo ambiente de cooperação entre governo local e os cidadãos, tendo como consequência uma verdadeira transformação na forma de comunicação entre governo e sociedade.

Conforme estudo realizado por Dalenogare e Araújo (2019), no que tange às subcategorias mais voltadas diretamente a dados públicos, figura a posição recorrente dos autores sobre a necessidade da definição de um *framework* ou modelo padrão de disponibilização de dados dos governos, a fim de que seja proporcionada maior unicidade e respeito a determinados critérios aceitos internacionalmente. Segundo a literatura verificada, a ausência de um *framework* ou modelo prejudica o desenvolvimento dos DAGs, bem como seu uso efetivo e a sua conversão em benefícios para a sociedade de uma forma geral.

Diante o exposto, quanto às possíveis contribuições, o presente trabalho pode colaborar, por exemplo, com iniciativas governamentais como a do Conselho Nacional da Justiça (CNJ) que, por meio da Portaria nº 133, de 28 de setembro de 2018, criou o Comitê Interinstitucional, cujo objetivo é “proceder estudos e apresentar proposta de integração das metas do Poder Judiciário com as metas e indicadores dos ODS, Agenda Global 2030” (BRASIL, 2018). Nesse sentido, a aplicabilidade e a generalização desta pesquisa têm grande potencial de contribuição científica e social.

Ainda segundo a Portaria nº 133, de 28 de setembro de 2018, a integração dos ODS com as Boas Práticas do Judiciário pode trazer diversos benefícios como: facilitar o acompanhamento do Tribunal de Contas da União (TCU) no desempenho de programas governamentais; aperfeiçoar os mecanismos de busca nos Portais de Transparência dos Tribunais, de forma associada aos ODS; auxiliar na medição da eficiência do Poder Judiciário em atingir os ODS, entre outros. Outra iniciativa que utiliza os ODS para acompanhamento de resultados é o Portal PPA Cidadão, que foi concebido para permitir ao cidadão acompanhar a implantação de metas e objetivos do Plano Plurianual do Governo Federal e os ODS, uma parceria do Governo Federal com o OGP e coordenado pela Corregedoria Geral da União (BRASIL, 2018).

Sendo assim, esta pesquisa se justifica ao propor um modelo comum para publicação de DAGs alinhado com o novo Modelo de Referência para Publicação de Dados Abertos do Governo Federal Brasileiro. Propõe facilitar o compartilhamento de um conjunto de dados públicos, que envolve a sua identificação, extração, tratamento e exposição. A pesquisa também vai ao encontro das temáticas de estudo do Programa de Pós-Graduação em Gestão da Informação (PPGInfo), ao promover a

qualidade das informações presentes nas Bases de Boas Práticas dos ODS, de acordo com atributos, tais como: acessibilidade e consistência.

A presente proposta de pesquisa pode, ainda, ao contribuir com o reuso e compartilhamento dos dados, permitir a construção de modelos para reconhecimento de padrões de DAGs. No caso do uso das Bases de Boas Práticas do Judiciário, possibilita a classificação automática das boas práticas e as temáticas de desenvolvimento sustentável, sendo a sua aplicação possível em diversas áreas, contribuindo para resolução de problemas comuns aos governos e cidadãos.

1.4 ESTRUTURA DO TRABALHO

O presente trabalho é composto por cinco seções, organizado da seguinte maneira: na seção dois são apresentados os conceitos fundamentais para entendimento da pesquisa; na seção três aborda-se a metodologia a ser utilizada para atingir os objetivos do trabalho. Na seção quatro debate-se sobre a estratégia prática por meio de um protótipo aplicado a uma base de dados de boas práticas e, por fim, são apresentadas as conclusões obtidas a partir da fundamentação teórica, desenvolvimento da pesquisa e cenário prático. Também são expostas sugestões, recomendações para trabalhos futuros e referências bibliográficas.

2 FUNDAMENTAÇÃO TEÓRICA

O objetivo da seção é construir um arcabouço teórico para o entendimento e aprofundamento das questões expostas como problemas de pesquisa. Além disso, apresentar uma visão geral sobre os conceitos pertinentes para a discussão a respeito do modelo proposto no presente trabalho.

Nesta seção são apresentados os fundamentos conceituais desta pesquisa: principais definições sobre a problemática do trabalho com DAGs sob a ótica da qualidade da informação; com uma discussão sobre este conceito na literatura; um breve levantamento sobre compartilhamento de dados na *web* com o advento dados abertos e DAGs; uma introdução aos princípios de *Findable, Accessible, Interoperable* e *Reusable* (FAIR); apresentação dos conceitos fundamentais da presente pesquisa sobre *Frictionless Data* e *Data Package* e uma breve descrição sobre os ODS.

Segundo Marconi e Lakatos (2002) um problema, para ser considerado relevante, precisar ser analisado sob a ótica de seu valor. De acordo com os autores, um problema de pesquisa deve ser viável, relevante, novo, exequível e oportuno. O presente trabalho se fundamentou em uma documentação direta sobre as abordagens para publicação de dados abertos governamentais. Buscou-se, por intermédio dos mecanismos de busca presentes no portal de Periódicos da CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior), base de dados utilizada na presente pesquisa, o uso de palavras-chave como modelo, *open government data*, *open data* e *Frictionless Data*, com o objetivo de identificar estudos com propostas de modelos para publicação de dados abertos governamentais. Por meio do levantamento bibliográfico, identificou-se na literatura as principais características presentes nos modelos para abertura de dados abertos governamentais e as barreiras inerentes ao processo de publicação.

A técnica de pesquisa de laboratório oferece um resultado mais exato, apesar de ser considerado um procedimento de investigação mais difícil. Além disso, exige instrumental específico, preciso e ambientes adequados (MARCONI; LAKATOS, 2003), o que se alinha aos objetivos delineados na presente pesquisa. Por essas razões, a técnica de pesquisa de laboratório foi selecionada para descrever e analisar a aplicação do modelo proposto em um ambiente controlado para fins específicos do presente trabalho. Foram utilizadas as ferramentas online disponíveis na iniciativa *Frictionless Data* para realizar os experimentos, usando a base de dados do CNJ

relacionadas aos Objetivos de Desenvolvimento Sustentável, e verificar todas as etapas propostas no presente modelo.

2.1 O PROBLEMA DA QUALIDADE DA INFORMAÇÃO NO ÂMBITO DOS DADOS ABERTOS GOVERNAMENTAIS

Dados precisos são um requisito fundamental para bons sistemas de informação. No entanto, a maioria dos sistemas de informação contém quantidades significativas de dados imprecisos e a maioria das empresas não possui um entendimento básico dos conceitos de qualidade da informação para reconhecer ou alterar a situação. “Dados precisos são uma dimensão importante da qualidade da informação – a dimensão mais importante” (OLSON, 2003).

Mas, antes de apresentar o conceito de qualidade da informação, é importante diferenciar, para melhor entendimento do presente trabalho, o conceito de dado e informação, além da abordagem que se assume nesta pesquisa. Primeiramente, não se assumiu, assim como em Barreto (2008), o conceito de informação apresentado por Shannon em 1949, na sua Teoria da Informação, a qual implica uma quantificação de dados, no processo via canal físico entre emissor e receptor:

[...] Deve-se esclarecer que não foi incluída a ‘informação’ no sentido de Shannon, na sua ‘Teoria da Informação’, parte da Matemática Aplicada, que implica a quantificação de dados, no processo de comunicação via um canal físico entre emissor e receptor [...] (BARRETO, 2008, p. 11).

De acordo com Roszac (1988 *apud* NEHMY, 1998) na obra de Shannon o termo informação recebeu um sentido técnico que a divorciaria do senso comum, separando-se do conteúdo semântico das frases e se tornando uma medida puramente quantitativa de trocas comunicativas. Dessa perspectiva, a qualidade da informação estaria relacionada à medida de capacidade de transmissão das sequências produzidas por uma fonte, podendo a informação ser desprovida de significado.

Segundo Davenport (1998), não é fácil distinguir, na prática, dados, informação e conhecimento, sendo no máximo possível elaborar um processo que inclua os três. O autor define dados como “observações sobre o estado do mundo e como exemplo

a observação de dados brutos, ou entidades quantificáveis, que pode ser feita por pessoas ou por uma tecnologia apropriada" (DAVENPORT, 1998). Sendo assim, os dados são, em síntese, os elementos por meio dos quais as informações são geradas. O entendimento desses conceitos é fundamental para o processo de desenvolvimento de sistemas de informação, uma vez que os sistemas registram, em suas bases de dados, os dados que proverão as informações necessárias ao atendimento das necessidades organizacionais (MACHADO, 2013).

Em relação ao conceito de informação, Davenport (1998) cita o definido por Peter Drucker: "informação são dados dotados de relevância e propósito", que no caso são atribuídos por seres humanos. Diversas são as manifestações conceituais sobre o significado da informação que surgem na trajetória da Ciência da Informação (CI), conforme levantamento realizado por Silva (2017), o qual apresentou uma representação conceitual de informação na trajetória da CI de 1945 até 2013. O Quadro 1 apresenta alguns conceitos relevantes para a presente pesquisa, que foram elencados por Silva (2017):

Quadro 1 - Algumas manifestações conceituais de informação na trajetória da Ciência da Informação

Autor/Instituição	Conceito	Ano
Harold Borko	A informação inserida no conceito de CI, sendo a informação compreendida, por um lado, no sentido de suas propriedades, comportamentos, fluxos e, por outro lado, por processos de conhecimentos relacionados à origem, coleção, organização, armazenamento, recuperação, interpretação, transmissão, transformação e utilização da informação.	1968
Robert Hayes	É uma propriedade dos dados resultante de ou produzida por um processo realizado sobre os dados. O processo pode ser simplesmente a transmissão de dados (em cujo caso são aplicáveis a definição e medida utilizada na Teoria da Comunicação); pode ser a seleção dos dados; pode ser a organização dos dados; pode ser a análise de dados.	1986
Michel Buckland	Informação como processo ("informação" é o ato de informar, comunicação do conhecimento ou novidade de algum fato ou ocorrência), informação como conhecimento (o conhecimento comunicado referente a algum fato particular, assunto ou evento; aquilo que é transmitido, inteligência ou notícias) e informação como coisa (atribuído para objetos, assim como dados para documentos, que são considerados como informação, porque são relacionados como informativos, tendo a qualidade de conhecimento comunicado ou comunicação, informação, algo informativo).	1991
Yves-François Le Coadic	É um conhecimento inscrito (gravado) sob a forma escrita (impressa ou digital), oral ou audiovisual.	1994

Fonte: Adaptado de Silva (2017, p. 224).

Dentre os conceitos sobre informação apresentados, optou-se por adotar a visão de Borko (1968), que comprehende a informação sob a perspectiva dos processos de conhecimento relacionados à origem, coleção, organização, armazenamento, recuperação, interpretação, transmissão, transformação e utilização da informação, além de abranger o sentido de suas propriedades, comportamentos e fluxos. A definição de informação apresentada por Borko (1968) suporta a abordagem do presente trabalho, uma vez que se teve como objetivo identificar problemas de qualidade da informação que se constituem barreiras quanto ao seu acesso e utilidade, conforme será abordado nas próximas seções.

No contexto das instituições públicas, foco do presente trabalho, tem-se que os órgãos públicos criam, armazenam e divulgam uma grande variedade de informações, desde dados demográficos e econômicos até renda familiar. Esses dados podem assumir formatos diferentes, divulgados como dados primários ou processados. Podem também estar relacionados a serviços públicos ou processos internos (REN; GLISSLAMM, 2012). Por exemplo, a informação judicial disponibilizada pelos sites dos Tribunais de Justiça, acessada diariamente pelo cidadão comum e por todos os operadores do direito, devido ao seu grande volume e importância, necessita ter boa qualidade (ALBUQUERQUE; BASTOS; LINO, 2009). No Brasil, apesar dos esforços para promover a transparência do governo brasileiro através da divulgação de seus dados públicos, ainda são comuns problemas como informações em formatos proprietários disponíveis nas Plataformas e Portais de Transparência, o que dificulta o acesso aos dados pelo cidadão.

A divulgação de dados abertos governamentais pode mostrar que a qualidade dos dados, sobre os quais são tomadas decisões importantes, é ruim. O sucesso na abertura de dados governamentais requer mais do que a simples provisão de acesso aos dados. Também são necessários o aprimoramento da qualidade das informações governamentais, a criação e a institucionalização de uma cultura de governo aberto e o fornecimento de ferramentas e instrumentos com os quais os dados serão utilizados. Essa perspectiva mais ampla precisa ser adotada pelos governos que agora estão apenas abrindo um portal para tornar os dados acessíveis. É necessária uma infraestrutura que ajude os usuários a entender os dados, e medidas institucionais são necessárias para garantir o envolvimento do público. Sob essas condições, os dados abertos podem potencialmente ir além do nível atual de envolvimento dos cidadãos e

resultar em um diálogo contínuo entre os governos e seus constituintes, baseando-se na inteligência coletiva do público (ZUIDERWIJK *et al.*, 2012).

Assim, diante do exposto, apresenta-se esta subseção organizada em cinco partes. Na primeira parte introduz-se uma discussão sobre o conceito de qualidade da informação presente na literatura, em que se enfatiza os *frameworks* para avaliação da qualidade da informação, por meio dos quais se destacam atributos e dimensões hipoteticamente importantes para mensuração no contexto dos DAGs. Na segunda parte é apresentado um histórico sobre as abordagens de abertura de dados, além de conceitos relevantes como melhores práticas para dados abertos. Na terceira parte apresenta-se um levantamento na literatura sobre DAGs, as principais barreiras e desafios para publicação desse tipo de dado, como também os benefícios para a sociedade quando as instituições os disponibilizam. Na quarta parte tem-se uma introdução sobre os princípios FAIR, relacionados com uma descrição detalhada sobre os princípios e uma discussão sobre as semelhanças e as dimensões da qualidade da informação. Na quinta e última parte serão abordados os conceitos, características e funcionamento relacionados à abordagem *Frictionless Data*, um levantamento sobre o estado da arte por meio da revisão da literatura, além de alguns exemplos de uso da iniciativa.

2.1.1 Qualidade da informação

A tarefa de definir objetivamente o que é qualidade da informação não é algo trivial. Na área da CI, por exemplo, observa-se que a definição de qualidade da informação ainda não é um consenso. No artigo intitulado ‘A desconstrução do conceito de qualidade da informação’, Schwuchow (1990 *apud* NEHMY; PAIM, 1998) apresenta uma análise sobre o conceito de qualidade da informação, de forma a “focalizar a discussão nas formas de abordagem da qualidade da informação na literatura, no esforço de se desvelarem limitações e desafios para a construção teórica do conceito”. De acordo com o estudo, não existe uma definição geralmente aceita sobre qualidade da informação, sendo um conceito vago e subjetivo.

Para Paim, Nehmy e Guimarães (1996), o tema qualidade da informação ainda esbarra na inerente dificuldade de conceituar o termo qualidade, não havendo um consenso na literatura sobre definições teóricas e operacionais da qualidade da informação. Os autores analisam as dificuldades existentes e a tendência na literatura

da qualidade da informação sob três vertentes: qualidade transcendente (ou filosófica, ou metafísica) da informação; uma vertente que se baseia em aspectos intrínsecos, a qual alguns autores entendem que é sinônimo da transcendente; e outra nos atributos contingencias.

Ainda de acordo com Paim, Nehmy e Guimarães (1996), pela perspectiva transcendente, a informação e seu valor é tido como absoluto e universalmente aceito, aproximando-se da ideia de excelência e coexistindo com uma essência independentemente do tempo-espacó histórico. Com relação ao ponto de vista contingencial ou prático:

[...] a tendência mais marcante na literatura é o enfoque no usuário. A proposição central é a de que o valor ou a qualidade da informação depende do usuário e do contexto em que é considerada. O usuário, quer individual ou coletivo, faz o julgamento da informação. O foco no usuário, em estudos de qualidade, segue o movimento mais geral da ciência da informação de apostar no paradigma do usuário (DERVIN, 1977; VAKKARI, 1994; WAGNER, 1990), em contraposição ao modelo de abordagem anteriormente dominante, que pensava a informação a partir da teoria matemática de Shannon. Nessa vertente, a qualidade da informação é relativa à medida quantitativa de eficácia técnica da transmissão de uma mensagem entre um emissor e um receptor (ROSZAC, 1989 *apud* PAIM; NEHMY; GUIMARÃES, 1996).

O estudo de Paim, Nehmy e Guimarães (1996) apresenta a classificação ‘transcendente, intrínseca e contingencial’ para qualidade da informação. Esse estudo tem como foco destacar os atributos de responsabilidade do provedor da informação, evitando o excessivo aspecto subjetivo das definições usuais de qualidade da informação.

O Quadro 2 mostra a classificação da qualidade da informação conforme os três aspectos mencionados.

Quadro 2 - Dimensões e atributos da qualidade da informação

Dimensões e Atributos			
	Intrínseca	Contingencial	
		Usuário	Formato do Produto
Valor filosófico ou metafísico	<ul style="list-style-type: none"> • Validez • Confiabilidade • Precisão • Completeza • Novidade • Atualidade • Significado através do tempo • Abrangência 	<ul style="list-style-type: none"> • Valor percebido • Eficácia • Relevância • Redundância 	<ul style="list-style-type: none"> • Eletrônico • Impresso • Oral • Microforma • Integral • Sintético • Formal • Informal

Fonte: Paim, Nehmy e Guimarães (1996, p. 115).

No Quadro 2, Paim, Nehmy e Guimarães (1996) apresentam o modelo multidimensional da qualidade da informação, o qual revela a complexidade típica ao processo de avaliação da informação. Os atributos intrínsecos se referem ao valor inerente à informação. Dimensões como confiabilidade, significam credibilidade no conteúdo e na fonte da informação, enquanto o formato, pertencente à categoria contingencial, representa a forma de apresentação do produto e seus atributos devem se adequar à preferência ou interesse do usuário.

Trabalhos como o de Trindade e Oliveira (2007) procuraram definir atributos para avaliação da qualidade da informação em sistemas que disponibilizam grandes quantidades de informação que podem ser compartilhadas e transformadas em conhecimento. No entendimento das autoras supracitadas, alguns atributos que possam ajudar a avaliar a qualidade da informação são: acessibilidade, clareza, completeza, concisão, consistência, credibilidade, disponibilidade, empacotamento, facilidade de uso, histórico, interface, interpretabilidade, objetividade, precisão, prontidão, quantidade, relevância, reputação, segurança, temporalidade e valor (TRINDADE; OLIVEIRA, 2007).

Ainda segundo Trindade e Oliveira (2007), a definição da qualidade da informação está relacionada aos requisitos de informação dos usuários, ou como a informação supre a necessidade dos consumidores de informação. Para as autoras, é de suma importância que os produtores de informação saibam quem são todos os consumidores da informação, quais os requisitos de qualidade, como o dado é utilizado e qual o custo do dado sem qualidade. Os consumidores de informação precisam acessar uma informação com qualidade de maneira a permitir o apoio às suas atividades.

Ainda de acordo com o levantamento feito por Trindade e Oliveira (2007), foram elencados atributos para avaliar a qualidade da informação e suas definições segundo diversos autores, conforme descritos a seguir:

- a) acessibilidade: a facilidade e a eficiência que o usuário pode navegar na aplicação para acessar a informação desejada (KAHN; STRONG; WANG, 2002; KIM; KISHORE; SANDERS, 2005);
- b) clareza: a informação é facilmente entendida (KAHN; STRONG; WANG, 2002);

- c) completeza: o sistema provê toda a informação necessária para a execução da atividade (KAHN; STRONG; WANG, 2002; SHANKARANARAYAN; ZIAD; WAND, 2003; WIXON; TODD, 2005; KIM, KISHORE; SANDERS, 2005);
- d) concisão: a informação é apresentada de forma compacta (KAHN; STRONG; WANG, 2002);
- e) consistência: a informação é apresentada no mesmo formato (KAHN; STRONG; WANG, 2002; WIXON; TODD, 2005);
- f) credibilidade: a informação é considerada verdadeira e confiável (KAHN; STRONG; WANG, 2002);
- g) disponibilidade: refere-se à flexibilidade e compreensão das características da aplicação disponibilizada aos usuários para especificar e controlar a relação temporal entre os vários componentes hipermídia, para que seja disponibilizada uma informação de hipermídia integrada (KIM; KISHORE; SANDERS, 2005);
- h) empacotamento: refere-se à variedade de informações que os vários tipos de mídias são empacotados dentro de uma interface web para apresentação ao usuário final (KIM; KISHORE; SANDERS, 2005);
- i) facilidade de uso: a informação é fácil de manipular e se aplica a diferentes atividades (KAHN; STRONG; WANG, 2002);
- j) histórico: refere-se à flexibilidade e compreensão das características da aplicação disponibilizadas para os usuários para especificar e manter um histórico das ações dos usuários e do estado da aplicação (KIM; KISHORE; SANDERS, 2005);
- k) interface: consistência no arranjo estrutural e no estilo do conteúdo da informação e *hiperlinks* dentro da aplicação, devendo permitir que o usuário conheça o conteúdo da informação disponibilizada nas páginas web (KIM; KISHORE; SANDERS, 2005);
- l) interpretabilidade: a informação está na linguagem e símbolos apropriados e as definições estão claras (KAHN; STRONG; WANG, 2002);
- m) objetividade: a informação não é tendenciosa e imparcial (KAHN; STRONG; WANG, 2002);
- n) precisão: a informação está correta, isenta de erros (KAHN; STRONG; WANG, 2002; WIXON; TODD, 2005; KIM; KISHORE; SANDERS, 2005);

- o) prontidão: a informação é fornecida sempre que for necessária (KAHN; STRONG; WANG, 2002);
- p) quantidade: o quanto o volume de informação é apropriado para atividade que está sendo executada (KAHN; STRONG; WANG, 2002);
- q) relevância: a informação deve estar relacionada com o interesse e as necessidades de informação do usuário (KAHN; STRONG; WANG, 2002; KIM; KISHORE; SANDERS, 2005);
- r) reputação: a informação é considerada verdadeira com relação à sua fonte ou ao seu conteúdo (KAHN; STRONG; WANG, 2002);
- s) segurança: o acesso à informação é mantido restrito apropriadamente para garantir a sua segurança (KAHN; STRONG; WANG, 2002);
- t) temporalidade: a informação está suficientemente atualizada (KAHN; STRONG; WANG, 2002; WIXON; TODD, 2005; KIM; KISHORE; SANDERS, 2005);
- u) valor: a informação provê benefícios e vantagens com o seu uso (KAHN; STRONG; WANG, 2002).

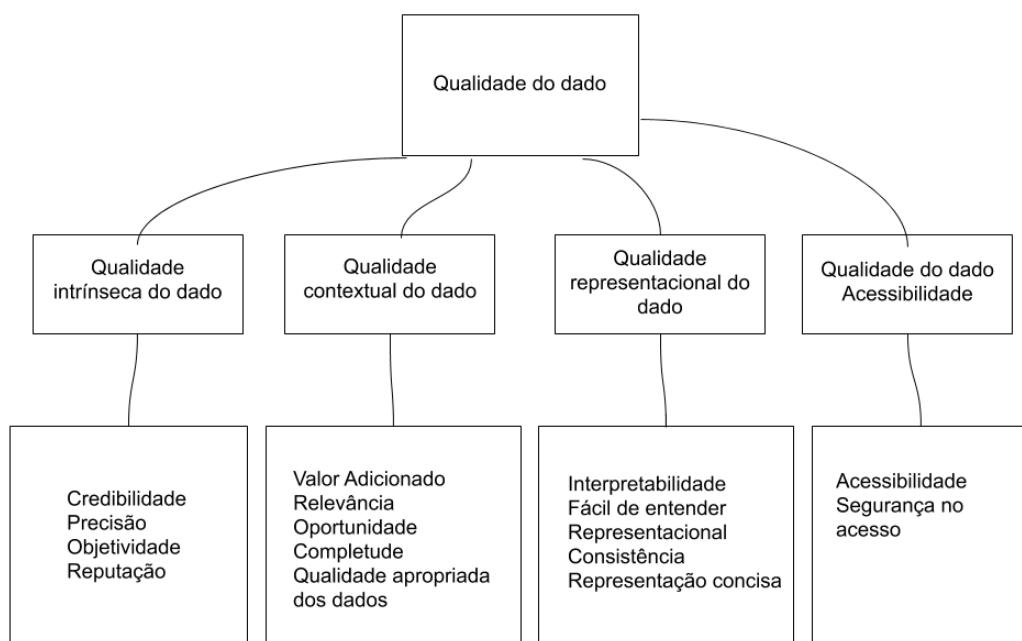
Trindade e Oliveira (2007) também argumentam que os atributos apresentados pela maioria dos autores considerados no seu estudo são completeza, precisão e temporalidade. Conclui-se que para obtenção da qualidade da informação em sistemas com grande complexidade de informações é indispensável obter a informação necessária para a execução da atividade, sendo a mesma correta e suficientemente atualizada.

Para Bobrowski, Marré e Yankelevich (1999), os problemas relacionados à qualidade da informação derivam de práticas equivocadas na engenharia de *software*. Foram levantados sete atributos para mensurar a qualidade da informação: completude, relevância, confiabilidade, consistência, temporalidade, precisão e concisão. Outros autores mostram a evolução sobre o tema qualidade da informação e sua aplicação no contexto organizacional. Segundo Calazans (2008), a falta de qualidade da informação em uma organização pode ter impactos sociais e no negócio. Fatores como informações com múltiplas origens, utilização de julgamentos subjetivos, sistemáticos erros na produção da informação, além da grande quantidade de informação em diferentes fontes influenciam a qualidade da informação (STRONG; LEE; WANG, 1997; CALAZANS, 2008).

Eppler e Wittig (2000) realizaram um levantamento sobre *frameworks* para avaliação da qualidade da informação. Foram identificadas vinte estruturas de qualidade da informação que definem e categorizam critérios para variados contextos. Ressalta-se que os autores adotam a definição de *framework* assumida por Porter (1991), no qual considera-se uma forma de ajudar a identificar variáveis relevantes e as perguntas que o usuário deve responder para desenvolver conclusões personalizadas para um determinado setor e empresa.

Ainda como uma forma de compreender o significado da qualidade de dados para os consumidores de informação, Wang e Strong (1996) desenvolveram um *framework* que captura aspectos da qualidade de dados que são importantes para quem usa a informação. Especificaram 118 atributos de qualidade de dados consolidados em 20 dimensões, agrupadas em 4 categorias. O objetivo era oferecer um *framework* para que profissionais que lidam com sistemas de informação pudessem melhorar e entender as necessidades dos usuários quanto à qualidade dos dados, conforme Figura 1:

Figura 1 - *Framework* conceitual de qualidade de dados



Fonte: Adaptado de Wang e Strong (1996, p 20).

A pesquisa de Wang e Strong (1996) adotou o ponto de vista de ‘adequação ao uso’ do consumidor da informação. O autor definiu a dimensão de qualidade dos dados como um conjunto de atributos de qualidade dos dados que representam um único aspecto ou construção.

Nas próximas subseções será retomada a discussão sobre qualidade da informação, dessa vez sob a perspectiva dos dados DAGs, mas antes serão apresentados conceitos sobre o que faz um dado ser considerado aberto.

2.2 DADOS ABERTOS

A *Open Knowledge Foundation* (OKF), no *Open Data Handbook*, define dados abertos como dados que podem ser livremente usados, reutilizados e redistribuídos por qualquer pessoa – sujeitos, tendo como única exigência a atribuição da fonte e compartimento pelas mesmas regras. Alguns pontos importantes relacionados à definição completa de ‘aberto’ são (OPEN DATA HANDBOOK, 2020):

- a) disponibilidade e acesso: os dados devem estar disponíveis como um todo e sob custo não maior que um custo razoável de reprodução, preferencialmente possíveis de serem baixados pela Internet. Os dados devem também estar disponíveis de uma forma conveniente e modificável;
- b) reutilização e redistribuição: os dados devem ser fornecidos sob termos que permitam a reutilização e a redistribuição, inclusive a combinação com outros conjuntos de dados;
- c) participação universal: todos devem ser capazes de usar, reutilizar e redistribuir – não deve haver discriminação contra áreas de atuação ou contra pessoas ou grupos. Por exemplo, restrições de uso ‘não comercial’ que impediriam o uso ‘comercial’, ou restrições de uso para certos fins (ex.: somente educativos) excluem determinados dados do conceito de ‘abertos’.

Duas qualidades importantes tornam os dados abertos: os dados devem estar ‘legalmente’ abertos, ou seja, significa que é irrestrito por termos restritivos de direitos autorais e pode ser compartilhado legalmente e usado por indivíduos, empresas, academia, organizações sem fins lucrativos e outros. Os dados também devem ser

abertos no sentido ‘técnico’ da palavra, que se refere à entrega dos dados em um formato padrão e bem definido, geralmente via Internet (AYRE; CRANER, 2017).

Outro conceito relevante relacionado aos dados abertos é a interoperabilidade, que significa a capacidade de diversos sistemas e organizações trabalharem juntos (interoperar). Ou seja, a capacidade de interoperar – ou combinar – diferentes conjuntos de dados. Conforme *World Wide Web Consortium* (W3C), um dado é definido como aberto quando permite que as informações sejam publicadas e disseminadas de maneira que terceiros consigam reutilizá-lo. Considera-se, ainda, que são três as leis dos dados abertos, conforme apresentado a seguir (WORLD WIDE WEB CONSORTIUM, 2009):

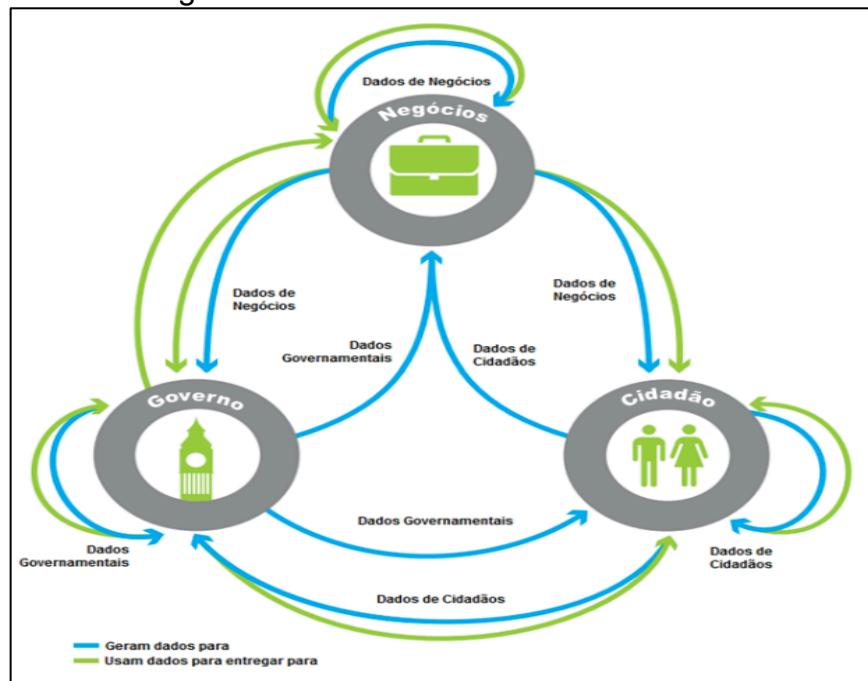
- a) os dados precisam ser encontrados e indexados na *web*;
- b) devem estar em formato aberto e comprehensíveis por máquina;
- c) sem restrições quanto à utilização.

Dados abertos se referem aos dados coletados e compartilhados com outras pessoas para usarem como desejam, sem restrições de direitos autorais ou uso. Exemplos tradicionais de dados abertos incluem dados coletados pelo governo (por exemplo, boletins meteorológicos, relatórios de incidentes criminais e códigos postais), bem como algumas fontes acadêmicas (periódicos de acesso aberto, pesquisas e dados de pesquisa e resultados de experimentos científicos). Outros exemplos de aplicação dos dados abertos são dados disponibilizados por governos estaduais, nacionais, agências de fomento e organizações não governamentais, que são usados para reduzir custos e melhorar os serviços que oferecem aos cidadãos (AYRE; CRANER, 2017).

De acordo com Germano (2013), existem três principais componentes em qualquer ecossistema de dados abertos bem-sucedido: governo, empresas e cidadãos. Cada integrante fornece dados para si mesmo e para os outros, ficando o governo e as empresas responsáveis por usar os dados para entregar serviços demandados por todos os componentes do ecossistema. Para tanto, são três as categorias de dados abertos fornecidos pelos componentes do ecossistema: os dados abertos governamentais (produzido por instituições públicas); os dados abertos de negócios (produzidos pelo setor privado) e os dados pessoais e não pessoais dos

cidadãos publicados em domínio aberto. A Figura 2 mostra o funcionamento do ecossistema e a interação entre seus atores quanto aos dados abertos.

Figura 2 - Ecossistema de dados abertos



Fonte: Germano (2013, p. 25).

Sob a perspectiva da qualidade da informação, a W3C criou o *Data on the Web Best Practices*, um guia que fornece as melhores práticas relacionadas à publicação e ao uso de dados na web, criado para ajudar e apoiar um ecossistema autossustentável para dados abertos. De acordo com o guia, composto por 35 melhores práticas, os dados precisam ser descobertos e compreensíveis por seres humanos e máquinas, sendo o objetivo dessas práticas recomendadas facilitar a interação entre o consumidor e o publicador de dados abertos. O Quadro 3 apresenta as melhores práticas, distribuídas em categorias e para cada melhor prática, uma breve descrição.

Quadro 3 - Melhores práticas para abertura de dados

Categoria	Melhor prática
Metadados	Fornecer metadados para usuários humanos e máquinas
	Fornecer metadados descritivos
	Fornecer metadados estruturados
Licenças	Fornecer informações de licença de dados
Proveniência de dados	Fornecer informações completas sobre as origens dos dados e quaisquer alterações feitas
Qualidade de dados	Disponibilizar informações sobre a qualidade de dados e adequações necessárias

Versão de dados	Atribuir versão para cada conjunto de dados Fornecer um histórico de versão
Identificadores de dados	Usar URLs persistentes de conjunto de dados Usar URLs persistentes como identificadores dentro de conjuntos de dados
	Atribuir URLs a versões de conjuntos de dados
Formato de dados	Usar formatos de dados padronizados Usar representações de dados locais neutras
	Fornecer dados em vários formatos
Vocabulário de dados	Reutilizar vocabulários, de preferência padronizados Escolher o nível de formalização correto
Acesso a dados	Permitir o acesso completo (em massa) Permitir o acesso parcial ao conjunto de dados Disponibilizar dados em vários formatos Permitir o acesso em tempo real Fornecer dados atualizados Explicar os motivos de quando os dados não estiverem mais disponíveis Disponibilizar dados através de uma API Usar padrões da web como base de APIs Fornecer documentação à medida que adicionar ou alterar uma API Evitar quebrar alterações na sua API
	Preservar o identificador e fornecer informações sobre o recurso arquivado
	Avaliar a cobertura de um conjunto de dados antes da sua preservação
	Coletar feedback dos consumidores
	Disponibilizar publicamente o feedback
	Enriquecer dados gerando novos dados Oferecer apresentações complementares
	Fornecer feedback ao publicador original
	Seguir termos de licença
	Citar publicação original

Fonte: Adaptado de Santarém Segundo(2018).

Ainda conforme o *Data on the Web Best Practices*, nem todos os dados e metadados devem ser compartilhados abertamente. Questões como segurança, sensibilidade comercial e, acima de tudo, privacidade dos indivíduos, precisam ser consideradas e deve ser de responsabilidade dos produtores de dados determinar quais políticas são adequadas e quais dados devem ser compartilhados e sob que circunstâncias.

Diante o exposto, faz-se necessário diferenciar dados abertos dos demais conceitos existentes na literatura. Nesse sentido, a próxima seção apresenta definições encontradas na revisão de literatura sobre DAGs, além de barreiras para o uso de dados públicos e seus benefícios proporcionados à sociedade.

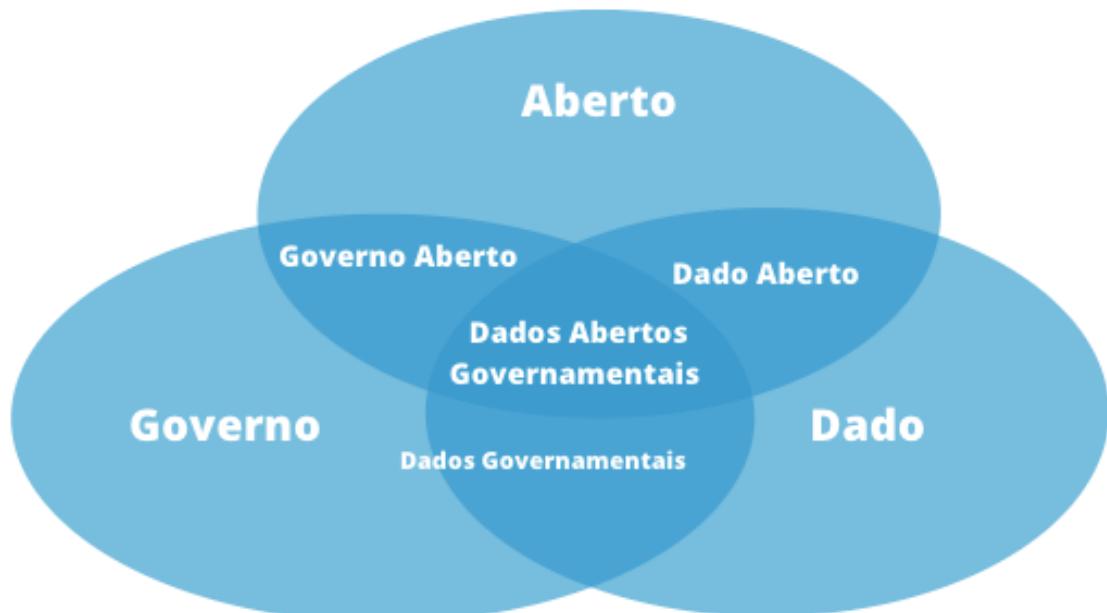
2.2.1 Dados abertos governamentais

Dados abertos governamentais também são caracterizados sob três perspectivas:

- a) **Dados do governo:** os governos coletam e geram uma grande quantidade de dados e conjuntos de dados. Esses conjuntos de dados contêm dados relacionados, por exemplo, a pagamentos de assistência social, impacto da estrutura de assistência social, dados de identificação exclusiva de cidadãos e dados relacionados ao serviço público.
- b) **Dados abertos:** preocupações em superar barreiras à acessibilidade, trazendo uniformidade na maneira como os dados são capturados, armazenados, processados e gerados. Dados abertos são quaisquer dados que podem ser usados livremente, reutilizados e redistribuídos por qualquer pessoa.
- c) **Governo aberto:** um governo pode ser denominado governo aberto se mostrar uma tendência de governança, priorizando a abertura de seus dados. A abertura também envolve a criação de atos de direito à informação, mais liberdades civis e mais transparência (KUMAR, 2016, p. 16).

Para maior compreensão, a Figura 3 demonstra os fundamentos dos DAGs:

Figura 3 - Fundamentos dos DAGs



Fonte: Gonzalez-Zapata e Heeks (2015, p. 5).

Quando se trata de governo aberto, existem princípios que se manifestam, conforme elencados por Ramírez-Alujas e Dassen (2014):

- a) efetuar melhoria dos níveis de transparência e acesso à informação ao abrir dados públicos (para exercer controle social e responsabilidade) e a reutilização de informações do setor público, a fim de promover a inovação e o desenvolvimento econômico;
- b) facilitar a participação dos cidadãos na criação e implementação de políticas públicas;
- c) favorecer a geração de espaços de colaboração entre os diversos atores, principalmente entre administrações públicas, sociedade civil e setor privado.

No ano de 2007, Larry Lessig e mais 29 pessoas consideradas ‘pioneiras em dados abertos’ estabeleceram um conjunto de oito valores que definem como o governo poderia fornecer os dados de maneira a capacitar outras pessoas a revisar, analisar e agir sobre os dados, tendo como objetivo final a melhoria dos serviços públicos (AYRE; CRANER, 2017).

Conhecidos como os oito princípios fundamentais dos dados abertos governamentais, esses valores determinam que um dado do governo somente será considerado aberto se forem divulgados publicamente, de acordo com os seguintes valores:

1. **Completo:** todos os dados públicos são disponibilizados. Dados públicos são dados que não estão sujeitos a limitações válidas de privacidade, segurança ou privilégio;
2. **Primário:** os dados são coletados na fonte, com o nível mais alto possível de granularidade, não em formas agregadas ou modificadas;
3. **Oportuno:** os dados são disponibilizados o mais rápido possível para preservar o valor dos dados;
4. **Acessível:** Os dados estão disponíveis para a maior variedade de usuários para a maior variedade de finalidades.
5. **Processável por máquina:** os dados são razoavelmente estruturados para permitir o processamento automatizado;
6. **Não discriminatório:** os dados estão disponíveis para qualquer pessoa, sem necessidade de registro;
7. **Não proprietário:** os dados estão disponíveis em um formato sobre o qual nenhuma entidade tem controle exclusivo;
8. **Sem licença:** os dados não estão sujeitos a nenhum direito autoral, patente, marca comercial ou regulamento de segredo comercial. Restrições razoáveis de privacidade, segurança e privilégios podem ser permitidas (AYRE; CRANER, 2017, p. 182, tradução nossa).

Para fins da presente pesquisa, destaca-se o Princípio 5 – Processável por máquina, o qual significa que os dados são razoavelmente estruturados para permitir

o processamento automatizado. Esse princípio orienta não somente a escolha do formato do arquivo, como também orienta como o formato deve ser usado. No Quadro 4 são apresentados os formatos recomendados para publicação de DAGs para diferentes tipos de mídia.

Quadro 4 - Formatos recomendados para publicação de DAGs

Mídia	Formato de dado recomendado
Planilhas	CSV (UTF-8); planilha OpenOffice ou XML
Documentos	XHTML/XML (estrutura); PDF/A-1b (paginação)
Imagens	JPEG, SVG ou PNG (as appropriate)
Áudio	WAV (sem compressão) ou Ogg Vorbis
Vídeo	AVI (sem compressão) ou Ogg Theora

Fonte: Tauberer (2014).

A abertura de dados públicos está intrinsecamente relacionada à criação da *Freedom of Information Act (FOIA)*, ou Lei de Liberdade de Informação. Pode ser definida como o direito de acessar informações mantidas por órgãos públicos, sendo parte integrante do direito fundamental à liberdade de expressão, conforme reconhecido em Resolução nº 59 da Assembleia Geral da ONU, adotada em 1946 e pelo Art. nº 19 da Declaração Universal dos Direitos Humanos de 1948 (UNESCO, 2017).

As Leis associadas ao *FOIA* refletem a premissa fundamental de que todas as informações mantidas por governos e instituições governamentais são, em princípio, públicas e só podem ser retidas por razões como privacidade e segurança. Quanto à adesão ao *FOIA*, em 1990, apenas 13 países adotaram Leis nacionais de *FOIA*, enquanto atualmente existem mais de 90 Leis adotadas em todo o mundo (UNESCO, 2017). As Leis *FOIA* determinam quem pode acessar dados públicos governamentais, quais informações estão disponíveis, quanto tempo os governos devem responder e quais penalidades resultam quando o governo descumpe a legislação. Além disso, a implementação das Leis *FOIA* frequentemente acompanha movimentos de democratização e anticorrupção em todo o mundo e o seu início ocorreu com a chamada ‘Década da Abertura’, que durou desde o final da Guerra Fria até aos ataques terroristas nos Estados Unidos da América (EUA) (FINK, 2018).

Dados públicos abertos são uma ferramenta poderosa e um derivado da *FOIA*. Sem o *FOIA*, os dados abertos não podem cumprir o que prometem: a verdadeira liberdade de informação (IZDEBSKI, 2015). Em 2009, o governo Obama esboçou um conjunto de princípios abertos do governo dos EUA que abrange três objetivos

fundamentais: colaboração, participação e transparência. Com o objetivo de divulgar os dados do governo americano ao contribuinte, de modo que possam ser usadas para gerar valor social e econômico, instituiu-se o *Open Government Directive*, um conjunto de diretrizes as quais exigem que todas as agências do governo americano publiquem conjuntos de dados em um site disponível ao público, fornecendo contexto e metadados importantes por meio de relatórios, resumo de serviços públicos, *links* para os principais programas e conjuntos de dados para *download* (DAWES; HELBIG, 2010).

No Brasil, de acordo com Oliveira e Lóscio (2014), a Lei de Acesso à Informação Pública, sancionada em 18 de novembro de 2011, e que começou a vigorar em 16 de maio de 2012, permitiu a qualquer pessoa ter acesso a documentos e informações que estejam sob a guarda de órgãos públicos, em todos os poderes e níveis do governo. Além de regulamentar o fornecimento de informações sob demanda do cidadão, o normativo também determina que órgãos e entidades devem se antecipar aos pedidos e publicar seus dados e informações na internet, bem como exige que os dados sejam publicados em formatos abertos e não-proprietários (OLIVEIRA, 2020).

Mais recentemente, o Brasil realizou oficinas de cocriação, entre maio e agosto de 2018, para elaborar o 4º Plano de Ação Nacional de Governo Aberto. O resultado dessas rodadas de reuniões, integradas por membros do governo e da sociedade civil, foi a definição de onze compromissos brasileiros, dentre os quais se destaca o número 2 sobre o ecossistema de dados abertos. Esse compromisso objetivou estabelecer, de forma colaborativa, um modelo de referência de política de dados abertos que promova integração, capacitação e sensibilização entre sociedade e as três esferas de governo, a partir do mapeamento das demandas sociais (OLIVEIRA, 2020).

A finalidade do documento Modelo de Referência de Abertura de Dados instituído pelo governo federal consiste em integrar, capacitar e sensibilizar a sociedade e as três esferas de governo no contexto da abertura de dados públicos e orientar gestores, agentes públicos e sociedade quanto à importância, ao uso, à publicação, à sustentação e ao monitoramento de dados abertos nos seus respectivos campos de atuação. A ideia é apresentar diretrizes que possam auxiliar estados e municípios a disponibilizar e a selecionar, de forma aberta, um conteúdo mínimo de informações e temas de sua competência, com base em pesquisas realizadas com especialistas e em boas experiências nacionais. O documento estabelece duas

abordagens distintas: a primeira buscou ouvir especialistas em diversas áreas para levantar demandas de cada setor com relação à abertura de dados nos níveis estadual e municipal. A segunda abordagem levou em consideração agendas amplas de discussão na arena pública, que demandam uma série de indicadores de estados e municípios, a saber: dados associados a dimensões de competência específica de cada ente subnacional; associados aos 17 ODS; dados que permitam o aprimoramento dos serviços públicos e dados que permitam o fomento à inovação e aos negócios (OLIVEIRA, 2020).

Outras iniciativas também surgiram com a criação da Lei de Acesso à Informação e muitas são as possibilidades com relação ao uso de dados abertos públicos como, por exemplo, promover a transparência quanto às ações dos governos. O lançamento do Portal Brasileiro de Dados Abertos possibilitou a interlocução entre atores da sociedade e governo, viabilizando a localização dos dados em formato aberto a partir de uma única fonte e lugar, por meio de uma ferramenta de busca. Segundo o entendimento de Barbalho (2018), algumas funções da transparência para o setor público são: confiança no governo e nas instituições; participação cidadã; planejamento estratégico; direcionamento de produtos e serviços e exercício da cidadania.

As Plataformas de Dados Abertos podem ajudar na criação de políticas públicas e consistem uma poderosa abordagem para tornar os dados localizáveis, interoperáveis e legíveis por máquina, melhorando drasticamente a eficiência de análises e *insights* (KLEIN; KLEIN; LUCIANO, 2018a).

No entanto, no Brasil, apesar de existirem programas bastante avançados de transparência pública – como, por exemplo, de transparência orçamentária –, ainda são raríssimos os órgãos ou secretarias que disponibilizam dados abertos. No melhor dos casos, há dados disponíveis para visualização, mas existem inúmeras barreiras técnicas, e até políticas, para que sejam reutilizados pela sociedade na criação de novos projetos e serviços (WORLD WIDE WEB CONSORTIUM, 2011, p. 11).

Ainda quanto à aplicação dos dados abertos públicos a W3C, em seu Manual de Dados Abertos, elenca algumas áreas de atuação para as quais são importantes dados abertos, como transparência e controle democrático; participação popular; empoderamento dos cidadãos; melhores ou novos produtos e serviços privados; inovação; melhora na eficiência e efetividade de serviços governamentais; medição

de impacto das políticas; conhecimento novo a partir da combinação de fontes e padrões (WORLD WIDE WEB CONSORTIUM, 2011).

O Quadro 5 mostra exemplos quanto ao uso dos dados DAGs no Brasil e no mundo. É importante salientar que para criar alguns desses serviços, no caso do Brasil, os desenvolvedores tiveram de ‘raspar’ informações de sítios públicos, ou seja, é como se a sociedade, maior interessada no uso do dado, estivesse fazendo todo o trabalho sozinho.

Quadro 5 - Iniciativas que fazem uso de DAGs

Iniciativa	Descrição
Citymapper	Dados sobre ônibus, trens e metrô, publicados em tempo real pela autoridade de trânsito londrina para formular um algoritmo que mostra sempre o melhor tempo de viagem entre um ponto e outro
Influence Explorer	Rastreia doações políticas em nível estadual e federal, permitindo que qualquer pessoa possa acompanhar o grau de influência por legislador, empresa ou indivíduo. A ferramenta oferece uma visão geral sobre financiamentos de campanha, lobby, verbas parlamentares, irregularidades em contratações e gastos públicos federais.
Radar Parlamentar	Aplicativo que ilustra as semelhanças entre partidos políticos com base na análise matemática dos dados de votações que ocorrem na casa legislativa.
CoronApp	Aplicativo que utiliza geolocalização para informar casos de contaminação pelo novo Coronavírus ao redor do celular em que está instalado.
Serenata de Amor	Usa tecnologia baseada em inteligência artificial para oferecer serviços de fiscalização dos gastos públicos e compartilhar informações com a sociedade.
Colaboradados	Iniciativa colaborativa que busca reunir, demonstrar, investigar e monitorar o acesso à informação no Brasil.
Appelou	Aplicativo baseado nas informações coletadas a partir da nota fiscal do consumidor eletrônica (NFC-e), que são armazenadas e utilizadas no sistema Economiza Alagoas.

Fonte: OLIVEIRA (2020, p. 47).

Os setores públicos e privados, por meio de movimentos conhecidos como *hackathons*, vêm estimulando a criação de soluções tecnológicas que utilizam DAGs. São também chamados de *hackday*, *hackfest* e *codefest*. Porém, dois grandes problemas são conhecidos ao se realizar essas competições: incentivar os desenvolvedores a transformar seus aplicativos interessantes em produtos

sustentáveis e ter público para usá-los. Assim, muitos aplicativos produzidos são rapidamente abandonados (OLIVEIRA; LÓSCIO, 2014).

Quando se fala do uso dos DAGs, gestores públicos geralmente enfrentam desafios e barreiras na implementação de projetos, os quais estão relacionados às múltiplas interações entre os atores, fluxos de informação, tecnologias e interesses nessas iniciativas. Albano (2014) elencou seis grandes grupos de barreiras e possíveis fatores inibidores: institucionais, complexidades das tarefas, uso e participação, legislação, qualidade da informação e técnicas, conforme descrito no Quadro 6:

Quadro 6 - Barreiras e possíveis fatores inibidores

Categoría	Barreiras ou fator inibidor
Institucionais	Ênfase de barreiras e negligência das oportunidades.
	Confusão entre os valores públicos (valores de transparência versus privacidade).
	Cultura de aversão ao risco (ao empreendedorismo).
	Falta de uniformidade nas políticas de publicação dos dados Institucionais.
	Publicação sem agregar valor aos dados.
	Falta de recursos específicos para a tarefa de publicar os dados.
	Sistema da receita é baseado na criação de renda a partir de dados.
	Priorizar interesses das organizações em detrimento dos interesses dos cidadãos.
	Falta de definição de processos para trabalhar e/ou interagir com o usuário.
	Qualidade discutível para trabalhar com o usuário.
Complexidade da tarefa	Falta de capacidade de descobrir os dados apropriados.
	Nenhuma explicação do significado dos dados.
	Falta de informações sobre a qualidade dos dados abertos.
	Aplicativos escondendo a complexidade, mas também potencial uso dos dados abertos.
	Duplicação de dados, os dados disponíveis em várias formas, ou antes/depois do processamento, resultando em discussões sobre qual é a fonte.
	Dificuldade em pesquisar e navegar por índices ou outros meios para garantir a fácil pesquisa para encontrar os dados corretos.
	Mesmo que os dados possam ser encontrados, os usuários podem não estar cientes de seus usos potenciais.
	Formatos de dados e conjuntos de dados são demasiadamente complexos para lidar e usar facilmente.
	Falta de suporte a ferramentas ou <i>helpdesk</i> foco.
	Foco está em fazer uso de conjuntos de dados individuais, ao passo que o valor real pode vir da combinação de vários conjuntos de dados.
Uso e participação	Resultados contraditórios baseados na utilização dos mesmos dados.
	Conclusões inválidas.
	Falta de incentivos para os usuários.
	Organizações públicas não reagem à entrada do usuário.
	Falta de informações sobre a qualidade dos dados abertos.
	Frustração com a existência de muitas iniciativas de dados.
	Falta de tempo para se aprofundar nos detalhes para prover maior entendimento.

	Falta de tempo para usar dados abertos. Falta de conhecimento para fazer uso ou para dar sentido ao dado. Falta de capacidade necessária para usar a informação. Nenhum conhecimento estatístico ou compreensão do potencial e das limitações das estatísticas.
Legislação	Violação de privacidade. Falta de segurança. Nenhuma licença para usar os dados, o que acaba limitando o uso. Disputas e litígios. Permissão prévia por escrito exigida para ter acesso aos dados e para reproduzi-los. Reutilização de contratos/convenções.
	Dados semelhantes armazenados em sistemas diferentes produzem resultados diferentes. Falta de precisão da informação.
	Dados obsoletos e não válidos fazem parecer que a informação é irrelevante quando vista isoladamente, mas quando ligada e analisada em conjunto pode resultar em novos conhecimentos.
	Informações incompletas, apenas parte do quadro total mostrado, ou apenas de um determinado intervalo.
	Falta de informação. Muita informação para processar e sem saber o que exatamente olhar Informação essencial perdida.
Técnicas	Os dados devem estar em um formato bem definido que é facilmente acessível: enquanto o formato de dados é arbitrário, o formato de definições de dados precisa ser rigorosamente definido. Ausência de padrões. Ausência de um portal centralizador. Ausência de suporte para disponibilização de dados. Falta de metas e padrões. Nenhum software padrão para o processamento de dados abertos. Fragmentação de software e aplicações. Sistemas legados que complicam a divulgação de dados.

Fonte: Albano (2014, p. 36).

Em pesquisa realizada por Martin *et al.* (2013), foram analisados três casos de publicação de dados abertos por meio de plataformas de cidades como França, Berlim e Reino Unido, cujo objetivo era traçar uma tipologia de desafios, riscos, limitações e barreiras, além de identificar como as barreiras foram superadas e os riscos gerenciados. Os riscos foram divididos em sete categorias: governança; questões econômicas; licenças e estruturas legais; características dos dados; metadados; acesso e habilidades.

Já Pedroso, Tanaka e Capelli (2013) identificaram barreiras na área de interoperabilidade, ao apresentar perspectivas de pesquisa envolvendo a Lei de Acesso à Informação e os desafios tecnológicos na área de DAGs e transparência. Os autores citam como exemplos de barreiras para publicação de DAGs, na área de interoperabilidade, a divulgação de dados em PDF ou em arquivos de imagem digitalizados como o maior entrave; e a ausência de divulgação em formato legível por

máquinas que também acarreta um retrabalho para os órgãos, que são obrigados a recriar documentos em PDF e imagens digitalizadas e publicá-los com a frequência exigida por Lei.

Quanto aos possíveis benefícios do uso de DAGs, Albano e Craveiro (2015) os classificam em três grandes grupos: políticos e sociais; econômicos e técnicos e operacionais, conforme apresentado no Quadro 7:

Quadro 7 - Possíveis benefícios

Categoria	Benefícios
Políticos e sociais	Promover mais transparência, maior participação e autorresponsabilização dos cidadãos, possibilitar maior e mais fácil acesso a dados públicos (governos), promover maior participação da sociedade (usuários), aprimorar a formulação de políticas públicas, promover estímulo ao desenvolvimento do conhecimento no setor público e criar novos serviços públicos para a sociedade (inovação).
Econômicos	Uso do conhecimento coletivo, sobre as informações públicas, possibilitando a criação de novos produtos e serviços, criados pela sociedade, fazendo uso de recursos de inovação. <u>Geração de novas oportunidades econômicas.</u>
Técnicos	Fatores relacionados à capacidade de reutilizar os dados, ou seja, não precisar recolher os mesmos novamente, evitando possíveis duplicações e custos associados, o que facilita algumas tarefas, como criar uma base de dados e integrar dados privados e públicos.
Operacionais	Possibilita otimização dos processos administrativos (nos governos), usando capacidades da sociedade para resolver problemas, possibilitando a melhora na tomada de decisão.

Fonte: Albano e Craveiro (2015, p. 29).

É notório que dados abertos implementados e mantidos por governos e suas agências, organizações da sociedade civil e setor de negócios oferecem diversos benefícios, como a criação de novos produtos e serviços e geração de novas oportunidades econômicas. Diante disso, surgiram alguns exemplos da utilização de dados abertos com a finalidade econômica, como a iniciativa do Banco Mundial de instituir um concurso para indicar a melhor solução (aplicativo desenvolvido) criada a partir de seus dados (dados que o banco publica em formato aberto) (ALBANO, 2014). Quando se fala em mercado de dados, estima-se que o seu valor se aproxima de 40 a 140 bilhões de euros anualmente (IZDEBSKI, 2015).

A respeito da qualidade dos DAGs, alguns trabalhos abordam a avaliação da qualidade das informações no setor público brasileiro como Albuquerque (2009), ao apresentar um modelo de avaliação para aferição de qualidade das informações publicadas em Portais dos Tribunais de Justiça, os quais têm vital importância para a comunidade que acessa diariamente o conteúdo judiciário e necessita de mecanismos e metodologias para avaliar a qualidade das informações disponíveis em seus sites.

Araujo (2012) realizou uma revisão na literatura e verificou quais aspectos da qualidade da informação são contemplados nas iniciativas de abertura e disponibilização de Bases de Dados Governamentais. A avaliação da qualidade da informação levou em consideração 16 dimensões evidenciadas no modelo desenvolvido no trabalho de Kahn, Strong e Wang (2002). A intenção do estudo foi evidenciar a necessidade da discussão das atuais iniciativas de DAGs, com enfoque na qualidade da informação (ARAUJO, 2012).

Trabalhos como Vetrò *et al.* (2016) focaram na medição da qualidade de DAGs por meio de uma estrutura de indicadores para mensurar a qualidade baseada em uma série de dimensões de qualidade de dados no nível mais granular de medição. Os autores aplicaram a abordagem a dois estudos de casos com conjunto de DAGs italianos. Os resultados mostraram diferenças na qualidade relacionadas ao uso de boas práticas e problemas comuns a um conjunto de dados. Com base na avaliação, os pesquisadores puderam fornecer diretrizes técnicas e políticas para superar os problemas observados, principalmente na política descentralizada de publicação de dados.

A próxima seção aborda uma visão geral sobre os princípios FAIR, tendo como ponto de partida uma breve introdução ao assunto, uma descrição detalhada sobre os princípios e uma discussão sobre os pontos convergentes entre os princípios e as dimensões da qualidade da informação, já discutidas anteriormente no presente trabalho.

2.3 PRINCÍPIOS FAIR E DADOS ABERTOS

Melhorar a qualidade de acesso aos dados pode ser um facilitador essencial para ciência, tecnologia e inovação. Um exemplo são as iniciativas para compartilhar dados de pesquisa, aumentando o ritmo da descoberta de conhecimento e do progresso científico. A reutilização de dados de pesquisa tem o potencial de evitar a duplicação de conjuntos de dados e trazer novas perspectivas a partir de múltiplas análises do mesmo conjunto de dados. Um exemplo são os estudos que utilizam dados de diversas fontes para analisar as variações genômicas associadas ao câncer, com o objetivo de selecionar qual a terapia mais apropriada para um paciente específico (FIGUEIREDO, 2017).

Em 2016, a *Scientific Data*, uma publicação focada em dados, publicou o *FAIR Guiding Principles for Scientific Data Management and Stewardship*. Os autores pretendiam fornecer diretrizes para melhorar a localização, acessibilidade, interoperabilidade e reuso dos ativos digitais. Os princípios enfatizam a capacidade dos sistemas computacionais de encontrar, acessar, interoperar e reutilizar os dados com pouca ou nenhuma intervenção humana, uma vez que se exige um maior suporte computacional para lidar com o crescente volume de dados (GOFAIR, 2020).

Segundo Wilkinson *et al.* (2016, p. 4), os princípios *FAIR* são:

Os princípios orientadores do *FAIR* precedem as escolhas de implementação e não sugerem nenhuma tecnologia, padrão ou solução de implementação específica; além disso, os Princípios não são, eles próprios, um padrão ou uma especificação. Eles atuam como um guia para editores de dados e administradores para ajudá-los a avaliar se suas opções de implementação específicas estão tornando seus artefatos de pesquisa digital localizáveis, acessíveis, interoperáveis e reutilizáveis.

Os princípios *FAIR* defendem o aumento da localização, acessibilidade, interoperabilidade e reutilização de dados de pesquisa e objetos digitais acadêmicos em geral. Sob a égide da sigla *FAIR*, um acrônimo para *Findable, Accessible, Interoperable and Reusable* (SCIENCE VIDENPORTAL, 2020), foram formulados 15 princípios para orientar as ações dos editores, administradores de dados e outras partes interessadas, os quais estão listados no Quadro 8.

Quadro 8 - Guia dos princípios *FAIR*

Princípios	Como aplicar
F - Findable	Para serem encontrados.
F1. Os (meta)dados devem ter identificadores globais, persistentes e identificáveis	Metadados com identificador único - Unique and Persistent Identifier (PID) ou equivalente, para cada conjunto de dados. Ex.: DOI, ARK, RRID, PID.
F2. Os dados devem ser descritos com metadados enriquecidos (impacta diretamente R1)	Conjuntos de dados devem possuir metadados ricos o suficiente para que, uma vez indexados para um mecanismo de busca, esses metadados possam ajudar o usuário dos dados a encontrá-los mesmo que não possua o seu identificador.
F3. Os metadados devem incluir claramente e explicitamente os identificadores dos dados que descrevem	Como não se pode prever que os dados e seus metadados estejam sempre juntos (por exemplo quando os metadados são indexados por um mecanismo de busca e, portanto, está em uma plataforma diferente dos dados), a associação entre eles deve ser feita pela inclusão do identificador dos dados pelos metadados.
F4. Os (meta)dados devem ser registrados ou indexados em recursos que ofereçam capacidades de busca	Para que contribuam para que os dados sejam encontrados, seus metadados devem ser indexados por mecanismos de busca que, por sua vez, permitem aos usuários encontrá-los por meio de elementos desses metadados.
A - Accessible	Para serem acessíveis

A1. (Meta)dados devem ser recuperáveis pelos seus identificadores usando protocolo de comunicação padronizado	Com o identificador do conjunto de dados ou de seus metadados, o usuário deve poder recuperá-los por meio de um protocolo de comunicação padronizado.
A1.1 O protocolo deve ser aberto, gratuito e universalmente implementável	Independente de licenciamento dos dados e dos metadados, o protocolo de comunicação usado para dar acesso a eles deve ser aberto, gratuito e passível de ser implementado por qualquer interessado.
A1.2 O protocolo deve permitir procedimentos de autenticação e autorização, quando necessário	Quando necessário, dependendo das restrições ao acesso aos dados e/ou metadados, um mecanismo para autenticação e autorização para o acesso deve ser permitido pelo protocolo de comunicação
A2. Metadados devem ser acessíveis, mesmo quando os dados não estão mais disponíveis	É preciso existir um conjunto de estratégias de preservação para os dados e metadados. Minimamente, os metadados devem ser sempre acessíveis, possibilitando a criação de índices para conjuntos de dados atuais e passados.
I - Interoperable	Para serem interoperáveis
I1. (Meta)dados devem ser representados por meio de uma linguagem formal, acessível, compartilhada e amplamente aplicável para a representação do conhecimento	Para que seja possível representar dados e metadados deve-se utilizar linguagens de representação do conhecimento que sejam formais, acessíveis e amplamente aplicáveis. Por exemplo, RDF, XML, DICOM etc.
I2. (Meta)dados devem usar vocabulários que seguem os princípios FAIR	Como os dados e metadados devem possuir referências aos vocabulários que contenham os conceitos utilizados, deve-se garantir que sejam utilizados vocabulários que também sigam os princípios FAIR.
I3. (Meta)dados devem incluir referências qualificadas para outros (Meta)dados	Referenciar os conjuntos de dados devidamente, possibilitando que conjuntos de dados gerados, a partir de outros conjuntos de dados, sejam ligados. Possibilitar a ligação semântica entre eles por meio dos padrões adotados.
R - Reusable	Para serem reutilizáveis
R1. (Meta)dados são descritos com uma pluralidade de atributos precisos e relevantes	Prover metadados que permitam aos potenciais usuários uma pluralidade de atributos precisos e avaliarem se é possível reusar os dados ou metadados relevantes, e se adequam às suas necessidades.
R1.1. (Meta)dados devem ser disponibilizados com licenças de uso claras e acessíveis	É fundamental que o responsável pelos dados e metadados defina explicitamente quem pode ter acesso a eles, para que e sob quais condições. Essas informações são definidas por meio de suas licenças de uso.
R1.2. (Meta)dados devem estar associados a proveniência	Especificar a proveniência (linhagem) dos dados é importante não só para que o usuário possa avaliar a utilidade dos dados ou metadados, mas também para que se possa atribuir o devido crédito a quem produziu, manteve ou editou esses dados. Dentre as informações relevantes à proveniência destaca-se: (a) A linhagem dos dados, ou seja, o processo de obtenção do dado (gerado ou coletado); (b) Particularidades ou limitações sobre os dados que outros usuários devem conhecer; (c) Data da geração do conjunto de dados, condições de laboratório, quem preparou os dados, configurações de parâmetros, nome e versão do software utilizado; (d) Explicitar se são dados brutos ou processados; (e) A versão dos dados arquivados e/ou reutilizados deve ser claramente especificada.
R1.3. (Meta)dados devem estar alinhados com padrões relevantes ao seu domínio.	Atender os padrões específicos da comunidade da área. Atender às boas práticas de arquivamento e ao compartilhamento do campo de pesquisa específico.

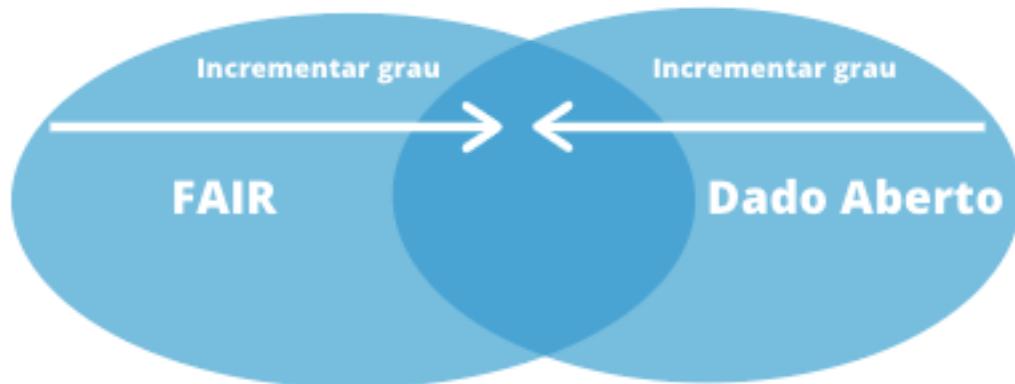
Fonte: Adaptado de Henning *et al.* (2018. p. 399-403).

O ponto central do conceito *FAIR* é sua aplicação para atividades guiadas por humanos e orientadas por máquinas. Além disso, *FAIR* não é binário (ou seja, *FAIR*/não *FAIR*), mas, sim, um espectro ao longo do qual são possíveis variáveis ‘*FAIRness*’, isto é, o máximo quanto for possível sua aplicação (HIGMAN; BANGERT; JONES, 2019).

Destaca-se que os princípios *FAIR* se aplicam aos dados e seus metadados, ou seja, registros sobre conjunto de dados. Por isso, o termo ‘metadados’ é declarado nos princípios. Outro aspecto importante sobre os princípios *FAIR* é que não se referem somente a dados abertos, sendo possível trabalhar com dados privados, por exemplo. Os princípios *FAIR* não representam um padrão de qualidade para avaliação de ferramentas, dados, políticas, entre outros, como também sua implementação pode ser uma adaptação gradual e sistemática de novas rotinas de trabalho (HANSEN, 2018).

Outra questão a ser enfatizada é que embora muitas políticas exijam dados abertos e se baseiem nos princípios *FAIR*, os dois não significam a mesma coisa. Os dados podem ser abertos e estar alinhados aos princípios *FAIR*, apenas um deles ou nenhum dos dois. Ambos conceitos podem ser compreendidos como escalas para representar se um dado pode ser cada vez mais aberto ou *FAIR*, conforme mostra a Figura 4:

Figura 4 - Relacionamento entre *FAIR* e dados abertos

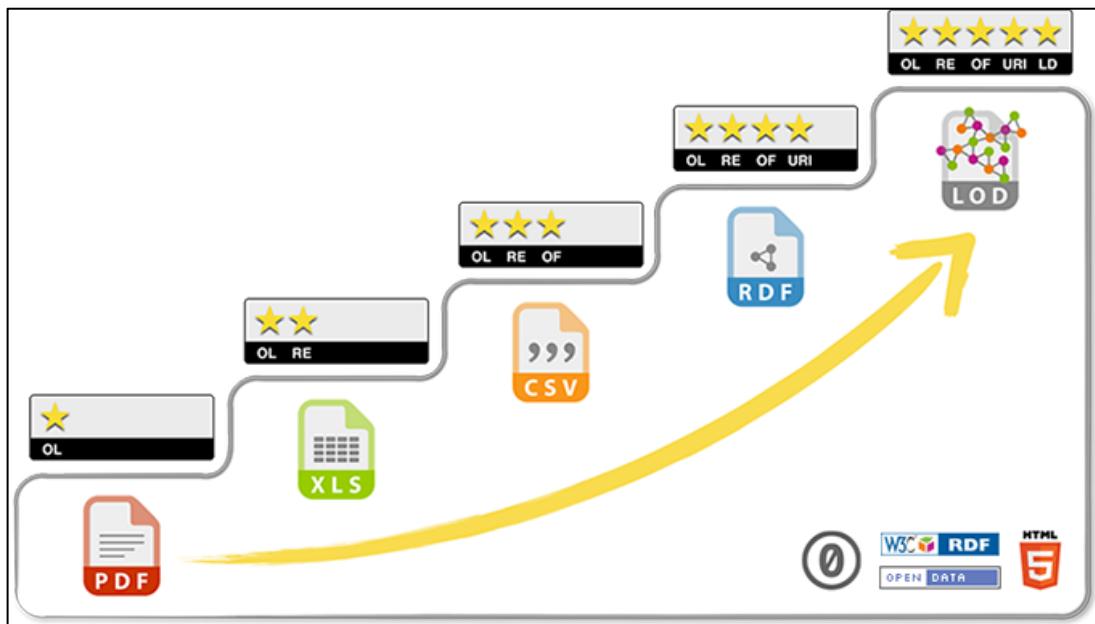


Fonte: Higman, Bangert e Jones (2019, p. 3).

No ano de 2010, com o objetivo de classificar o grau de abertura e conexão dos dados, Tim Berners-Lee propôs cinco princípios conhecidos como ‘Sistema de 5 estrelas’ ou *5-Star Open Data Principles*. Os princípios apresentam um esquema, o

qual determina que quanto maior o número de estrelas, maior será o grau de abertura e conexão dos dados (VILELA, 2018), como pode-se verificar na Figura 5.

Figura 5 - 5-Star Open Data



Fonte: 5-Star Open Data (2012).

O esquema de 5 estrelas fornece diretrizes para provedores e editores de dados, de modo a tornar os dados mais acessíveis, disponíveis e reutilizáveis na web. O Quadro 9 cada uma das cinco estrelas (VILELA, 2018):

Quadro 9 - Definição de cada uma das cinco estrelas dos dados abertos

Número de estrelas	Descrição
*	Disponível na Internet, em qualquer formato, desde que com licença aberta.
**	Disponível na Internet de maneira estruturada.
***	Disponível na Internet de maneira estruturada e em formato não proprietário.
****	Mesmas regras anteriores, além da utilização de URIs para identificar os recursos na web.
*****	Mesmas regras anteriores além da conexão com outros dados com o intuito de fornecer um contexto para os dados.

Fonte: Vilela (2018, p. 24).

De acordo com Hasnain e Rehholz-Schuhmann (2018), assim como a estrutura 5-Star, os princípios FAIR são incrementais e quanto melhor a adoção das diretrizes FAIR, maior o grau de 'FAIRness'. O autor faz um comparativo entre os princípios 5-Star e FAIR, conforme o Quadro 10.

Quadro 10 - Mapeamento dos princípios FAIR com esquema 5-Star

Princípios FAIR	*	**	***	****	*****
F - Findable					
F1. Os (meta)dados devem ter identificadores globais, persistentes e identificáveis.				X	
F2. Os dados devem ser descritos com metadados enriquecidos (impacta diretamente R1).					X
F3. Os metadados devem incluir claramente e explicitamente os identificadores dos dados que descrevem.				X	
F4. Os (meta)dados devem ser registrados ou indexados em recursos que ofereçam capacidades de busca.	X				
A - Accessible					
A1. (Meta)dados devem ser recuperáveis pelos seus identificadores usando protocolo de comunicação padronizado.		X	X	X	
A1.1 O protocolo deve ser aberto, gratuito e universalmente implementável.	X	X	X	X	
A1.2 O protocolo deve permitir procedimentos de autenticação e autorização, quando necessário.	X				
A2. Metadados devem ser acessíveis, mesmo quando os dados não estão mais disponíveis.					
I - Interoperable					
I1. (Meta)dados devem ser representados por meio de uma linguagem formal, acessível, compartilhada e amplamente aplicável para a representação do conhecimento.	X	X	X		
I2. (Meta)dados devem usar vocabulários que seguem os princípios FAIR.	X	X			X
I3. (Meta)dados devem incluir referências qualificadas para outros (Meta)dados.				X	
R - Reusable					
R1. (Meta)dados são descritos.		X			
R1.1. (Meta)dados devem ser disponibilizados com licenças de uso claras e acessíveis.	X				
R1.2. (Meta)dados devem estar associados a proveniência.				X	X
R1.3. (Meta)dados devem estar alinhados com padrões relevantes ao seu domínio.	X	X			X

Fonte: Adaptado de Hasnain e Rehholz-Schuhmann (2018)

As abordagens 5-Star e FAIR possuem pontos convergentes, no entanto, ressalta-se que os princípios FAIR cobrem um escopo mais amplo e incluem dados não abertos. Já o esquema 5-Star são orientados para abertura dos dados. Quanto às definições de qualidade no contexto de dados abertos, essas variam consideravelmente. Por exemplo, o Portal Europeu de Dados considera como alta qualidade dos dados abertos se 'os seres humanos puderem entendê-los e as máquinas puderem manipulá-los', e estiverem alinhados com o sistema de classificação 5-Star Open Data. Outros, como o G8 Open Data Charter e os Open

Data Institute Certification Badges, se concentram no fornecimento de metadados, descrições de esquemas de dados, uso de dicionários de dados compartilhados, licença usada, formato de arquivo e suporte ao editor para interagir com os usuários de dados (CORSAR; EDWARDS, 2017).

Com relação aos princípios *FAIR*, conforme comentado anteriormente, não se trata de um padrão de qualidade para avaliação de ferramentas, dados, nem políticas, mas pode ser vista como uma recomendação que qualifica a aptidão para a reutilização de dados de maneira imparcial e mensurável. Para fins da presente pesquisa, realizou-se um estudo comparativo com os atributos de QI, *FAIR* e 5-Star, com o objetivo de mapear os pontos convergentes entre as abordagens *FAIR*, 5-Star e os atributos para avaliar Qualidade da Informação (QI) e fornecer base conceitual para o modelo proposto no presente trabalho, conforme indicado no Quadro 11.

Quadro 11 - Comparativo entre os atributos para avaliar a QI, princípios *FAIR* e 5-Star Open Data

Nº	Atributos de QI	FAIR	5-Star
1	Acessibilidade - a facilidade e a eficiência que o usuário pode navegar na aplicação para acessar a informação de desejada (KAHN; STRONG; WANG, 2002; KIM, KISHORE; SANDERS, 2005).	A1, A1.1, A1.2 e A2	* ** *** ****
2	Clareza - a informação é facilmente entendida (KAHN; STRONG; WANG, 2002); Completeza - o sistema provê toda a informação necessária para a execução da atividade (KAHN; STRONG; WANG, 2002; SHANKARANARAYAN; ZIAD; WAND, 2003; WIXON; TODD, 2005; KIM, KISHORE; SANDERS, 2005); completeza ficou junto com clareza.	Não atende	Não atende
3	Concisão - a informação é apresentada de forma compacta (KAHN; STRONG; WANG, 2002).	Todos princípios	Todos princípios
4	Consistência - a informação é apresentada no mesmo formato (KAHN; STRONG; WANG, 2002; WIXON; TODD, 2005).	Todos princípios	Todos princípios
5	Credibilidade - a informação é considerada verdadeira e confiável (KAHN; STRONG; WANG, 2002).	Não atende	Não atende
6	Disponibilidade - refere-se à flexibilidade e compreensão das características da aplicação disponibilizada para os usuários para especificar e controlar a relação temporal entre os vários componentes hipermídia para que seja disponibilizada uma informação de hipermídia integrada (KIM; KISHORE; SANDERS; 2005).	F1, F2, F3 e F4	* **** *****
7	Empacotamento - refere-se à variedade de informações que os vários tipos de mídias são empacotados dentro de uma interface web para apresentação ao usuário final (KIM; KISHORE; SANDERS, 2005).	I1, I2 e I3	* ** *** **** *****
8	Facilidade de uso - a informação é fácil de manipular e se aplica a diferentes atividades (KAHN; STRONG; WANG, 2002).	Não atende	Não atende
9	Histórico - refere-se à flexibilidade e compreensão das características da aplicação disponibilizadas para os usuários para especificar e	Não atende	Não atende

	manter um histórico das ações dos usuários e do estado da aplicação (KIM; KISHORE; SANDERS, 2005).		
10	Interface - consistência no arranjo estrutural e no estilo do conteúdo da informação e hiperlinks dentro da aplicação, devendo permitir que o usuário conheça o conteúdo da informação disponibilizada nas páginas web (KIM; KISHORE; SANDERS, 2005).	Não atende	Não atende
11	Interpretabilidade - a informação está na linguagem e símbolos apropriados e as definições estão claras (KAHN; STRONG; WANG, 2002).	F1, F2, F3 F4, I1, I2, I3 e R1.3	* ** *** **** *****
12	Objetividade - a informação é não tendenciosa e imparcial (KAHN; STRONG; WANG, 2002).	Não atende	Não atende
13	Precisão - a informação está correta, isenta de erros (KAHN; STRONG; WANG, 2002; WIXON; TODD, 2005; KIM; KISHORE; SANDERS, 2005).	Não atende	Não atende
14	Prontidão - a informação é fornecida sempre que for necessária; Quantidade - O quanto o volume de informação é apropriado para atividade que está sendo executada (KAHN; STRONG; WANG, 2002).	F1, F2, F3, F4 e A2	* **** *****
15	Relevância - a informação deve estar relacionada com o interesse e as necessidades de informação do usuário (KAHN; STRONG; WANG, 2002; KIM; KISHORE; SANDERS, 2005).	Não atende	Não atende
16	Reputação - a informação é considerada verdadeira com relação a sua fonte ou conteúdo (KAHN; STRONG; WANG, 2002).	Não atende	Não atende
17	Segurança - o acesso à informação é mantido restrito apropriadamente para garantir a sua segurança (KAHN; STRONG; WANG, 2002).	A1.2	*
18	Temporalidade - a informação está suficientemente atualizada (KAHN; STRONG; WANG, 2002; WIXON; TODD, 2005; KIM; KISHORE; SANDERS, 2005).	Não atende	Não atende
19	Valor - a informação provê benefícios e vantagens com o seu uso (KAHN; STRONG; WANG, 2002).	Não atende	Não atende

Fonte: Elaborado pela autora (2020).

Foi possível identificar, dentre os atributos listados, que alguns não possuíam características correspondentes nos princípios *FAIR*, como também no esquema 5-*Star*. Importante destacar o atributo 'Acessibilidade', que tem conceito diferente entre os princípios *FAIR* e 5-*Star*. Para o primeiro, acessibilidade significa que o ser humano ou a máquina recebe, por meio de metadados, as condições precisas pelas quais os dados são acessíveis (COLLINS *et al.*, 2018), não tendo como pré-requisito o dado ser aberto; já dentro do esquema 5-*Star* o conceito de acessível corresponde à disponível na Internet, em qualquer formato, desde que com licença aberta.

Nesta seção foi feito um breve levantamento sobre os princípios *FAIR*, com enfoque na comparação entre *FAIR*, dados abertos e qualidade da informação. Na próxima seção será abordada a iniciativa *Frictionless Data*, a qual tem sido usada para tratar problemas de publicação de dados abertos.

2.4 FRICTIONLESS DATA E DATA PACKAGE

Frictionless Data é uma iniciativa da *OKF* que pretende remover o ‘atrito’ no trabalho com os dados. Entende-se como ‘atrito’ ou ‘fricção’ quando os consumidores de dados perdem tempo e recursos demais apenas para entender e trabalhar com as informações. Por outro lado, os produtores de informação encontram dificuldades para descrever os tipos de dados permitidos em cada coluna de uma planilha, por exemplo (OPEN KNOWLEDGE FOUNDATION, 2019).

Ainda segundo o site oficial da iniciativa, *Frictionless Data* é um *progressive framework* para construção de infraestrutura de dados – gerenciamento, integração, fluxo, entre outros, ou seja, ao contrário de outros *frameworks*, *Frictionless Data* foi projetado para ser adotado de forma incremental e ‘progressiva’. O objetivo do *framework* é trabalhar, desenvolver e aprimorar os dados e ferramentas existentes usadas no trabalho com dados, ao invés de substituí-los. Logo, não se trata de um grande banco de dados ou um sistema de gerenciamento de dados, mas, sim, uma abordagem que auxilia na automatização de tarefas repetitivas sobre conjunto de dados, como obtenção, limpeza e utilização dos dados.

Uma boa analogia para entender o atrito no trabalho com os dados é o processo de fazer um bolo, conforme exemplifica Rufus Pollock, criador e fundador do *OKF*: ao decidir cozinhar, tem-se os ingredientes prontamente disponíveis no mercado ou na sua própria cozinha. Não é necessário viajar para uma fazenda, coletar ovos, moer o milho, fazer o bacon. Ao invés disso, graças à logística, transporte (contêineres) e pagamento, os ingredientes estão disponíveis em um supermercado, sem necessidade de buscá-los na fonte. O problema é que, quando se trata de dados, ainda existe grande custo para obter e trabalhar com eles. Ao realizar uma análise ou construir um aplicativo, por exemplo, dependendo do conjunto de dados necessário, precisa-se descobrir a fonte de dados, extraí-los, limpá-los e prepará-los antes que possa iniciar o trabalho (POLLOCK, 2017).

O *Frictionless Data* oferece, ainda, uma gama de ferramentas, padrões e boas práticas para publicação de dados. O *Frictionless Data* se baseia em princípios, tais como (OPEN KNOWLEDGE FOUNDATION, 2019):

- a) Foco: concentra em tipos específicos de dados;

- b) orientado à web: usa formatos que funcionam bem na web, como o JSON, formato comumente usado para troca de dados por APIs;
- c) distribuição: projeto com visão de um ecossistema sem centralização de dados;
- d) abertura: qualquer pessoa pode usar e reusar livremente e abertamente os dados;
- e) uso de ferramentas existentes: integração com ferramentas tecnológicas existentes;
- f) simplicidade: manutenção dos formatos e metadados de forma simples e leve, fácil de interpretar e usar.

O Quadro 12 apresenta alguns estudos de caso com aplicações do *Frictionless Data* encontrados na literatura:

Quadro 12 - Estudos de caso com aplicações de *Frictionless Data*

Artigo/Caso de Uso	Descrição	Ano	Fonte/Autores
Developing a modern data workflow for regularly updated data	Workflow com uma combinação de ferramentas, incluindo a iniciativa <i>Frictionless Data</i> , para reduzir erros de dados, importar e reestrutura dados, arquivar e criar versões dos dados e automatizar a maioria das etapas do pipeline de dados para reduzir o tempo e esforço necessários aos pesquisados.	2019	Glenda Yennil, Erica Christensenl, Ellen K. Bledsoe, Sarah R. Suppl., Renata Diazl Ethan P. Whitel, S. K. Morgan, Ernest
<i>Frictionless Data: Making Research Data Quality Visible</i>	Artigo dos pesquisadores participantes do projeto <i>Frictionless data</i> , que traz o estado da arte da iniciativa e relata casos de uso, como foco na qualidade de dados.	2018	Fowler, Dan Barratt, Jo Walsh, Paul
TEDDINET	Como parte do Data Management for TEDDINET (DM4T), iniciativa da Open Knowledge International, em parceria com a University of Bath, para testar o uso de especificações e ferramentas de <i>Frictionless Data</i> para conjuntos de dados criados por meio do TEDDINET (Transforming Energy Demand for Digital Innovation NETwork). O objetivo do projeto é usar <i>Frictionless Data</i> para preparar e publicar dados de pesquisa, a fim de facilitar sua reutilização.	2018	Fowler, Dan Barratt, Jo Walsh, Paul
Active Data Biology at Pacific Northwest National Laboratory We	Projeto que desenvolveu uma ferramenta de análise de dados colaborativa para dados biológicos. Iniciativa usou <i>Frictionless Data</i> para tratar restrições de tipos e valor de dados.	2018	Fowler, Dan Barratt, Jo Walsh, Paul

Archaeology Working Group	Grupo de Trabalho de Arqueologia Aberta para explorar abordagens abertas para trabalhar com dados arqueológicos.	2018	Fowler, Dan Barratt, Jo Walsh, Paul
---------------------------	--	------	-------------------------------------

Fonte: Elaborado pela autora (2020).

Os autores Wiese *et al.* (2019) usaram *Frictionless Data* para resolver o problema da grande quantidade de dados fornecidos pelas empresas pertencentes ao Sistema Europeu de Energia com documentação insuficiente, além de dados em formatos inconsistentes. O trabalho elencou alguns desafios que, com o uso do *Frictionless Data*, obtiveram resultados satisfatórios para resolução de problemas relacionados à identificação de fontes apropriadas, a diferentes fontes de dados, à falta de nomenclatura padronizada e classificação de fontes de energia, à falta de metadados e à qualidade dos dados (inconsistências, erros, lacunas nos dados).

A ideia central da iniciativa *Frictionless Data* é o conceito de ‘conteinerização de dados’, ou seja, a forma como o dado é compartilhado segue o conceito de contêiner como se entende no transporte de mercadorias em um navio de carga, por exemplo. O objetivo é desenvolver padrões, especificando os contêineres de modo a reduzir os custos e tempo para transporte de dados (OPEN KNOWLEDGE FOUNDATION, 2019).

Um conceito-chave presente na iniciativa *Frictionless Data* é o *Data Package* ou pacote de dados. Trata-se de uma forma para simplificar o agrupamento de dados e suas descrições em um único local, com a finalidade de tornar os dados mais facilmente compartilhados e usados. O formato do pacote de dados é muito simples, amigável e extensível (OPEN KNOWLEDGE FOUNDATION, 2019).

Um *Data Package* consiste em: metadados que descrevem a estrutura e o conteúdo do pacote, além de recursos como arquivo de dados que formam o conteúdo do pacote. Os metadados do pacote são armazenados em um ‘descritor’, que por sua vez transforma uma coleção de dados em um pacote de dados (OPEN KNOWLEDGE FOUNDATION, 2019). A estrutura desse descritor é o conteúdo principal da especificação do pacote de dados. A especificação do *Data Package* não impõe nenhum requisito à sua forma ou estrutura e, portanto, pode ser usada para ‘empacotar’ qualquer tipo de dado, contribuindo para a publicação de dados, por exemplo (OPEN KNOWLEDGE FOUNDATION, 2019).

O descritor é o arquivo central em um pacote de dados e fornece metadados gerais como título do pacote, licença, editor etc. Inclui também uma lista dos recursos

de dados que compõem o pacote, incluindo sua localização online ou em disco físico. Um descritor do *Data Package* deve ser um objeto JSON válido e quando disponível como um arquivo, ele deve ser denominado *datapackage.json*. O descritor pode estender outras propriedades além das descritas como propriedades obrigatórias e opcionais. Por exemplo, se os dados contêm séries temporais e seja necessário listar a cobertura temporal dos dados no *Data Package*, é possível adicionar uma propriedade temporal. Essa flexibilidade permite a extensão do *Data Package* conforme a necessidade. A seguir estão as propriedades comumente usadas que o descritor *Data Package* pode conter, assim como as propriedades opcionais:

a) propriedades recomendadas:

- *name*: um nome que pode ser usado por URL; deve ser minúsculo e conter apenas caracteres alfanuméricos junto com caracteres “.”, “_” ou “-”; essa propriedade funciona como um identificador único;
- *id*: uma propriedade reservada para identificadores exclusivos globalmente. Exemplos de identificadores exclusivos incluem UUIDs e DOIs;
- *license*: uma ou mais licenças sob as quais o pacote é fornecido; esta propriedade não é legalmente vinculativa e não garante o que o pacote seja licenciado de acordo com os termos definidos nesta propriedade.
- *profile*: identifica o perfil deste descritor de acordo com a especificação de perfis.

b) propriedades opcionais:

- *title*: fornece um título ou uma descrição para o *Data Package*;
- *description*: fornece uma descrição do *Data Package*; a descrição permite texto simples; o primeiro parágrafo (até a primeira quebra de linha dupla) deve ser usado como um resumo para descrever o *Data Package*;
- *resources*: representam as fontes brutas para um *Data Package*; possui o título e um propriedade com o endereço do recurso ou um *email*;
- *contributors*: pessoas ou organizações que contribuíram para o *Data Package*; cada colaborador deve ter uma propriedade de título, além de *email*, função e nome da organização a qual pertence (FRICTIONLESS DATA SPECS, [2020]).

O processo de compartilhamento de dados usando *Data Package* possui, em apenas uma camada, uma estrutura de transporte de dados básica que reduz consideravelmente o ‘atrito’ na integração e distribuição de dados, apoiando a automação, sem impor mudanças drásticas nos dados para serem empacotados. A estrutura, por sua natureza simples e leve, torna fácil sua adoção pelos publicadores de conteúdo, usuário de dados e criadores de ferramentas de software (OPEN KNOWLEDGE FOUNDATION, 2019).

Existem diversas iniciativas de ‘empacotamento’ de dados que compartilham um modelo de ‘arquivos em disco’ para compactar dados com metadados. Estes incluem *BagIt* e *CSVW* na web do W3C. *BagIt* é um formato de pacote de conjunto de dados popular entre as bibliotecas, incluindo a Biblioteca do Congresso nos EUA, projetado para oferecer suporte ao armazenamento baseado em disco e à transferência de rede de conteúdo digital arbitrário (KUNZE *et al.*, 2016). *BagIt* se concentra no arquivamento de ativos digitais (principalmente conteúdo), enquanto o *Data Package* se concentra em descrever a estrutura dos dados empacotados. Além disso, um elemento-chave da abordagem de *Data Package* é a integração com as ferramentas existentes e a extensibilidade para muitos tipos de dados, ambos os quais não são prioridades para *BagIt* (KUNZE *et al.*, 2016). *CSVW* é o resultado de um grupo de trabalho W3C concluído em 2016, é mais diretamente comparável. O *CSVW* começou como um esforço para padronizar a especificação do então protótipo do Tabular *Data Package* com várias modificações para garantir a compatibilidade com a web (FOWLER; BARRATT; WALSH, 2017).

Na próxima seção será apresentado o modelo baseado em *Frictionless Data* para colaborar com a publicação e compartilhamento de DAGs, além de uma contextualização sobre os detalhes do modelo e sua aplicação.

3 MODELO BASEADO EM *FRictionless DATA* PARA DAGs

Nesta seção será apresentada uma visão geral do modelo baseado em *Frictionless Data*, para auxiliar na publicação e compartilhamento de DAGs. Também será demonstrada uma contextualização sobre os sujeitos da pesquisa, contexto e coleta de dados; as ferramentas utilizadas para aplicação prática do modelo, tendo como subsídios para a pesquisa os DAGs presentes nas BBPODS da ONU divulgados pelo CNJ; a aplicação do modelo usando as ferramentas disponibilizadas pela iniciativa *Frictionless Data*; bem como a apresentação dos resultados da pesquisa.

3.1 VISÃO GERAL DO MODELO

Nesta seção será apresentado o modelo proposto para publicação de DAGs baseado em *Frictionless Data*. Antes, contudo, são apresentados alguns trabalhos relacionados à pesquisa, a fim de comparar o modelo proposto com abordagens semelhantes para resolver problemas inerentes a publicação de DAGs.

3.1.1 Trabalhos relacionados

Um levantamento na literatura sobre publicação de DAGs mostrou trabalhos como o de Ávila (2015), que propôs um modelo de processo de publicação de dados considerando o nível de maturidade da instituição publicadora, contribuindo com as estratégias de produção e disponibilização de dados abertos e dados abertos conectados do setor público.

No entanto, apesar de o trabalho considerar o nível de maturidade das instituições e contemplar várias etapas do processo de publicação de dados abertos, não houve maior detalhamento técnico (envolvendo ferramentas, por exemplo) de como deve ser desenvolvida a publicação de dados abertos. Consequentemente, não contempla nenhuma recomendação referente às boas práticas para publicação de dados abertos de natureza mais técnica (ÁVILA, 2015).

Trabalhos, como o de Dzikrullah e Rinjani (2017), apresentam o modelo chamado *SmartGov* para integração de sistemas de dados baseados em *Big Data* com foco em e-gov. A proposta é facilitar a troca de DAGs entre agências do governo, criando um ambiente de serviços integrados. O foco do modelo foi resolver problemas

de interoperabilidade de dados e, apesar do modelo contemplar diversas etapas inerentes à publicação de dados abertos, como uma camada chamada *Open Data Layer*, a qual utiliza o *Comprehensive Knowledge Archive Network* (CKAN), a plataforma desenvolvida pela OKF não possui uma etapa de *feedback* para reportar problemas de qualidade encontrado nos dados.

O estudo de Ruijer, Grimmelikhuijsen e Meijer (2017), apresenta um modelo heurístico para compreender a relação entre DAGs e as atividades dos cidadãos e governos do ponto de vista dos processos democráticos monitorial, deliberativo e participativo. O modelo adapta a abertura de dados abertos de acordo com o tipo de processo democrático e ajuda a identificar desafios do uso de dados abertos para a democracia. O trabalho tem como ponto forte considerar a complexidade dos processos democráticos, mas não apresenta etapa para tratamento de qualidade dos dados, apesar de mencionar que a importância do fornecimento de informações e a qualidade dos dados são cruciais.

Em Saxena (2016) foi proposto um modelo para avaliar a intenção comportamental para o uso dos DAGs. Foi realizada uma análise quantitativa usando regressão múltipla para analisar 341 catálogos com dados governamentais. O ponto forte desse trabalho foi verificar a popularidade de um conjunto de dados abertos entre usuários finais.

Já o modelo proposto por Zhu e Freeman (2019) forneceu uma estrutura para avaliar o desempenho dos portais que publicam DAGs, com foco nas possíveis interações e experiências dos usuários com os portais, no que diz respeito ao uso dos portais. O ponto forte do estudo foi a avaliação de 34 portais baseada em critérios abrangentes que tinham como objetivo avaliar o senso de confiança dos usuários nos portais, além da capacidade de compreender o conteúdo, capacidade de integrar dados e formas de participação do cidadão. Porém, ambos modelos mencionados com foco em participação cidadã não contemplam outras etapas importantes para publicação de dados abertos, como o tratamento da informação publicada.

No artigo de Tambouris (2016) foi delineado um processo para publicação, expansão e consumo de DAGs usando um modelo multidimensional de dados, como a abordagem de cubo de dados juntamente com ferramentas de *Linked Data*. Como ponto forte do trabalho, seguindo o processo proposto, os produtores de dados podem transformar dados brutos em cubos de dados em formato RDF. Com isso, os consumidores de dados podem encontrar, visualizar, mesclar e analisar dados com

maior facilidade. No entanto, o trabalho não abrange outros processos fundamentais para publicação de DAGs, como tratamento dos dados e *feedback* dos usuários.

No trabalho de Brandt *et al.* (2018) foi elaborado um modelo de dados abertos conectados (*Linked Open Data - LOD*) para um conjunto de dados abertos legislativos da Câmara dos Deputados. O modelo foi elaborado com base no conjunto real de dados fornecidos pela Câmara dos Deputados usando *RDF* para estruturar os dados. Destaca-se como ponto forte do trabalho a contribuição para áreas da CI, tais como a representação da informação e do conhecimento, com o uso de vocabulários e ontologias para representar e trazer significado aos relacionamentos entre as entidades modeladas. No entanto, a pesquisa não contempla questões como alinhamento com melhores práticas para publicação de dados e uso de metadados.

O modelo proposto por Kalampokis, Tambouris e Tarabanis (2011) possui duas dimensões principais: complexidade organizacional, tecnológica e valor agregado para os consumidores de dados. A ideia era complementar os modelos de governo eletrônico existentes para reuso de DAGs, com o objetivo de fornecer um roteiro para reutilização e DAGs e permitir a avaliação das iniciativas relevantes. O ponto forte do modelo foi a integração dos dados e, além das duas dimensões mencionadas, possui quatro fases: agregação de dados governamentais; integração de dados governamentais; integração de dados governamentais com dados sociais e formais não governamentais. Entretanto, não houve implementação de um protótipo para cada estágio do modelo, como também não foram contempladas o tratamento dos dados.

Na próxima seção apresenta-se o modelo para publicação de DAGs baseado em *Frictionless Data*, explicando como ele foi concebido, o detalhamento das etapas inerentes ao modelo e a metodologia utilizada para o seu desenvolvimento.

3.1.2 Apresentação do modelo e metodologia

Os órgãos públicos recolhem, geram e armazenam enormes quantidades de dados, que são necessários para o serviço ao cidadão, à tomada de decisões, à concepção de políticas públicas, à reutilização no seu interesse social e econômico, ou então pela transparência da atividade pública (OBSERVATORIO DE ADMINISTRACIÓN ELECTRÓNICA, 2017).

Do ponto de vista econômico e social, embora as competições de dados abertos conhecidas como *hackathons* tenham dado origem a algumas histórias de

sucesso em termos de *startups* e aplicativos, há também algumas evidências de que o tempo para atribuir valor e trabalhar esses conjuntos de dados permanece longo, principalmente devido à falta de conhecimento sobre as características de qualidade dos dados e o esforço resultante de tornar os dados prontos para uso (SADIQ; INDULSKA, 2017).

Ao mesmo tempo, os metadados, bem como a qualidade dos dados subjacentes para esses conjuntos de dados são conhecidos por serem deficientes. Por exemplo, muitos conjuntos de dados abertos têm dados duplicados, inconsistentes e ausentes, geralmente carecendo de descrições de esquema facilmente acessíveis (SADIQ; INDULSKA 2017).

Com relação ao ecossistema necessário para abertura de dados abertos, de acordo com Zuiderwijk, Janssen e Davis (2014), para criar um ecossistema de dados abertos, pelo menos quatro elementos-chave devem ser capturados, quais sejam:

- a) liberação e publicação de dados abertos na Internet;
- b) pesquisa, localização, avaliação e visualização de dados e suas licenças relacionadas;
- c) limpeza, análise, enriquecendo, combinando, vinculando e visualizando dados;
- d) interpretando e discutindo dados e fornecendo *feedback* para o provedor de dados e outras partes interessadas.

Além disso, para integrar os elementos do ecossistema e permitir que eles atuem como um todo integrado, deve haver três elementos adicionais: caminhos do usuário mostrando direções para como os dados abertos podem ser usados; um sistema de gestão da qualidade; e diferentes tipos de metadados para ser capaz de conectar os elementos (ZUIDERWIJK; JANSSEN; DAVIS, 2014).

Sob a perspectiva técnica, para Zuiderwijk, Janssen e Parnia (2013), as infraestruturas de dados abertos devem fornecer ferramentas e mecanismos que permitam o compartilhamento, troca e reutilização de dados, como também permitir a descoberta, transformação, visualização e *download* de dados, de preferência gratuitamente, garantindo a interoperabilidade por meio do uso de metadados. Ainda conforme os autores, abordagens relacionadas ao *Linked Data* fornecem parte da infraestrutura que pode suportar o fornecimento de dados abertos, porém uma

infraestrutura baseada em *Linked Data* não é suficiente, uma vez que não atende aos requisitos de todo o processo de dados abertos.

Ainda sob a ótica técnica, de acordo com Zuiderwijk, Jeffery e Janssen (2012), o processo de dados abertos encontra impedimentos que foram categorizados em 10 categorias:

1. Disponibilidade e acesso: dados não são coletados ou sem detalhes; dados com licença proprietária; dados não estão disponíveis, apenas mediante solicitação; dados desatualizados.
2. Capacidade de encontrar: inexistência de mecanismo de pesquisa avançada; dificuldade de pesquisa e navegação; dados não podem ser encontrados em formato legível por máquina.
3. Usabilidade: falta de serviços prestados pelo provedor de dados para usar dados brutos; dados requerem carga de trabalho humana substancial para serem limpos e torná-los comprehensíveis por máquina;
4. Capacidade de compreensão: nenhuma explicação sobre o significado dos dados; conjunto de dados incompletos;
5. Qualidade: dados não são confiáveis; difícil determinar a qualidade dos dados; dados limitados ou de má qualidade; falta de precisão dos dados.
6. Ligar e combinar dados: dados não podem ser vinculados a outros conjuntos de dados;
7. Comparabilidade e compatibilidade: diferenças nas definições dos dados; diferentes tipos de políticas de dados abertos; fontes diferentes fornecem informações diferentes sobre o mesmo assunto;
8. Metadados: inexistência de metadados sobre a forma como os dados foram coletados;
9. Interação com o provedor de dados: nenhuma discussão entre o provedor de dados e o usuário de dados é possível; nenhuma plataforma ou mecanismo de participação disponível.
10. Abertura e *upload*: dificuldades com o upload de conjuntos de dados; falta de princípios orientadores e suporte para abertura de dados; processo de dados abertos não é visto como um processo de interação entre governo e cidadão.

Já sob o enfoque da qualidade de dados, um dos maiores riscos relacionados ao uso de dados abertos é a falta de consciência da qualidade inerente aos dados. Os dados abertos são frequentemente usados para uma finalidade que não foi originalmente pretendida no momento da coleta de dados, portanto, um conjunto de dados que pode ser de qualidade suficiente para um propósito pode não ser adequado para outro propósito. Os consumidores de dados abertos normalmente não são os produtores e, portanto, não há uma estratégia bem definida para a limpeza de dados, o que muitas vezes resulta em curadoria e transformação equivocada de dados (SADIQ; INDULSKA, 2017).

Os consumidores de dados abertos podem, portanto, investir esforços significativos para gerar resultados valiosos a partir dos dados apenas para perceber que os resultados são inadequados, ou pior, podem não perceber que os dados são de baixa qualidade e confiar em resultados errôneos. Tem-se uma necessidade crítica de ferramentas exploratórias e abordagens que permitam aos usuários estar cientes das deficiências dos dados em termos de uso pretendido (SADIQ; INDULSKA, 2017).

Do ponto de vista legal e sob o enfoque das políticas para publicação de dados abertos, tem-se que muitos governos e organizações internacionais disponibilizaram grandes coleções de dados ao público, que quer se beneficiar de uma licença padrão que permite o uso extensivo dos dados, ou que adaptam suas próprias licenças nacionais aos padrões de licenciamento promovidos por organizações como a OKF.

Em todas essas iniciativas de dados abertos, alguns princípios básicos são assumidos, sobre os quais as informações devem ser publicadas e as licenças devem ser configuradas. Segundo a ONU, esses princípios são: informação acessível na Internet; em formatos processáveis eletronicamente; licença aberta que permite a reutilização comercial; gratuito e sem restrições de reutilização (RAMOS-SIMÓN, 2017).

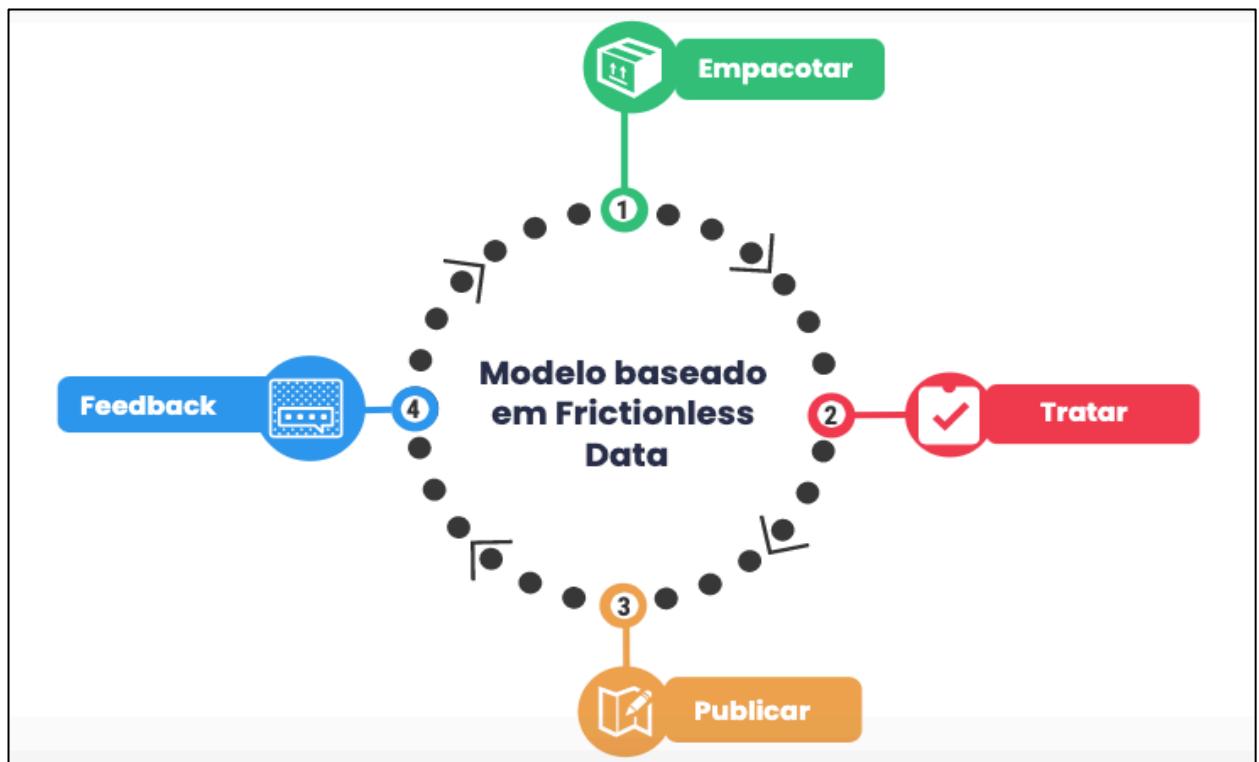
Diante do exposto, surgiu a ideia do presente trabalho para contribuir com a publicação de DAGs através de um modelo baseado em *Frictionless Data*, iniciativa da OKF para tratar o atrito em base de dados. O modelo também tem como base, além dos estudos de trabalhos relacionados, o guia com os princípios *FAIR* mencionados anteriormente no trabalho como suporte para atender as características principais de interoperabilidade; o *Frictionless Data Spec*, um conjunto de especificações para uso e publicação de dados usando *Frictionless Data*; os

elementos-chave para a criação de um ecossistema de dados abertos conforme exposto por Zuiderwijk, Janssen e Davis (2014); a aderência às melhores práticas para publicação de dados abertos e ao Modelo de Referência para Publicação de Dados Abertos do governo federal brasileiro (BRASIL, 2020).

As etapas do modelo foram delineadas de forma a responder ao problema delimitado e os objetivos centrais desta dissertação. Ou seja, o modelo pretende auxiliar a publicação de DAGs, tendo em vista diminuir o tempo que se gasta trabalhando com os DAGs, assim como promover a qualidade da informação publicada e fornecer uma estrutura básica que permita qualquer cidadão publicar e usar conjuntos de dados com mais facilidade e sem passar por um intermediário.

A Figura 6 mostra as etapas do modelo baseado em *Frictionless Data* para publicação de DAGs.

Figura 6 - Modelo baseado em *Frictionless Data* para publicação de DAGs



Fonte: Elaborado pela autora (2020).

As quatro etapas serão apresentadas com mais detalhes nas subseções apresentadas a seguir.

3.1.2.1 *Etapa empacotar*

No modelo proposto, a etapa empacotar se refere à transformação de um conjunto de dados em um pacote de dados, de modo a permitir a interoperabilidade e o reuso dos dados. Os procedimentos para esta etapa estão descritos a seguir:

- a) conversão do conjunto de dados em formato aberto;
- b) transformação do conjunto de dados em *Data Package*.

O procedimento “conversão do conjunto de dados em formato aberto” tem como principal objetivo converter o conjunto de dados em um formato aberto, de acordo com as três leis dos dados abertos: os dados precisam ser encontrados e indexados na web; devem estar em formato aberto e compreensível por máquina e sem restrições quanto à utilização (WORLD WIDE WEB CONSORTIUM, 2009). Destaca-se que não é objetivo da presente pesquisa indicar tecnologias para conversão dos dados em formato aberto.

Já o procedimento “transformação do conjunto de dados em *Data Package*” tem como fundamento a especificação *Frictionless Data*, que contém o conceito chamado *Data Package*, explicado em seções anteriores com mais detalhes. Porém, destaca-se que um *Data Package* consiste em: metadados que descrevem a estrutura e o conteúdo do pacote de dados e recursos como arquivos de dados que formam o conteúdo do pacote (FRICTIONLESS DATA SPECS, 2017). Os metadados do *Data Package* são armazenados em um descritor. Esse descritor é o conteúdo principal da especificação *Frictionless Data*. Além do descritor, um *Data Package* incluirá outros recursos, como arquivo de dados (FRICTIONLESS DATA SPECS, 2017).

Um dos pontos fortes da especificação é que o *Data Package* não impõe nenhum requisito em sua forma ou estrutura e, portanto, pode ser usada para empacotar qualquer tipo de dados (FRICTIONLESS DATA SPECS, 2017). Como os órgãos públicos disponibilizam grande quantidade de informações, porém essas bases geralmente estão em diferentes formatos ou em formatos que impossibilitam acesso pelas partes interessadas aos dados (ARAÚJO *et al.*, 2012), essa etapa auxilia na disponibilização dos dados independente do formato.

A filosofia de dados abertos não define exaustivamente uma lista de formatos permitidos e, para ser considerado um dado aberto, o conjunto de dados deve estar disponível em um formato de especificação aberta, não proprietário, e estruturado, ou seja, que possibilite seu uso irrestrito e automatizado através da Web. Além disso, é

imprescindível que seja utilizado um formato amplamente conhecido. É importante enfatizar a importância dos dados estruturados, pois um erro recorrente cometido por diversas instituições é a publicação em formato PDF de planilhas de dados. O PDF é um formato não estruturado, e ao fazer isso – desestruturação dos dados – o publicador está inviabilizando, ou dificultado, a reutilização daqueles dados (BRASIL, 2020).

Outra vantagem dessa etapa é o alinhamento aos 15 princípios *FAIR*, os quais defendem o aumento da localização, acessibilidade, interoperabilidade e reutilização de dados. Apesar do foco inicial dos princípios *FAIR* serem os dados de pesquisa e objetos digitais acadêmicos em geral, podem ser usados para qualquer tipo de dado. Logo, trata-se de uma contribuição relevante da presente pesquisa a aplicação inédita dos princípios *FAIR* aos DAGs, por meio do modelo baseado em *Frictionless Data*.

3.1.2.2 *Etapa tratar*

Outra problemática é com relação à qualidade das informações presentes nas bases de DAGs. A difusão dos DAGs por parte das instituições públicas se manteve em um ritmo muito acelerado, no entanto, a divulgação de dados sem controle de qualidade adequado prejudica a sua reutilização e afeta negativamente a participação cívica (VETRÒ *et al.*, 2016).

Pesquisas mostram que existem diversas interpretações e maneiras diferentes de medir a qualidade de dados, como comentado anteriormente no presente trabalho. Geralmente, as abordagens incluem interpretabilidade de dados, precisão, oportunidade de publicação, confiabilidade, acessibilidade, capacidade de descoberta e processamento ou integridade. No entanto, uma vez que as pessoas usam dados para finalidades distintas, certas qualidades de dados são mais importantes para um grupo de usuários que para outros.

Muitas métricas de qualidade de dados são focadas nos usuários, porém é fundamental que o governo, como produtor de dados, compreenda, monitore e melhore a qualidade inerente dos dados que produz. A *Open Knowledge International* tem como alvo os produtores de dados e problemas de qualidade dos arquivos de dados, principalmente por meio da iniciativa *Frictionless Data* (OPEN KNOWLEDGE FOUNDATION, 2017).

GoodTables.io é um serviço *web* para validação contínua da qualidade dos dados e permite a verificação da qualidade de um conjunto de dados de forma automática (FOWLER; BARRATT; WALSH, 2017). A ferramenta permite a verificação de erros e problemas comuns em arquivos tabulares, aumentando a qualidade e sua confiança no conjunto de dados (OPEN KNOWLEDGE FOUNDATION, [2020]). Na presente pesquisa, ao propor o modelo baseado nas especificações contidas na iniciativa *Frictionless Data*, sugere-se que a etapa tratar conte cole aspectos essenciais de qualidade para arquivos de dados tabulares contidos no *Data Quality Spec*, projeto apoiado pela *Open Knowledge International*, além do uso da ferramenta *GoodTables.io* para automatização da etapa Tratar, cujo objetivo é validar o conjunto de dados para posterior publicação.

Por fim, para essa etapa se recomenda atender à Boa Prática 6 do documento que contém as melhores práticas para abertura de dados, expostas anteriormente no Quadro 3, que descreve a importância de prover informações sobre a qualidade de dados, também descritos no Modelo de Referência para publicação de dados abertos do governo federal (BRASIL, 2020).

3.1.2.3 *Etapa publicar*

O principal desafio é que os dados abertos não têm valor em si, ele só se torna valioso quando usado (JANSSEN; CHARALABIDIS; ZUIDERWIJK, 2012). Além disso, o estabelecimento de portais de dados governamentais pode mudar a forma como os cidadãos acessam, recuperam e usam esses dados (LOURENÇO, 2015; SAXENA, 2016). Resta saber se esses portais de dados estão apresentando dados de uma forma que permita aos cidadãos usá-los de forma eficaz, criando, assim, os resultados esperados pelas agências governamentais.

Por causa dessa questão, vários autores apontam para a importância de avaliar os resultados e impactos das iniciativas de DAGs. Eles sugerem que essa avaliação é necessária para fornecer evidências empíricas da eficácia dos DAGs, para orientar a prática e a formulação de políticas e para aumentar a confiança do público (SAXENA, 2016). Muitas organizações públicas aderiram ao movimento de disponibilizar dados sem ter uma política sólida. Isso resultou na publicação de dados em portais centrais que já estavam disponíveis publicamente, como dados de endereço (JANSSEN; CHARALABIDIS; ZUIDERWIJK, 2012). No entanto, muitas

vezes os dados sequer estão disponíveis em formatos abertos nas Plataformas e Portais Governamentais, sendo necessário que o cidadão de maneira autônoma busque, trabalhe os dados e os disponibilize em sítios alternativos. Para isso, na presente pesquisa, utilizou-se a ferramenta *DataHub*, plataforma de gerenciamento de dados criada pela *Open Knowledge International* baseada no sistema de gerenciamento de dados CKAN.

O *DataHub* fornece acesso gratuito a muitos dos principais recursos do CKAN, permitindo: pesquisa de dados, o registro de conjuntos de dados publicados, a criação e gerenciamento de grupos de conjuntos de dados e acompanhar as atualizações dos conjuntos de dados e grupos de interesse. Esses recursos estão disponíveis por meio de uma interface web ou uma API, a CKAN API.

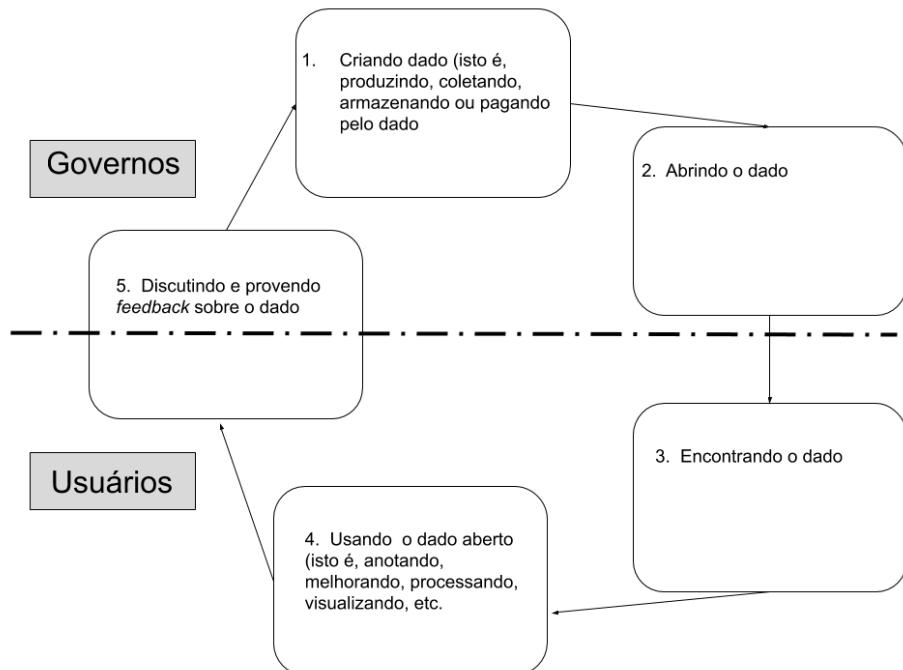
3.1.2.4 *Etapa feedback*

O aprimoramento da qualidade pode ser alcançado pelo acréscimo de participação e o inerente *feedback* sobre as publicações (KLEIN; KLEIN; LUCIANO, 2018b). Segundo Janssen, Charalabidis e Zuiderwijk (2012), a implicação da noção de *feedback* na teoria dos sistemas é que, ao abrir os dados, os governos não devem simplesmente instigar a comunicação unilateral de seus dados, mas devem esperar ou solicitarativamente o *feedback* e ser capazes de dar sentido a esse *feedback*. A abertura dos dados oferece a oportunidade de criar ciclos de *feedback* nos quais o governo pode aprender com o público.

De acordo com Janssen, Charalabidis e Zuiderwijk (2012), além disso, no processo de dados abertos ideal, informações de *feedback* sobre o uso dos dados são fornecidas aos órgãos públicos que produziram os dados, para que esses dados possam ser usados para melhorar os processos de trabalho, como processos de formulação de políticas de órgãos públicos.

No presente trabalho, incluiu-se a etapa *feedback* com o objetivo de tornar o modelo aderente aos padrões de processo para abertura de dados. Logo, seguiu-se o modelo recomendado por Zuiderwijk *et al.* (2012), conforme demonstrado na Figura 7:

Figura 7 - Modelo ideal de processo para publicação de DAGs



Fonte: Zuiderwijk *et al* (2012, p. 157).

Buscou-se cobrir a lacuna encontrada dentre os modelos pesquisados na literatura acrescentando a etapa *feedback* ao modelo proposto, sem a intenção de determinar qualquer uso específico de ferramentas ou processo, priorizando a flexibilização do uso da tecnologia, ferramenta ou metodologia, conforme a necessidade do órgão público. No entanto, para fins da presente pesquisa, reforça-se que a iniciativa *Frictionless Data* ainda não possui uma ferramenta que permita verificar o ciclo de *feedback* de dados abertos. Porém, entende-se que a ausência de uma ferramenta na iniciativa *Frictionless Data* para a verificação da etapa *feedback* não invalida a importância de sua presença no modelo e, como forma de contribuição da presente pesquisa sugere-se, como possibilidade de verificação da etapa *feedback*, a inclusão do conceito de *Net Promoter Score* (NPS) como forma de acompanhar o nível de satisfação dos usuários em relação à qualidade dos dados presentes nos serviços de informação.

O NPS foi apresentado por Reichheld (2003), uma forma simples de medir a fidelidade do cliente. O NPS se baseia na seguinte pergunta: Qual a probabilidade de você recomendar uma marca ou empresa a um amigo ou colega? Classificados em uma escala de 0 a 10. Os clientes são agrupados em 'promotores' (classificação de 9-10), 'passivamente satisfeitos' (classificação de 7-8) e 'detratores' (classificação de

0-6). Segundo Reichheld (2003), os promotores são considerados clientes fiéis que provavelmente continuaram usando o serviço.

Outras metodologias se propõem a avaliar serviços de informação como as propostas por Rocha e Souza (2010):

[...] uma metodologia para avaliar produtos e serviços de informação, na qual a triangulação de métodos qualitativos e quantitativos permite, ao mesmo tempo, obter uma visão contextualizada das práticas informacionais dos usuários e enriquecer as condições de uso dos produtos e serviços oferecidos, além de conhecer o grau de satisfação dos usuários e a extensão de acesso e uso dos produtos avaliados, bem como obter informações sobre a facilidade com que os usuários se relacionam com as interfaces desses produtos e serviços.

Na presente pesquisa entende-se serviço de informação sob a perspectiva da UNISIST (1978), organização que estuda aspectos relativos à informação disponibilizada – ao buscar definir os serviços de informação. Também o faz através de sua função principal que, segundo a organização, é a de servir de enlace entre uma população particular de usuários e o universo dos recursos de informação em forma impressa ou não impressa (ROZADOS, 2006).

O mesmo documento faz uma alusão ao papel passivo e ao papel ativo destes serviços. Num papel relativamente passivo, a função de um serviço de informação é assegurar que qualquer informação requerida por um membro da comunidade de usuários esteja à sua disposição, dentro do possível, no momento em que for solicitada. Em seu papel mais ativo, a função de um serviço de informação é alertar a comunidade de usuários sobre os documentos ou dados que possam ser de seu interesse. Para a organização, alguns serviços de informação incluem também as atividades de ‘análise da informação’ (ROZADOS, 2006).

No contexto dos serviços de informação relacionados aos DAGs, tem-se que o direito de acesso à informação pública no Brasil foi previsto na Constituição Federal no inciso XXXIII do Capítulo I - dos Direitos e Deveres Individuais e Coletivos – que dispõe que: “todos têm direito a receber dos órgãos públicos informações de seu interesse particular, ou de interesse coletivo ou geral, que serão prestadas no prazo da Lei, sob pena de responsabilidade, ressalvadas aquelas cujo sigilo seja imprescindível à segurança da sociedade e do Estado” (BRASIL, 1988) .

A Constituição também tratou do acesso à informação pública no Art. 5º, inciso XIV, Art. 37, § 3º, inciso II e no Art. 216, § 2º. São estes os dispositivos que a Lei de

Acesso à Informações regulamenta, estabelecendo requisitos mínimos para a divulgação de informações públicas e procedimentos para facilitar e agilizar o seu acesso por qualquer pessoa através do Serviço de Informação ao Cidadão (INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA, 2018).

Segundo a Lei de Acesso à Informações, os órgãos e entidades públicas deverão utilizar de todos os meios e instrumentos legítimos de que dispuserem, sendo obrigatória a divulgação em sítios oficiais na Internet e devem seguir requisitos tais como:

- I - conter ferramenta de pesquisa de conteúdo que permita o acesso à informação de forma objetiva, transparente, clara e em linguagem de fácil compreensão;
 - II - possibilitar a gravação de relatórios em diversos formatos eletrônicos, inclusive abertos e não proprietários, tais como planilhas e textos, de modo a facilitar a análise das informações;
 - III - possibilitar o acesso automatizado por sistemas externos em formatos abertos, estruturados e legíveis por máquina;
 - IV - divulgar em detalhes os formatos utilizados para estruturação da informação;
 - V - garantir a autenticidade e a integridade das informações disponíveis para acesso;
 - VI - manter atualizadas as informações disponíveis para acesso;
 - VII - indicar local e instruções que permitam ao interessado comunicar-se, por via eletrônica ou telefônica, com o órgão ou entidade detentora do sítio [...].
- (BRASIL, 2011).

Esse último citado possibilita ao cidadão realizar o *feedback* sobre os dados, mas a Lei de Acesso às Informações não oferece detalhes de como isso pode ser feito. Diante o exposto, a etapa feedback no presente modelo pretende cobrir a lacuna encontrada dentre os modelos pesquisados. No entanto, não se intenciona determinar um uso específico de ferramentas ou processo fixo, tendo como premissa a flexibilização da escolha da tecnologia ou ferramenta de acordo com a necessidade do órgão público.

Na próxima seção são apresentados os sujeitos da pesquisa, contexto e coleta de dados usados para validação do modelo.

3.2 SUJEITOS DA PESQUISA, CONTEXTO E COLETA DE DADOS

A Agenda Global 2030 é um compromisso assumido por líderes de 193 países, incluindo o Brasil, e coordenado pelas Nações Unidas, por meio do PNUD, nos termos da Resolução A/RES/72/279.OP32, de 2018, da Assembleia Geral da ONU. No ano

de 2012, aconteceu o Rio+20, um dos eventos mais importantes que discutiu questões importantes relacionadas à sustentabilidade. Durante a conferência, surgiram as primeiras discussões sobre os ODS e o que substituiria os ODM já em 2015, tornando-se os novos ODS o centro da Agenda Global 2030 (BRASIL, 2018).

Considera-se a Agenda Global 2030 um plano de ação para pessoas, planeta e prosperidade (BRASIL, 2018). Busca-se promover o desenvolvimento dentro dos limites de recursos do planeta. Pretende-se erradicar a pobreza por meio de uma parceria entre diversas nações, orientada a encontrar as melhores escolhas para proporcionar a melhoria da vida das pessoas no presente e futuro (BRASIL, 2018).

A Agenda Global 2030 representa uma declaração com 17 ODS, conforme mostra a Figura 8. Além disso, a Agenda Global 2030 contém 169 metas que, em conjunto com os objetivos, revelam a escala para mensurar e acompanhar a agenda. Todos os ODS estão integrados e não é possível acompanhá-los separadamente.

Figura 81 - Objetivos de desenvolvimento sustentável da ONU



Fonte: Secretaria de Governo da Presidência da República (2017, p. 28).

A partir de 2016, os ODS começaram a vigorar e se definiu uma agenda com ações no âmbito mundial. As áreas contempladas são: erradicação da pobreza; segurança alimentar; agricultura sustentável; saúde; educação; igualdade de gênero; redução de desigualdades; energia; água e saneamento; padrões sustentáveis de produção e consumo; mudança de clima; cidades sustentáveis; proteção e uso sustentável dos oceanos e ecossistemas e uso sustentável dos oceanos e ecossistemas terrestres; crescimento econômico inclusivo; infraestrutura e industrialização (BRASIL, 2018).

Pode-se agrupar os ODS em quatro dimensões principais:

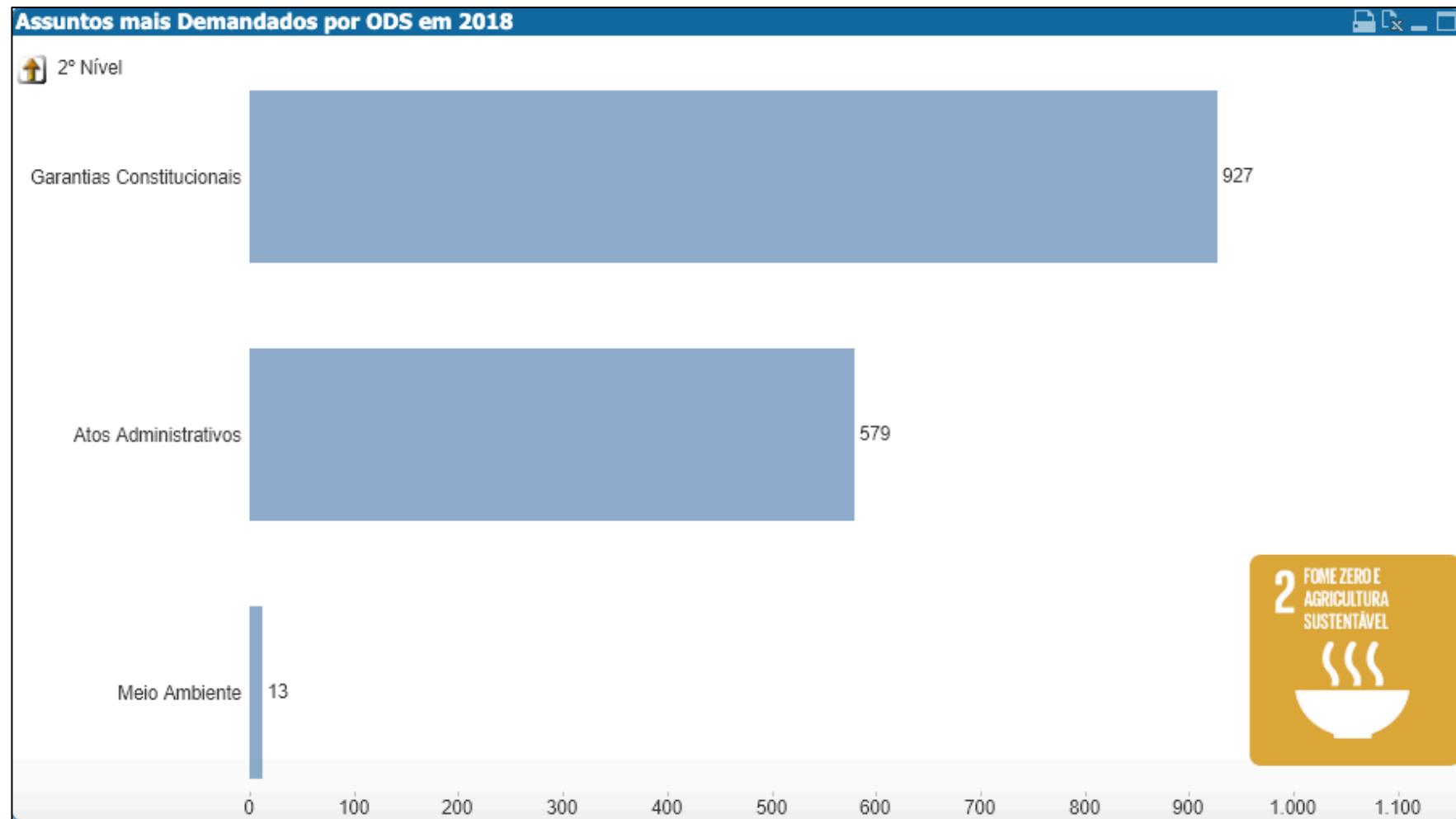
[...] **ambiental**: abrange a preservação e conservação do meio ambiente, com ações que vão da reversão do desmatamento, proteção das florestas e da biodiversidade, combate à desertificação, uso sustentável dos oceanos e recursos marinhos até a adoção de medidas efetivas contra mudanças climáticas; **social**: relacionada às necessidades humanas, de saúde, educação, melhoria da qualidade de vida e justiça; **econômica**: aborda o uso e o esgotamento dos recursos naturais, a produção de resíduos, o consumo de energia, entre outros; e **institucional**: diz respeito à capacidade de colocar em prática os ODS. (BRASIL, 2017, grifo nosso).

Por meio da Portaria CNJ nº 133/2018, foi instituído, no CNJ, um Comitê Interinstitucional destinado a avaliar a integração das metas do Poder Judiciário às metas e aos indicadores dos ODS, Agenda Global 2030, e elaborar relatório de trabalho com apoio de todos os tribunais do país, cuja composição consta da Portaria CNJ nº 148/2018 (CONSELHO NACIONAL DE JUSTIÇA, 2018).

Segundo a Portaria CNJ nº 133/2018, o alinhamento da atuação do Poder Judiciário à Agenda Global 2030, da ONU, pode representar avanço no campo na concretização dos direitos fundamentais dos cidadãos, pois o judiciário poderá fornecer informações relevantes e necessárias – cuja base de dados é produzida e mantida pelo próprio Poder Judiciário – para o cumprimento das metas dos ODS, que vincula todas as pessoas, de todos os países, por se traduzir em um Pacto Global para erradicar a pobreza e promover vida digna para todos, dentro dos limites do planeta (CONSELHO NACIONAL DE JUSTIÇA, 2018).

A Figura 9 mostra um exemplo de um painel de *Business Intelligence* (BI) com os assuntos mais demandados pelo judiciário relacionados ao ODS 2 – Fome Zero e Agricultura Sustentável, no ano de 2018.

Figura 92 - Assuntos mais demandados por ODS em 2018



Para mostrar que é possível essa correlação do Poder Judiciário aos ODS, o Comitê mapeou todas as Resoluções do CNJ, bem como as recomendações, os provimentos e as orientações da Corregedoria Nacional de Justiça, além dos Atos Normativos e Boas Práticas dos Tribunais Superiores, Tribunais Regionais Federais, Seções Judiciárias, Tribunais de Justiça dos Estados e do Distrito Federal, Tribunais Regionais do Trabalho, Tribunais Regionais Eleitorais, Tribunais Militares, sendo este último conjunto de dados utilizado na presente pesquisa para validação do modelo baseado em *Frictionless Data*.

Por meio da Portaria nº 19/2019, foi criado também o Laboratório de Inovação, Inteligência e Objetivos de Desenvolvimento Sustentável (LIODS), um programa que une o conhecimento institucional, a inovação e a cooperação. É um espaço administrativo e horizontal de diálogo e articulação de políticas públicas entre o Poder Judiciário, os entes federativos e a sociedade civil, com o objetivo de se alcançar a paz, a justiça e a eficiência institucional.

O programa é uma forma de institucionalizar o uso da inovação e da inteligência, promovendo o desenvolvimento de alianças estratégicas e projetos de cooperação, considerando o envolvimento de especialistas de todos os setores com o objetivo de trabalhar a Agenda Global 2030 no âmbito do Poder Judiciário, contribuindo, assim, com o Plano Estratégico do Poder Judiciário e com a melhoria das políticas públicas. O LIODS também tem o desafio de unir os laboratórios de inovação e os centros de inteligência já existentes no judiciário, criando uma rede de colaboração entre seus agentes. Estes espaços institucionais oportunizam a participação, favorecem o diálogo e a análise conjunta de demandas repetitivas ou com grande repercussão social, analisando dados, inclusive o custo econômico, e permitindo a construção de soluções estratégicas e a ampliação da transparência na atuação do Poder Judiciário, garantindo o acesso à Justiça por meio de instituições mais eficazes, responsáveis e transparentes, alinhadas com o desenvolvimento sustentável (CONSELHO NACIONAL DA JUSTIÇA, 2020).

Com a estruturação do LIODS, a ideia é, além de criar a rede de colaboração entre os laboratórios e centros de inteligência do Poder Judiciário, ampliar a rede para conexão com os laboratórios e centros de inteligência de todo o setor público, facilitando a troca de informações, a integração de políticas e a solução de problemas complexos, com foco na Agenda Global 2030 (CONSELHO NACIONAL DA JUSTIÇA, 2020). No sítio do CNJ se encontra um painel de BI no qual estão indexadas as

normativas e boas práticas publicadas no âmbito do Poder Judiciário com os ODS da Agenda Global 2030. A Figura 10 mostra a base de dados com as boas práticas dos tribunais que contém informações atualizadas e enviadas periodicamente para o CNJ.

Figura 103 - Base de dados com boas práticas dos tribunais relacionados aos ODS

A	B	C	D	E	F	G	H	I	J	K
1 ODS Relacionado	Unidade de Implantacao	Titulo	Implantacao	Autores	Email Autores	Tema	Objetivo	Finalidade	Roteiro de Implantacao	Resultado Alcançado
2 ODS1 - Erradicação da Pobreza	Área do Antigo Lixão do Aurá - Escola	Acreditar no Amanhã	01/06/2017 DESEMBARGADOR LEONARDO LEONARDO.tavares@tjpe.jus.br Gestão Social			Contribuir para a melhoria da c 1- Desenvolver atividades de 1- Realizar oficina de escavação2- Cor	Foram realizados atendimento			
3 ODS1 - Erradicação da Pobreza	Campo Grande	Replicação do Programa Cozinha e Voz	Paola Carosella	far@fli.org		Empregabilidade de mulheres Oferecer curso de formação e Inserção no mercado de trabalho	Pedido de autorização aos idealizadores 20 mulheres participaram do			
4 ODS1 - Erradicação da Pobreza	Locais mais carentes	Estradeiro Natalino	Justiça Comunitária	justica.comunitaria@tjmt.jus.br		Realizar, anualmente, doação	Erradicação da pobreza.			
5 ODS1 - Erradicação da Pobreza	Núcleo de Responsabilidade Socioambiental	COLETA SELETIVA DESTIN	02/07/1905 SILVANA MARIA TEIXEIRA	Comissão Permanente de Inclusão dos Catadores e		Promover Políticas Afirmativas	Promover a Inclusão dos Catadores			
6 ODS1 - Erradicação da Pobreza	Poder Judiciário do Estado	Programa Ribeirinho Cidadão	Justiça Comunitária	justica.comunitaria@tjmt.jus.br		Levar atendimento à pessoas	ca			
7 ODS1 - Erradicação da Pobreza	Socioambiental	Projeto Valente	26/06/1905 Socioambiental	socioambiental@tjrt11.jus.br		Inclusão social de catadores	Proporcionar trabalho aos assos	Acabar com a pobreza		Criação de programa de trabalho sustentável com a pobreza
8 ODS1 - Erradicação da Pobreza e ODS4	3ª Vara da Infância e da Juventude	Clube da Leitura	02/05/2018 Laercio Lopes Pinto e Ana Maercio.pinto@tjpa.jus.br	gestão social		Posibilitar acesso à incentivo	1- Incentivar o hábito e o go 1,2 e 3 etapa encontro de leitura.4- C	O projeto atingiu os objetivos		
9 ODS1 - Erradicação da Pobreza e ODS16 Município de Belém	Minha escola, Meu refúgio	Mônica Maciel Soares Fonseca	monica.fonseca@tjpa.jus.br	Gestão Social		Promover a proteção integral à	Promover reunião com escolas 1- Fase do levantamento e contatos	Foram realizadas 32 vistas e		
10 ODS2 - Fome Zero e Agricultura Sustentável	Comissão da Cesta Básica	Cesta Básica	Ángela de Almeida Rezende	angela.rezende@tre-mg.jus.br		Recolhimento de valor monetário	Contribuir para melhoria da	Recolhimento mensal dos valores mor Nos primeiros 13 anos do pro		
11 ODS2 - Fome Zero e Agricultura Sustentável	Coordenadoria de Planejamento e Gestão	Feira de produtos agroecológicos	09/01/2018 Yoanna Fernandes Guerra	yoanna@tjrt21.jus.br		A feira de produtos agroecológicos	Contribuir com a saúde do c	Identificação e seleção de instituições	Consumo de frutas, verduras	
12 ODS2 - Fome Zero e Agricultura Sustentável	Núcleo de Responsabilidade Socioambiental	CAMPANHA DE DOAÇÃO DE LARANJAS	10/07/1905 SILVANA MARIA TEIXEIRA	Comissão Permanente de Doação de Laranja ao IPRED		Promover Campanha para Arre	Promover o Voluntariado e possibilizar a adequada alimentação de crianças carentes, mediante o			
13 ODS2 - Fome Zero e Agricultura Sustentável	NUCLEO SOCIOAMBIENTAL	FEIRA ORGÂNICA		socioambiental@tjrt5.jus.br		auxiliar alimentação saudável				
14 ODS2 - Fome Zero e Agricultura Sustentável	Região Metropolitana de Belém	Projeto Feira Orgânica	02/05/2017 Evelise de Oliveira Rodrigues	evelise.rodrigues@tjpa.jus.br		Responsabilidade Socioam	Construir uma democracia susi 1- Melhorar a qualidade de vi	1- Formalização do processo2-Divulga	O projeto já realizou 43 feiras	
15 ODS3 - Saúde e Bem-Estar	Assessoria de Comunicação Social	Campanha de Doação de S	10/07/1905			Ação Social para doação	1- Promover ação social na institu	Campanha desenvolvida pela	Produção de peças publicitárias, confe	Resultado de grande repercus
16 ODS3 - Saúde e Bem-Estar	Assessoria de Políticas de Saúde/SGP	Convênio/Contrato	19/02/2018 Viviane Alves Gomes de Pa	viviane.paula@tre-ms.jus.br		Ginástica Laboral	Benefício fisiológico, psicológico Atua de forma preventiva e	Procedimento licitatório	Restultados satisfatórios na	
17 ODS3 - Saúde e Bem-Estar	Assessoria de Políticas de Saúde/SGP	Convênio/Contrato	19/04/2018 Viviane Alves Gomes de Pa	viviane.paula@tre-ms.jus.br		Vacina contra Influenza (Evitar a falta ao trabalho por	Promoção da saúde de serv	Procedimento licitatório	Resultados muito satisfatório
18 ODS3 - Saúde e Bem-Estar	Assessoria de Políticas de Saúde/SGP	Convênio/Contrato	01/08/2018 Viviane Alves Gomes de Pa	viviane.paula@tre-ms.jus.br		Tratar preventivamente a saúde	Promoção de assistência de	Procedimento licitatório		Resultados muito satisfatório
19 ODS3 - Saúde e Bem-Estar	CODSA - Coordenadoria de Saúde	TRT8SAÚDE	CODSA	cods@tjrt8.jus.br	TRT8SAÚDE	Acompanhamento e prevenção	Acompanhamento e prevenção			
20 ODS3 - Saúde e Bem-Estar	COORDENADORIA DE SAÚDE	CAMPANHA DE VACINAÇÃO (*GRIPE)	COORDENADORIA DE SAÚDE	saude@tjrt5.jus.br			diminuição do afastamento por gripe			
21 ODS3 - Saúde e Bem-Estar	COORDENADORIA DE SAÚDE	EXAME BUCAL PERIÓDICO	COORDENADORIA DE SAÚDE	saude@tjrt5.jus.br						
22 ODS3 - Saúde e Bem-Estar	COORDENADORIA DE SAÚDE	EXAME PERIÓDICO	COORDENADORIA DE SAÚDE	saude@tjrt5.jus.br						
23 ODS3 - Saúde e Bem-Estar	Coordenadoria de Serviços Gerais/SAF	Convênio/Contrato	09/04/2019 Luís Maciel Málvez da Silva	luis.silva@tre-ms.jus.br		l; Desratização, desinsetização, Eliminação de vetores de doen	Mantener a salubridade dos loc	Procedimento licitatório		Eradicacão realizada em sua
24 ODS3 - Saúde e Bem-Estar	DG/SGP	Comitê de Acompanhamento do Clima Organizacional	MARIA DA GLÓRIA ARAÚJO	gloria.araujo@tjmg.jus.br		Clima Organizacional	Aggerar à Alta Administração	o Melhor na gestão das ações	Publicação da Portaria nº 17/2017	Discussão dos resultados da
25 ODS3 - Saúde e Bem-Estar	Divisão de Saúde - TRT-7ª Região	AVALIAÇÃO ERGONÔMICA	10/07/1905 PATRÍCIA MAIA MOTA FAL	Seção Médico-Odontológico		Cuidados com a Postura e Promover a Saúde e o Bem Est	Promover a Saúde do Corpo Funcional do TRT-7ª Região e trazer	Melhorias na Qualidade de Vida		
26 ODS3 - Saúde e Bem-Estar	Divisão de Saúde - TRT-7ª Região	CAMPANHA JANEIRO BRAS	16/01/2017 MARIANA DE BRITO LIMA	Seção Psicosocial - (85)		Cuidados com a Saúde Me	Promover a Saúde do Corpo Funcional do TRT-7ª Região e trazer	Melhorias na Qualidade de Vida		
27 ODS3 - Saúde e Bem-Estar	Divisão de Saúde - TRT-7ª Região	CAMPANHA NOVEMBRO AZUL	THAIS EVANGELISTA FERNA	Divisão de Saúde - (85)		Cuidados com a Saúde e - (85) 3 Cuidados e Prevenção do	Promover a Saúde e o Bem Est	Promover a Saúde do Corpo Funcional do TRT-7ª Região e trazer	Melhorias na Qualidade de Vida	
28 ODS3 - Saúde e Bem-Estar	Divisão de Saúde - TRT-7ª Região	CAMPANHA OUTUBRO ROSA	THAIS EVANGELISTA FERNA	Divisão de Saúde - (85)		3 Cuidados e Prevenção do	Promover a Saúde e o Bem Est	Promover a Saúde do Corpo Funcional do TRT-7ª Região e trazer	Melhorias na Qualidade de Vida	
29 ODS3 - Saúde e Bem-Estar	Divisão de Saúde - TRT-7ª Região	CAMPANHA SETEMBRO AMARELHO	MARIANA DE BRITO LIMA	Seção Psicosocial - (85)		Cuidados e Prevenção do	Promover a Saúde e o Bem Est	Promover a Saúde do Corpo Funcional do TRT-7ª Região e trazer	Melhorias na Qualidade de Vida	
30 ODS3 - Saúde e Bem-Estar	Divisão de Saúde - TRT-7ª Região	PROJETO DE SUPORTE PSI	10/07/1905 MARIANA DE BRITO LIMA	Seção Psicosocial - (85)		Cuidados com a Saúde em Promover a Saúde e o Bem Est	Promover a Saúde do Corpo Funcional do TRT-7ª Região e trazer	Melhorias na Qualidade de Vida		
31 ODS3 - Saúde e Bem-Estar	Divisão de Saúde - TRT-7ª Região	PROJETO DE SUPORTE PSI	10/07/1905 MARIANA DE BRITO LIMA	Seção Psicosocial - (85)		Cuidados com a Saúde em Promover a Saúde e o Bem Est	Promover a Saúde do Corpo Funcional do TRT-7ª Região e trazer	Melhorias na Qualidade de Vida		
32 ODS3 - Saúde e Bem-Estar	Estado de Mato Grosso do Sul	Empenhado@ pela Saúde				Coordinadoria Estadual da	Criarão de um curso para profi	Promover esclarecimentos	reunião para firmar as parcerias	reúne 457 pessoas inscritas
33 ODS3 - Saúde e Bem-Estar	Foro Central de Porto Alegre - Foro Ce	DOAR É LEGAL	14/08/2009 Vara da Direção do Foro Ce	frpacientecardin@tjrs.jus.br		doação de órgãos.	Promover a conscientização da	Criarão de um curso para profi	Promover esclarecimentos	
34 ODS3 - Saúde e Bem-Estar	Gestão de Pessoas	Programa de prevenção de	02/12/2015 Maria do Socorro Chaves de dir.pessoal@tjrt11.jus.br			Prevenção de doenças	Reduzir absentismo			
35 ODS3 - Saúde e Bem-Estar	Gestão de Pessoas	Programa Institucional de	04/06/2018 Maria do Socorro Chaves de dir.pessoal@tjrt11.jus.br			Assegurar vida saudável e Promover a qualidade de vida / Minimizar e/ou eliminar situaç	Realização de programa de prevenç			
36 ODS3 - Saúde e Bem-Estar	Gestão de Pessoas	Projeto Mais Saúde para Todos	Maria do Socorro Chaves de dir.pessoal@tjrt11.jus.br			Assegurar vida saudável e Promover a qualidade de vida / Desenvolvimento de Capacida				
37 ODS3 - Saúde e Bem-Estar	Gestão de Pessoas	Projeto Resolução de Confl	30/01/2018 Maria do Socorro Chaves de dir.pessoal@tjrt11.jus.br			Assegurar vida saudável e Promover a qualidade de vida / Incentivar a realização de ati				
38 ODS3 - Saúde e Bem-Estar	NGP	Campanha anual de vacinação	Gizelma Rita	gizelma.rita@tjpe.jus.br		Assegurar vida saudável e Promover a qualidade de vida	Reducir a rotatividade de ser	Oferecer suporte para auxiliar pessoa	Promover o convívio harmônio	
39 ODS3 - Saúde e Bem-Estar	NGP	Promoção de ações voltadas à Qualidade de Vida	Núcleo de Gestão de Pessoas	ngp@tjfsc.jus.br		Prevenção de doenças	Reducir absentismo			
40 ODS3 - Saúde e Bem-Estar	Núcleo de Acompanhamento e Desenv.	Enfrentando Desafios	18/10/2018 Daniel Philippo de Negreiros	natdh.dircreao@tjfsc.jus.br		Saúde e Qualidade de vida	Prevenção de doenças ocupaci	Reducir o absentismo		
41 ODS3 - Saúde e Bem-Estar	Núcleo de Acompanhamento e Desenv.	Na Justiça somos todos m	21/11/2018 Leonardo Guillermo Cal	natdh.dircreao@tjfsc.jus.br		Gestão de Pessoas	Oportunizar um espaço de info	Saúde e Ben Estar		
42 ODS3 - Saúde e Bem-Estar	Núcleo de Gestão de Pessoas	Argões de Qualidade de Vida no Trabalho				Gestão de Pessoas	Estimular a Cultura de Prevenç	Saúde e Ben Estar		
43 ODS3 - Saúde e Bem-Estar	Núcleo de Responsabilidade Socioambiental	CAMPANHA "VOU DE ESCA	10/07/1905 SILVANA MARIA TEIXEIRA	Comissão Permanente de		Saúde e Qualidade de vida	Promoção de saúde no ambiente	Reduzir o absentismo		
44 ODS3 - Saúde e Bem-Estar	Núcleo de Responsabilidade Socioambiental	RODA DE COMUNISMA SOBRE	11/07/1905 THAIS EVANGELISTA FERNA	thais.evangelista@tjrt7.jus.br		Cuidados com a Saúde	Promover a Saúde e o Bem Est	Reduzir a rotatividade de		

Fonte: Conselho Nacional de Justiça (2020).

A base de dados se encontra em formato proprietário, possui 436 registros e contém campos do tipo texto, tais como:

- a) ODS relacionado: objetivo de desenvolvimento sustentável relacionado à boa prática;
- b) unidade de implantação: local de realização da boa prática;
- c) título: nome da boa prática;
- d) implantação: data da implantação da boa prática;
- e) autores: envolvidos na implantação da boa prática;
- f) e-mail autores: endereço eletrônico dos envolvidos na implantação da boa prática;
- g) tema: classificação da boa prática de acordo com temática pré-definida;
- h) objetivo: objetivos a alcançar com a boa prática;
- i) finalidade: ações para alcançar os objetivos da boa prática;
- j) roteiro: passo a passo para implantação da boa prática;
- k) resultados alcançados: resultados alcançados com a implantação da boa prática.

Nesta seção foram apresentados o contexto, os sujeitos da pesquisa, além da exposição da base de dados usada na pesquisa. Na próxima seção serão abordadas as ferramentas disponíveis no sítio da iniciativa *Frictionless Data*, as quais foram usadas para validar a aplicação do modelo proposto.

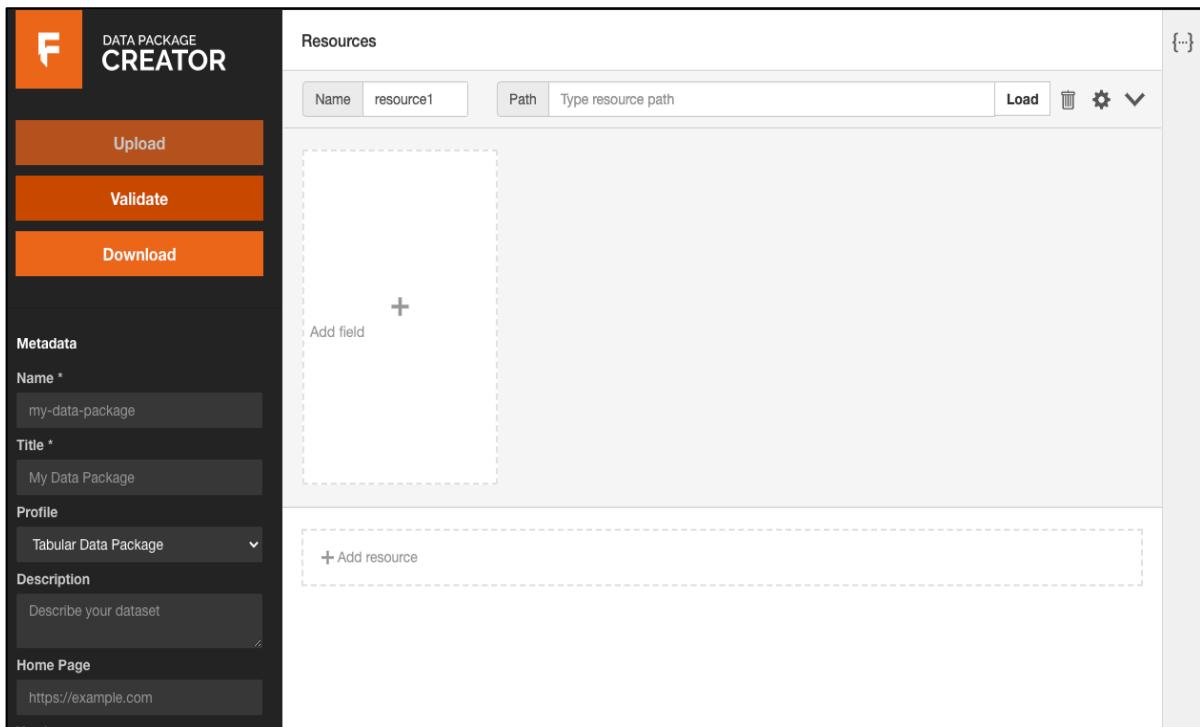
3.3 FERRAMENTAS

De acordo com Fowler, Barratt e Walsh (2017), existe atrito significativo na aquisição, compartilhamento e reutilização de dados de pesquisa, por exemplo. Entende-se por atrito o tempo e esforço para trabalhar com os dados. Estima-se que oitenta por cento da análise de dados sejam investidos na limpeza e mapeamento de dados. O *Frictionless Data* é um projeto em andamento na *Open Knowledge International* focado na remoção desse atrito. A iniciativa disponibiliza um conjunto de ferramentas, especificações e práticas recomendadas para descrever, publicar e validar dados. O coração deste projeto é o *Data Package*, um formato de contêiner

para dados com base nas práticas existentes para publicação de *software* de código aberto (FOWLER; BARRATT; WALSH, 2017).

Ao compartilhar vários conjuntos de dados sobre um assunto específico com um público variado, é importante garantir que quem acessa os dados entenda o seu contexto e possa acessar rapidamente informações sobre licenças e outras atribuições. *Data Package* é um formato que possibilita colocar dados e informações relevantes que fornecem contexto sobre dados em um contêiner antes de compartilhá-los. Todas as informações contextuais, como metadados e esquema de dados são publicadas em um único arquivo JSON chamado `datapackage.json`. O *Data Package Creator* é um serviço on-line que facilita a criação e edição de um *data package*. O serviço gera automaticamente um arquivo `datapackage.json`, à medida que é possível adicionar e editar dados que fazem parte da base de dados (FRICTIONLESS DATA, 2018). A Figura 114 mostra a ferramenta *Data Package Creator*, a qual será detalhada na seção de aplicação do modelo.

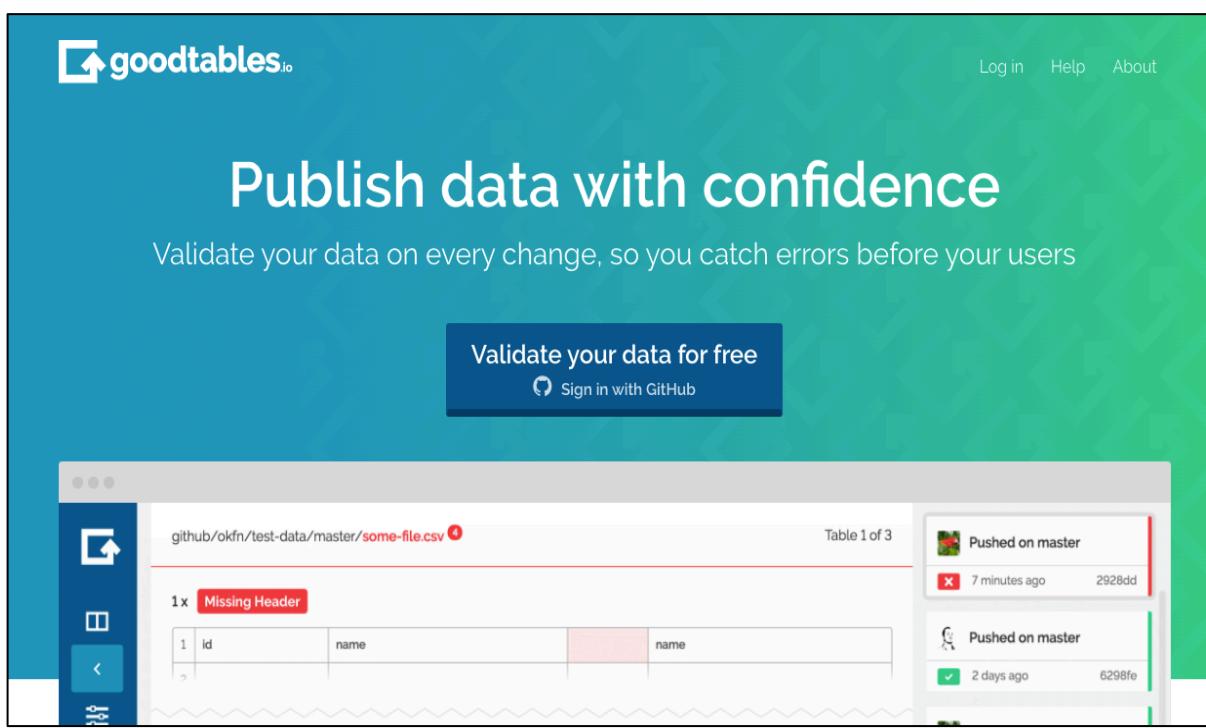
Figura 114 - Ferramenta *Data Package Creator*



Fonte: Data Package Creator (2020).

A presença de erros nos seus dados é muito comum e costumam atrapalhar a análise de dados para muitos usuários de dados. A validação de dados ajuda a facilitar o processo de coleta de dados, verificando a qualidade e a validade de uma base de dados antes de publicá-la. Dentro da iniciativa *Frictionless Data* existe uma ferramenta para validação dos dados transformados em *Data Package* chamada *Goodtables*, que se trata de um conjunto de bibliotecas e ferramentas para validar e transformar dados tabulares. Essas bibliotecas existem para identificar erros estruturais e de conteúdo em seus dados tabulares, para que possam ser corrigidos rapidamente. Por exemplo, um esquema de tabela contém informações sobre os campos e seus tipos de dados atribuídos, possibilitando destacar tipos de dados extraviados (por exemplo, uma sequência de caracteres em uma coluna de idade em que um número inteiro é esperado, ou um número inteiro em uma coluna de e-mail em que uma sequência de caracteres é esperada) (FRICTIONLESS DATA, 2018). Essa ferramenta foi selecionada para validação da etapa tratar do modelo proposto. A Figura 12 mostra a página principal do *Goodtables*.

Figura 125 - Página inicial da ferramenta *Goodtables*

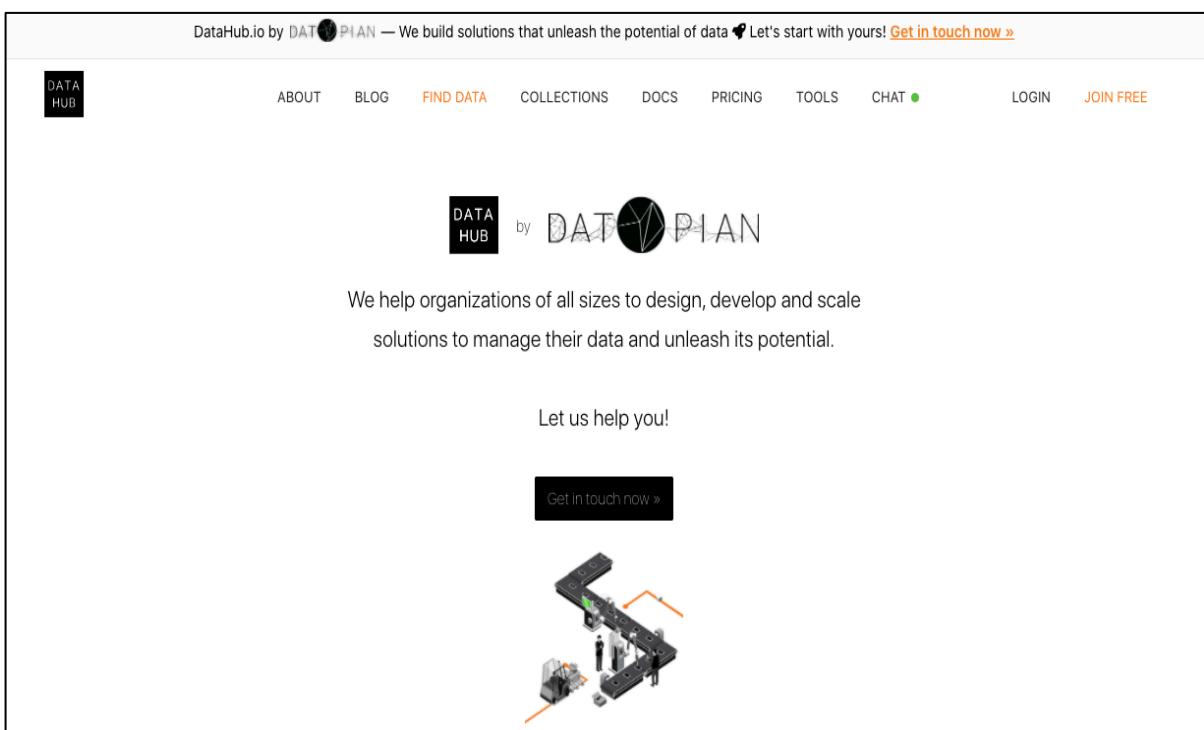


Fonte: Goodtables (2020).

Umas das ferramentas usadas para publicação dos dados disponível no sítio da iniciativa *Frictionless Data* é o *DataHub*, uma plataforma para encontrar, compartilhar e publicar dados de alta qualidade. O *DataHub* começou como um projeto da *Datopian* e da *Open Knowledge International*, sendo essa a mesma criadora do CKAN, a principal plataforma de portal de dados do mundo que fornece suporte a sites como data.gov e data.gov.uk. A proposta do *DataHub* é ser uma ferramenta para transformar criar e usar dados de qualidade, trazendo melhorias drásticas em facilidade, velocidade e confiabilidade. No entanto, a ferramenta é paga e possui um plano de publicação sem custos com capacidade de armazenamento de até 1GB (DATAHUB, 2018).

Logo, para fins da presente pesquisa, optou-se por publicar os dados no *DataHub*, porém futuramente se pretende aplicar o modelo usando o *Data Factory*, que se trata de um *framework* e um kit de ferramentas *open source* para criar fluxos de dados os quais abrangem as etapas de coletar, inspecionar, processar e publicar dados. É possível automatizar totalmente a coleta, processamento e publicação de dados com o *Data Factory* (DATAHUB, 2018). A Figura 13 mostra a página inicial da ferramenta *DataHub*.

Figura 136 - Página inicial da ferramenta *DataHub*



Fonte: DataHub (2020).

Nesta seção foram apresentadas as ferramentas usadas na presente pesquisa e na próxima seção são detalhados os procedimentos usados na aplicação do modelo baseado em *Frictionless Data*. Com o objetivo de validar o modelo proposto foram aplicadas as ferramentas disponíveis no sítio da iniciativa *Frictionless Data* à uma base de dados de boas práticas do CNJ, relacionadas aos ODS.

3.3.1 Aplicação do modelo e resultados

Nessa seção são expostos detalhes da aplicação do modelo proposto com objetivo de validá-lo por meio do uso das ferramentas disponíveis na iniciativa *Frictionless Data* e os seus correspondentes resultados. Pensando no leitor com pouca familiaridade com tecnologia, a seção aborda os passos principais para facilitar a reprodução da pesquisa, com definições essenciais ao entendimento da aplicação e foco nos dados das Boas Práticas do Judiciário, relacionadas aos ODS.

Primeiramente, definiu-se a base de dados que se deseja para realizar a abertura de dados. No caso da presente pesquisa, a base de dados com as boas práticas escolhida foi a do CNJ relacionadas aos ODS. A base se encontra em formato proprietário, sendo um conjunto de dados adequado para validar o modelo. Destaca-se que o modelo proposto pode ser adaptado para qualquer base de dados, inclusive em diferentes formatos, uma vez que o *Data Package*, fundamento da etapa empacotar, apresentada na seção que descreve o modelo, possui um formato de contêiner simples para descrever uma coleção coerente de dados em um único pacote, fornecendo a base para entrega, instalação e gerenciamento conveniente de um conjunto de dados (FRICTIONLESS DATA SPECS, 2017).

Um *Data Package* consiste em metadados, que descrevem a estrutura e o conteúdo do pacote, e recursos, como arquivos de dados que formam o conteúdo do pacote. Os metadados do *Data Package* são armazenados em um ‘descritor’. Este descritor é o que torna uma coleção de dados um *Data Package*. A estrutura desse descritor é o conteúdo principal da especificação. Além desse descritor, um *Data Package* incluirá outros recursos, como arquivos de dados. A especificação do *Data Package* NÃO impõe quaisquer requisitos em sua forma ou estrutura e, portanto, pode ser usada para empacotar qualquer tipo de dados. Os dados incluídos no pacote podem ser fornecidos como: arquivos agrupados localmente com o descritor de pacote; recursos remotos, referenciados por URL e dados ‘embutidos’ que estão

incluídos diretamente na descrição (FRICTIONLESS DATA SPECS, 2017). Todas as informações contextuais, como metadados e seu esquema de dados, são publicadas em um arquivo JSON denominado `datapackage.json`.

3.3.1.1 Verificação da etapa empacotar

Para validar a etapa empacotar do presente modelo foi usada a ferramenta *Data Package Creator*, que facilita a criação e edição de um *Data Package*. A ferramenta gera automaticamente um arquivo `datapackage.json` que contém o conjunto de dados que se deseja empacotar. Refere-se a cada parte dos dados de um conjunto de dados como um recurso de dados.

Existem três maneiras de adicionar base de dados no *Data Package Creator*: (1) através de um *hiperlink* para um repositório que contém a base de dados, como o *GitHub*; (2) por meio da criação da base de dados dentro do *Data Package Creator*, porém essa opção é mais complexa dependendo da estrutura de dados; ou (3) carregar um arquivo `datapackage.json` pré-existente para visualizar e editar os metadados e campos da base de dados (FRICTIONLESS DATA, 2018).

Para a presente pesquisa foi escolhida a opção de realizar o *upload* e disponibilizar a base de dados, conforme mostra a Figura 14.

Figura 147 - Ferramenta Data Package Creator

DATA PACKAGE CREATOR

Upload

Validate

Download

Metadata

Name *
cnj_ods

Title *
Boas práticas CNJ ODS

Profile
Data Package

Description
Boas práticas do judiciário relacionados aos ODS

Home Page
<https://example.com>

Version

Data package is valid!

Resources

Name	res
Path	cnj_ods_csv.csv
Load	
More	
ODS Relacionado	Unidade de Implan
1 ODS1 - Erradicação	1 Área do Antigo Lixão
Title	Title
Description	Description
Data Type	Data Type
string	string
Data Format	Data Format
default	default
Implantacao	Autores
1 01/06/2017	1 DESEMBARGADOR
Title	Title
Description	Description
Email Autores	
1 des.leonardo.tavares@tjpa.j	
Title	
Description	

Preview

```
{
  "profile": "data-package",
  "resources": [
    {
      "name": "resource1",
      "path": "cnj_ods_csv.csv",
      "profile": "data-resource",
      "schema": {
        "fields": [
          {
            "name": "ODS Relacionado",
            "type": "string",
            "format": "default"
          },
          {
            "name": "Unidade de Implan",
            "type": "string",
            "format": "default"
          },
          {
            "name": "Titulo",
            "type": "string",
            "format": "default"
          },
          {
            "name": "Implantacao",
            "type": "any",
            "format": "default"
          },
          {
            "name": "Autores",
            "type": "string",
            "format": "default"
          },
          {
            "name": "Email Autores",
            "type": "string",
            "format": "default"
          }
        ]
      }
    }
  ]
}
```

Fonte: Elaborado pela autora (2020).

A ferramenta permite incluir metadados para um conjunto de dados através da barra lateral esquerda, onde é possível incluir informações como o nome do conjunto de dados, descrição, autor, tipo de licença do dado, entre outras informações. Segundo a especificação *Frictionless Data*, algumas propriedades são recomendadas para um metadado e devem ser incluídas em cada descritor do pacote (FRICTIONLESS DATA, 2018):

- a) nome: um nome curto que pode ser usado por URL; deve ser minúsculo e conter apenas caracteres alfanuméricos junto com caracteres (‘.’, ‘_’ ou ‘-’); esse nome funcionará como um identificador único, de preferência globalmente único; o nome também deve ser invariável, ou seja, não deve mudar quando um pacote de dados é atualizado, ao menos que uma nova versão do pacote resulte em um pacote com mudanças significativas na estrutura ou interpretação;
- b) id: um identificador de uso exclusivo para um *Data Package*; exemplos de identificador exclusivo inclui UUIDs e DOIs;
- c) licença: licença sob as quais um *Data Package* é fornecido; essa propriedade não é legalmente vinculativa e não garante que o *Data Package* seja licenciado de acordo com os termos definidos nesta propriedade; a função é informativa;
- d) descrição: uma descrição do pacote; a descrição deve ser formatada com marcação, isso também permite texto simples e deve conter uma informação resumida do pacote;
- e) URL: uma URL para a página principal na web relacionada ao *Data Package*;
- f) versão: identifica a versão do pacote; deve estar em conformidade com os requisitos *Semantic Versioning1* e *Data Package Version2*;
- g) *profile* ou perfil: permite especificar o tipo de coleção de dados que está sendo empacotado; na presente pesquisa, usou-se o perfil *Data Package* por se tratar de uma planilha Excel; são três os tipos de *profile*: *Data Package*: perfil básico, usado para quando o conjunto de dados contiver recursos de formatos diversos, como dados tabulares ou geográficos; *tabular Data Package*: quando os dados contiverem apenas recursos tabulares, como CSVs e planilhas em formato aberto; *fiscal Data Package*:

se os dados contiverem informações fiscais, como dados de orçamentos e despesas.

Ainda é possível incluir palavras-chave relacionadas ao *Data Package*. No presente trabalho se optou por preencher esse campo com possíveis palavras-chave, tais como ‘ODS’ e ‘boas práticas’, o que se inferiu serem palavras adequadas para tornar o *Data Package* mais detectável nas buscas. Após a criação do *Data Package*, é possível validá-lo ou fazer o *download* do datapackage.json. A cada alteração na base de dados, o *Data Package Creator* preenche o arquivo datapackage.json com as novas informações. Na própria ferramenta se pode visualizar o conteúdo do novo datapackage.json. Após a criação do *Data Package*, é possível pré-visualizar o conteúdo do datapackage.json, além do conteúdo do arquivo datapackage.json (Figura 15).

Figura 158 - Preview do arquivo datapackage.json



```

▶ Preview
{
  "profile": "data-package",
  "resources": [
    {
      "name": "resource1",
      "path": "cnj_ods_csv.csv",
      "profile": "data-resource",
      "schema": {
        "fields": [
          {
            "name": "ODS Relacionado",
            "type": "string",
            "format": "default"
          },
          {
            "name": "Unidade de Implantação",
            "type": "string",
            "format": "default"
          },
          {
            "name": "Título",
            "type": "string",
            "format": "default"
          },
          {
            "name": "Implantação",
            "type": "any",
            "format": "default"
          },
          {
            "name": "Autores",
            "type": "string",
            "format": "default"
          },
          {
            "name": "Email Autores",
            "type": "string"
          }
        ]
      }
    }
  ]
}

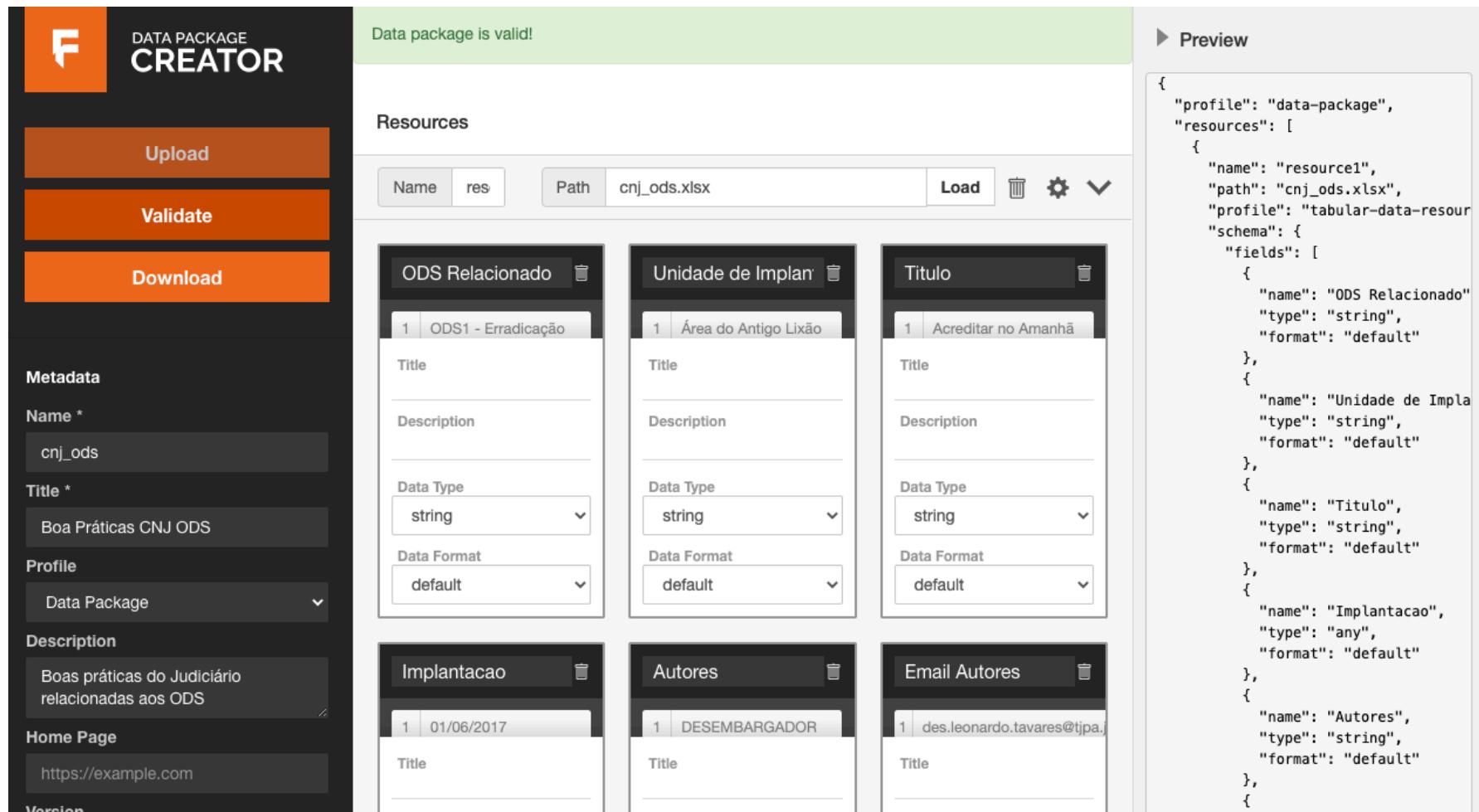
```

Fonte: Elaborado pela autora (2020).

A opção validar permite confirmar se o profile escolhido está adequado ao conjunto de dados. Os resultados possíveis para a validação são: ‘*Data Package is invalid*’, o que significa que há algum erro de validação, como, por exemplo, o formato do nome do arquivo ou a escolha de um perfil incorreto. Caso a mensagem informar ‘*Data Package is valid*’, então não há erros no processo de empacotamento, sendo possível realizar o *download* do arquivo.

A partir do processo de validação do datapackage.json, pode-se visualizar a estrutura dos dados, como mostra a Figura 16, com o resultado do processo de empacotamento usando a ferramenta *Data Package Creator*.

Figura 169 - Resultado da criação do arquivo datapackage.json



The screenshot shows the Data Package Creator interface. On the left, there are buttons for Upload, Validate, and Download. Below these are sections for Metadata, Profile, Description, Home Page, and Version. The Metadata section includes fields for Name (cnj_ods), Title (Boa Práticas CNJ ODS), Profile (Data Package), and Description (Boas práticas do Judiciário relacionadas aos ODS). The Home Page is set to <https://example.com>. The Version field is partially visible.

The main area shows a green header bar stating "Data package is valid!". Below this is a "Resources" section with a table. The table has columns for Name, res, Path, and actions. A row is shown with Name: "res", Path: "cnj_ods.xlsx", and actions including Load, Delete, and Settings. The table contains six rows, each representing a resource:

- ODS Relacionado: Title, Description, Data Type (string), Data Format (default)
- Unidade de Implan: Title, Description, Data Type (string), Data Format (default)
- Titulo: Title, Description, Data Type (string), Data Format (default)
- Implantacao: Title, Description, Data Type (any), Data Format (default)
- Autores: Title, Description, Data Type (string), Data Format (default)
- Email Autores: Title, Description, Data Type (string), Data Format (default)

On the right, a "Preview" section displays the generated `datapackage.json` file content:

```
{
  "profile": "data-package",
  "resources": [
    {
      "name": "resource1",
      "path": "cnj_ods.xlsx",
      "profile": "tabular-data-resource",
      "schema": {
        "fields": [
          {
            "name": "ODS Relacionado",
            "type": "string",
            "format": "default"
          },
          {
            "name": "Unidade de Implan",
            "type": "string",
            "format": "default"
          },
          {
            "name": "Titulo",
            "type": "string",
            "format": "default"
          },
          {
            "name": "Implantacao",
            "type": "any",
            "format": "default"
          },
          {
            "name": "Autores",
            "type": "string",
            "format": "default"
          },
          {
            "name": "Email Autores",
            "type": "string",
            "format": "default"
          }
        ]
      }
    }
  ]
}
```

Fonte: Elaborado pela autora (2020).

Os resultados da etapa empacotar permitiram a transformação da base de boas práticas relacionadas ao ODS em um pacote de dados, conceito chamado *Data Package*, dentro da especificação *Frictionless Data*. Considerou-se que a etapa empacotar é o coração do modelo proposto e os resultados mostraram que o objetivo da etapa foi atingido de maneira positiva.

Na próxima seção será apresentada a validação da etapa tratar do modelo e os resultados obtidos.

3.3.1.2 Verificação da etapa tratar

O próximo passo é validar a etapa tratar do presente modelo. A ferramenta usada foi o *Goodtables*, a qual facilita a validação dos dados, evitando a publicação de dados com erros. Com o *Goodtables* é possível verificar automaticamente se há erros como cabeçalhos em branco, linhas duplicadas, dados com tipo errado (por exemplo, deve ser um número, as é uma data).

Após a etapa empacotar, optou-se por disponibilizar o arquivo *datapackage.json* no repositório *Github*, com o objetivo de facilitar a realização da etapa tratar. O processo de validação dos dados por meio da ferramenta *Goodtables* consiste em três passos (FRICTIONLESS DATA, 2018):

- a) acessar o sítio <https://Goodtables.io>;
- b) no campo ‘Manage Sources’, clicar na opção *Github*; em seguida, clicar no ícone  , para selecionar o arquivo *datapackage.json* criado na etapa empacotamento; o *Goodtables* também permite integração com outros repositórios em nuvem, como o Amazon S3 e integração de APIs;
- c) clicar em ‘Validate’ para realizar a validação.

Em seguida, ao selecionar a opção *Jobs*, no menu esquerdo da ferramenta, pode-se verificar se os dados são válidos ou não. A Figura 17 mostra claramente o resultado da validação do arquivo *datapackage.json* usado na presente pesquisa.

Figura 17 - Ferramenta Goodtables com resultado da validação do arquivo datapackage.json

The screenshot shows the Goodtables.io interface. On the left, a sidebar with a dark blue background contains the following navigation items: Dashboard (with a bar chart icon), Jobs (selected, with a document icon), Manage Sources (with a gear icon), Help (with a question mark icon), Feedback (with a speech bubble icon), Privacy policy (with a shield icon), and Log out (with a user profile icon). The main content area has a light gray background. At the top, it shows the GitHub repository `melffagundes/cnj_ods` with a status badge indicating "data valid". To the right of the badge is a GitHub icon. Below this, a green horizontal bar contains the text "Job #3 success" with a checkmark icon and "5 days ago". To the right of the bar, it says "Pushed by melffagundes bc7087 on master". On the right side of the main content area, there is a sidebar titled "3 jobs" showing a list of recent activities:

- Add files via upload** (5 days ago, #3 (bc7087))
- Update README.md** (8 days ago, #2 (b549e5))
- Add files via upload** (23 days ago, #1 (991b22))

The main content area displays two tables: "cnj_ods.xlsx" and "cnj_ods_csv.csv". For "cnj_ods.xlsx", it shows "Table 1 of 2" and a "Valid Table" button. For "cnj_ods_csv.csv", it shows "Table 2 of 2" and a "Valid Table" button. At the bottom, there is a section titled "Badge image URL" with a text input field containing the URL https://goodtables.io/badge/github/melffagundes/cnj_ods.svg.

Fonte: Elaborado pela autora (2020).

A etapa tratar do presente modelo contemplou o tratamento da base de dados usada na pesquisa utilizando a ferramenta *Goodtables*, a qual faz parte da iniciativa *Frictionless Data*. O *Goodtables* se mostrou uma eficiente ferramenta para identificação de erros estruturais e de conteúdo, além de permitir rapidamente a correção desses problemas.

Na base de dados usada na presente pesquisa não foram entrados tais erros, o que mostra que no futuro se faz necessário testes com outras bases de dados com problemas como cabeçalhos em branco, linhas duplicadas ou dados com tipo errado.

Na seção seguinte é apresentada a verificação da etapa publicar e seus respectivos resultados.

3.3.1.3 Verificação da etapa publicar

O passo seguinte é validar a etapa publicar do modelo proposto. Utilizou-se a ferramenta Datahub.io, disponível no sítio <https://datahub.io/tools>. Através da instalação de um arquivo disponível para diversas plataformas como Windows e MacOs. Para realizar a tarefa de instalação no ambiente com o sistema operacional MacOs e validar do modelo, seguiu-se as instruções conforme a Figura 18.

Figura 18 - Instalação da ferramenta Datahub.io em ambiente de validação

```
melffagundes@MacBook-Air-de-Melissa Mestrado % gunzip -f data-macos.gz
chmod +x data-macos
mv data-macos /usr/local/bin/data
melffagundes@MacBook-Air-de-Melissa Mestrado % npm install -g data-cli
```

Fonte: Elaborado pela autora (2020).

Em seguida, realizou-se a publicação do arquivo *datapackage.json* na ferramenta Datahub.io, através dos comandos abaixo:

- a) necessário realizar o *login* através do comando *data login*; para fazer o *login* é necessário escolher entre uma conta do *Github* ou *Googl*; ao selecionar, faz-se o redirecionamento para o sítio escolhido para fazer o *login* e acessar a ferramenta Datahub.io;
- b) em seguida, o comando *data push datapackage.json* efetiva a publicação do arquivo *datapackage.json*.

A Figura 19 mostra o resultado da publicação.

Figura 1910 - Publicação do arquivo *datapackage.json* na ferramenta Datahub.io

```
mac@MacBook-Pro-de-Mac Mestrado % sudo data login
? Login with... Github
> Opening browser and waiting for you to authenticate online
> Note: If nothing is loaded in browser please run 'data login -i'
(node:2638) [DEP0066] DeprecationWarning: OutgoingMessage.prototype._headers is deprecated
> Error! listen EADDRINUSE: address already in use 127.0.0.1:3000
mac@MacBook-Pro-de-Mac Mestrado % sudo data push datapackage.json
Completed: datapackage.json
Completed: cnj_ods.xlsx
Uploading [*****] 99% (0.0s left)  cnj_ods.xlsx
👉 your data is published!
🔗 https://datahub.io/melffagundes/cnj_ods/v/1 (copied to clipboard)
```

Fonte: Elaborada pela autora (2020).

Após a publicação, a ferramenta Datahub.io disponibiliza um sítio para acesso aos dados. Com isso, é possível acessá-los através de uma interface *web* e fazer *download* dos arquivos publicados, conforme a Figura 20.

Figura 20 - Sítio para *download* do arquivo datapackage_json

DataHub.io by  — We build solutions that unleash the potential of data. Let's start with yours! [Get in touch now »](#)

DATA HUB ABOUT BLOG FIND DATA COLLECTIONS DOCS PRICING TOOLS CHAT 

Boas Práticas CNJ ODS

 [melffagundes](#)

Files	Size	Format	Created	Updated	License	Source
2	1MB	csv zip	3 months ago	33 seconds ago		

[Download](#) [Developers](#)

Data Files

Download files in this dataset

File	Description	Size	Last changed	Download
 resource1		302kB		csv (302kB) , json (447kB)
 cnj_ods.zip	Compressed versions of dataset. Includes normalized CSV...	275kB		zip (275kB)

[resource1](#) 

Fonte: Elaborado pela autora.

Os resultados da etapa publicar mostraram que a ferramenta Data.io atendeu aos requisitos para disponibilização dos dados após verificação das etapas empacotar e publicar. Além disso, contemplou aspectos como permitir acesso, *upload* e *download* da Base de Boas Práticas do CNJ.

4 CONCLUSÕES E TRABALHOS FUTUROS

Diante do crescimento nos últimos anos da quantidade de dados produzida pelas organizações públicas e a divergência presente no formato desses dados disponíveis para consulta pelo cidadão, que dificulta o trabalho com os dados e seu compartilhamento, além da baixa qualidade dos DAGs descrita na literatura, tendo como possíveis consequências o prejuízo à participação popular e o enfraquecimento da democracia, é que este trabalho se propôs a responder o questionamento de: Como um modelo baseado na especificação *Frictionless Data* pode contribuir para a abertura dos DAGS?

Ao decorrer desta pesquisa foi realizado um levantamento nas bases de produção científica nacionais e internacionais, além de uma busca sobre as boas práticas para abertura de DAGs. Conceitos como qualidade da informação sob a perspectiva da CI, DAGs, princípios *FAIR* e *Frictionless Data* foram abordados no presente trabalho, com o objetivo de embasar a proposição do modelo baseado em *Frictionless Data* para auxiliar a publicação e o compartilhamento de DAGs .

Definiu-se, então, os objetivos geral e específicos que conduziram a concepção do modelo. Por meio do levantamento da literatura e da construção da fundamentação teórica, como também das discussões e aprofundamentos sobre os benefícios e barreiras para abertura de DAGs, conseguiu-se atingir os objetivos: a) identificar e avaliar abordagens recentemente empregadas para publicação de DAGs e b) identificar dificuldades e barreiras apresentadas no processo de publicação de DAGs.

O objetivo c) mapear os dados presentes nas BBPODS foi alcançado com o acesso à Base de Boas Práticas do CNJ relacionadas ao ODS cujo detalhamento se encontra na seção 3.3 SUJEITO DE PESQUISA, CONTEXTO E COLETA DE DADOS. O objetivo d) Conceber estrutura de modelo para a publicação das BBPODS foi cumprido por meio da descrição da estrutura do modelo e suas respectivas etapas, expostas em detalhes na seção 3.2.2 Apresentação do Modelo e metodologia, com a com a especificação das etapas: empacotar, tratar, publicar e *feedback*.

Já o objetivo e) Demonstrar a aplicação prática do modelo baseado em *Frictionless Data* em uma BBPODS foi obtido com a verificação das etapas presentes no modelo utilizando ferramentas da iniciativa *Frictionless Data*, cuja descrição se encontra na seção 3.4.1 Aplicação do modelo e resultados.

Os trabalhos relacionados ao contexto desta pesquisa mostraram um consenso sobre a necessidade de um *framework* ou modelo de disponibilização de dados dos governos, com o objetivo de trazer maior unicidade e respeito a determinados critérios internacionalmente aceitos. Além disso, ao padronizar métodos e desenvolvimento de metadados, por exemplo, pode-se aumentar o acesso ao dado. Outro problema presente no cenário de abertura de dados é a falta de *feedback* que mostre o que é feito com os dados abertos. Percebeu-se uma lacuna relacionada à possibilidade de *feedback* para reportar problemas na qualidade dos dados, assim como viabilizar que os cidadãos enviem sugestões de melhorias nos conjuntos de DAGs.

Os resultados decorrentes da verificação do modelo mostraram que a etapa empacotar permitiu a transformação da base de boas práticas relacionadas ao ODS em um pacote de dados, conceito chamado *Data Package*, dentro da especificação *Frictionless Data*. A etapa colaborou com o objetivo geral da presente pesquisa ao disponibilizar os dados independente do formato, uma vez que o *Frictionless Data* não impõe nenhum requisito em sua forma ou estrutura, podendo ser usado para empacotar qualquer tipo de dado. Logo, considera-se que a etapa empacotar é o coração do modelo proposto e os resultados mostraram que o objetivo da etapa foi atingido de maneira afirmativa.

A etapa tratar do presente modelo buscou contemplar o tratamento da base de dados usada na pesquisa utilizando a ferramenta *Goodtables*, a qual faz parte da iniciativa *Frictionless Data*. Os resultados apresentados nessa etapa sugerem que a etapa tratar contemplou dimensões essenciais de qualidade da informação como concisão, consistência e precisão, ao realizar a validação do conjunto de dados, verificando problemas de qualidade da informação como cabeçalhos em branco, linhas duplicadas e dados com tipo errado.

Os resultados da etapa publicar do presente modelo foram apresentados na seção 3.4.1.3, na qual se buscou atender o principal desafio: DAGs não têm valor em si mesmos, senão quando permite que os usuários os utilizem de forma eficaz. Logo, faz-se necessário disponibilizar os dados em formato aberto e com qualidade. Durante a presente pesquisa, após a verificação das etapas de empacotar e tratar, expostas em seções anteriores, foi possível publicar os dados por meio da ferramenta Data.io, pertencente à iniciativa *Frictionless Data*.

A etapa publicar cumpriu o seu objetivo ao contemplar aspectos fundamentais para publicação e disponibilidade, ao permitir o acesso, *upload* e *download* da Base

de Boas Práticas do CNJ relacionadas aos ODS. Requisitos como visualização foram atendidos com o uso da ferramenta Data.io.

Finalmente, etapa *feedback* proposta no presente modelo foi estabelecida com o objetivo de atender a lacuna encontrada na literatura com relação aos modelos existentes para publicação e abertura de dados abertos, ao acrescentar a participação e *feedback* sobre as publicações. A intenção da etapa era uma possível contribuição por meio da criação de ciclos de *feedback*, com os quais o governo pode aprender com o público e promover melhorias sobre a qualidade dos dados abertos. Durante o decorrer da aplicação do modelo, não foi possível verificar a etapa *feedback* utilizando ferramentas da iniciativa *Frictionless Data*, posto que a especificação não possui uma ferramenta que atenta ao ciclo de *feedback* dos dados.

No entanto, a etapa *feedback* no presente modelo foi concebida sem a intenção de determinar uma ferramenta específica ou processo fixo, de modo a flexibilizar a escolha da tecnologia ou ferramenta conforme necessidade do órgão público. Logo, a ausência de uma ferramenta na iniciativa *Frictionless Data* que permita acompanhar o ciclo de *feedback* não compromete o uso do modelo, uma vez que a presença da etapa atende a lacuna identificada na literatura, tornando-se uma oportunidade para desenvolvimento futuro de uma ferramenta pela comunidade.

Os resultados do presente trabalho foram considerados positivos e conduziram ao alcance dos objetivos elencados, através da apresentação do modelo e respectiva verificação das etapas. Um dos pontos diferenciais com relação aos modelos para publicação de DAGs encontrados na literatura estão o uso da especificação *Frictionless Data* e o Modelo de Referência para Publicação de Dados Abertos do governo brasileiro como base para um modelo de abertura dos DAGs no contexto do judiciário brasileiro, como também a inclusão da etapa *feedback* para atender a lacuna encontrada nos demais modelos.

O *Frictionless Data* tem sido usado em iniciativas internacionais no âmbito governamental como a *Workforce Data Initiative*, a qual pretende modernizar a força de trabalho dos EUA por meio do uso dos dados. O objetivo dessa iniciativa é ajudar os conselhos estaduais e locais de força de trabalho a coletar, agregar e distribuir estatísticas sobre a eficácia dos provedores de treinamento e programas para esses provedores.

O Departamento de Trabalho dos EUA exige que os provedores de treinamento elegível (ETP) trabalhem com as *Workforces* dos conselhos estaduais para

acompanhar os resultados de seus alunos a fim de receber financiamento federal. Para tal tarefa, estão construindo um conjunto de ferramentas de código aberto baseado na especificação *Frictionless Data* chamada *etp-uploader*, uma ferramenta que permite aos provedores de treinamento carregar os dados em uma única plataforma.

São muitas centenas ou milhares de provedores de treinamento sob a alçada de cada Conselho e cada um deve carregar com segurança seus dados e, ao mesmo tempo, deve ser fácil para a própria placa processar e validar automaticamente os conjuntos de dados. Uma ferramenta como o *etp-uploader* poderia ser desenvolvida com base no modelo proposto e usada para enviar os dados das boas práticas relacionadas aos ODS dos vários Tribunais de Justiça para o CNJ em formato *Data Package*.

Outra possibilidade como trabalho futuro seria o uso do *Data Package Pipelines*, uma estrutura declarativa baseada em fluxo para a construção de *pipelines* de processamento de dados tabulares. *Pipelines* são definidos como uma forma de paralelização de tarefas em um sistema que seja executado de forma decomposta, ou seja, permite que um fluxo de tarefas possa dividir suas etapas de trabalho em unidades ou estágio e as executa simultaneamente. Cada estágio do *pipeline* é normalmente a entrada de sua fila, que é a fila de saída do estágio anterior, e depois de processado o dado, serve como entrada do próximo estágio.

O *Data Package Pipelines* poderia ser usado para todas as tarefas de extração, transformação e carregamento (ETL) como estrutura para facilitar a implementação do modelo proposto, pois é particularmente adequado para trabalhar com fontes de dados diversas e heterogêneas de qualidade variável e desconhecida. Existem muitas ferramentas e estruturas para fazer trabalho de ETL com dados, no entanto, o *Data Package Pipelines* tem como foco aqui organizar e controlar dados confusos de várias fontes e usar as ferramentas de dados sem fricção para transformar esses dados em um fluxo de dados limpo e consistente.

Quanto às contribuições, a presente pesquisa, ao promover a abertura dos dados das boas práticas do judiciário relacionadas ao ODS, colaborou com a Agenda Global 2030 no sentido de disponibilizar os dados em formato aberto, de modo que possam ser usados como subsídios para aumentar a transparência sobre as bases de boas práticas desenvolvidas no âmbito do judiciário brasileiro. A partir da abertura desses dados, por exemplo, seria possível ao cidadão consumir esses dados,

desenvolvendo ferramentas inovadoras que poderiam auxiliar os tribunais na implementação dos ODS.

A pesquisa também contribuiu com as temáticas de estudo do PPGINFO, ao contemplar conceitos como qualidade da informação no contexto dos DAGs e apresentar caráter original quanto à interdisciplinaridade entre diversas áreas, como Ciência da Informação e Tecnologia.

Ainda sobre outras possíveis contribuições, o presente trabalho pode colaborar, por exemplo, de forma semelhante, com outra iniciativa que utiliza os ODS para acompanhamento de resultados, Portal PPA Cidadão, que foi concebido para permitir ao cidadão acompanhar a implantação de metas e objetivos do Plano Plurianual do Governo Federal e os ODS, sendo uma parceria do Governo Federal com o OGP e coordenado pela Corregedoria Geral da União. Logo, a aplicabilidade e a generalização desta pesquisa têm grande potencial de contribuição científica e social.

O modelo proposto se mantém aberto a novas aplicações e sugestão de melhorias, principalmente no que tange a sua aplicabilidade em outros contextos como, por exemplo, integração e abertura de bases públicas relacionadas aos ODS. O aperfeiçoamento da etapa *feedback*, com o uso de outras ferramentas e adição de metodologias como Net Promoter Score, poderia contribuir para medir a satisfação do cidadão com relação aos DAGs publicados, podendo permitir um aprimoramento da qualidade dos dados por parte dos governos.

A presente proposta de pesquisa pode ainda, ao contribuir com o reuso e compartilhamento dos dados, permitir a construção de modelos para reconhecimento de padrões de DAGs e, no caso do uso das Bases de Boas Práticas do Judiciário, possibilitar a classificação automática das boas práticas e as temáticas de desenvolvimento sustentável, sendo a sua aplicação possível em diversas áreas, contribuindo para com a implementação da Agenda Global 2030.

REFERÊNCIAS

ALBANO, Cláudio Sonáglia. **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: <https://teses.usp.br/teses/disponiveis/12/12139/tde-03062014-170642/publico/ClaudioSonaglioAlbanoVC.pdf>. Acesso em: 10 out. 2019.

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**, Florianópolis, v. 7, n. 1, p. 73-92, jan./mar. 2017. DOI: <http://dx.doi.org/10.22279/navus.2017.v7n1.p73-92.462>. Disponível em: <http://navus.sc.senac.br/index.php/navus/article/view/462/pdf>. Acesso em: 10 jun. 2019.

ALBANO, Cláudio Sonáglia; CRAVEIRO, Gisele da Silva. Lições aprendidas com a utilização de dados orçamentários em formato aberto: um estudo exploratório no ecossistema brasileiro. **Revista de Gestão e Projetos**, São Paulo, v. 6, n. 3, p. 17-27, set./dez. 2015. DOI: <https://doi.org/10.5585/gep.v6i3.370>. Disponível em: <http://www.spell.org.br/documentos/ver/40161/licoess-aprendidas-com-a-utilizacao-de-dados-orcamentarios-em-formato-aberto--um-estudo-exploratorio-no-ecossistema-brasileiro>. Acesso em: 10 set. 2019.

ALBUQUERQUE, Almir Dos Santos; BASTOS, Rogério Cid; LINO, Manuel Rosa de Oliveira. A qualidade da informação em portais do judiciário. **Encontros Bibli**, Florianópolis, v. 14, n. 27, p. 115-134, 2009. DOI: <http://dx.doi.org/10.5007/1518-2924.2009v14n27p115>. Disponível em: <https://brapci.inf.br/index.php/res/v/36133>. Acesso em: 10 mar. 2020.

ALDEBERT, Geoffrey; AUGUSTI, Antoine. **Advancing open data quality**. 2020. Disponível em: <https://frictionlessdata.io/blog/2020/05/22/etalab-case-study-schemas-data-gouv-fr/#what%E2%80%99s-a-schema>. Acesso em: 23 jun. 2020.

ARAUJO, Marcelo Henrique de *et al.* Dados governamentais abertos: uma análise sob a ótica das dimensões de qualidade da informação. In: ENCONTRO DA ASSOCIAÇÃO NACIONAL DE PÓS-GRADUAÇÃO E PESQUISA EM ADMINISTRAÇÃO, 36., 2012, Rio de Janeiro. **Anais** [...]. Rio de Janeiro: UERJ, 2012. p. 1-16. Disponível em: https://www.researchgate.net/publication/236342272_Dados_Governamentais_Abertos_Uma_analise_sob_a_otica_das_dimensoes_de_Qualidade_da_Informacao. Acesso em: 10 fev. 2020.

ASSEMBLY, U. G. Calling of an international conference on freedom of information. **Resolution**, v. 59, n. 1, p. 14, 1946.

ÁVILA, Thiago José Tavares. **Uma proposta de modelo de processo para publicação de dados abertos conectados governamentais**. 2015. Dissertação

(Mestrado em Modelagem Computacional do Conhecimento) – Instituto de Computação, Universidade Federal de Alagoas, Maceió, 2015. Disponível em: <http://www.repositorio.ufal.br/handle/riufal/2209>. Acesso em: 10 jun. 2019.

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. Disponível em: <https://doi.org/10.1080/01616846.2017.1313045>. Acesso em: 10 set. 2019.

BARBALHO, Fernando Almeida. A emergência do campo de políticas públicas de dados abertos governamentais no Brasil. **Conhecer: debate entre o público e o privado**, Ceará, v. 8, n. 20, p. 118-137, 2018. DOI: <https://doi.org/10.32335/2238-0426.2018.8.20.1056>. Disponível em: <https://revistas.uece.br/index.php/revistaconhecer/article/view/1056/962>. Acesso em: 12 jan. 2020.

BARRETO, Aldo de Albuquerque. Uma quase história da Ciência da Informação. **DataGamaZero**, Rio de Janeiro, v. 9, n. 2, p. 1-18, abr. 2008. Disponível em: http://eprints.rclis.org/17637/1/DataGamaZero%20-20Revista%20de%20Ci%C3%A3Ancia%20da%20Informa%C3%A7%C3%A3o%20-Artigo%2001_Aldo.pdf. Acesso em: 14 ago. 2019.

BARRY, Emily; BANNISTER, Frank. Barriers to open data release: a view from the top. **Information Polity**, v. 19, n. 1-2, p. 129-152, 2014. Disponível em: <https://pdfs.semanticscholar.org/9f8d/7653ab54e63141487fa9e18c8cb4fa920cbf.pdf>. Acesso em: 20 out. 2019.

BERNERS-LEE, Tim. **Linked Data - Design issues**. 2006. Disponível em: <https://www.w3.org/DesignIssues/LinkedData.html>. Acesso 25 ago. 2019.

BOBROWSKI, Mónica; MARRÉ, Martina; YANKELEVICH, Daniel. A Homogeneous framework to measure data quality. **Fourth Conference on Information Quality**, v. 25, n. 3, p. 115-124, 1999. Disponível em: https://www.researchgate.net/publication/220918947_A_Homogeneous_Framework_to_Measure_Data_Quality. Acesso em: 10 set. 2019.

BORKO, H. Information science: what is it? **American Documentation**, v. 3, n. 5, 1968.

BRASIL. Constituição (1988). **Constituição da República Federativa do Brasil**. Brasília, DF: Centro Gráfico, 1988.

BRASIL. **Decreto nº 8.777, de 12 de maio de 2016**. Institui a Política de Dados Abertos do Poder Executivo Federal. Brasília, DF: Presidência da República, 2016. Disponível em: https://www.in.gov.br/materia-/asset_publisher/Kujrw0TzC2Mb/content/id/21520099/do1-2016-05-12-decreto-n-8-777-de-11-de-maio-de-2016-21520002. Acesso em: 9 abr. 2020.

BRASIL. Lei nº 12.527, de 18 de novembro de 2011. Regula o acesso a informações previsto no inciso XXXIII do art. 5º, no inciso II do § 3º do art. 37 e no § 2º do art. 216 da Constituição Federal; altera a Lei nº 8.112, de 11 de dezembro de 1990; revoga a Lei nº 11.111, de 5 de maio de 2005, e dispositivos da Lei nº 8.159, de 8 de janeiro de 1991; e dá outras providências. Brasília, DF: Presidência da República, 2011. Disponível em: http://www.planalto.gov.br/ccivil_03/_Ato2011-2014/2011/Lei/L12527.htm. Acesso em: 23 ago. 2019.

BRASIL. Lei nº 16.817, de 02 de fevereiro de 2018. Adota a Agenda 2030 para o Desenvolvimento Sustentável da Organização das Nações Unidas (ONU) como diretriz de políticas públicas em âmbito municipal, institui o Programa de sua implementação, autoriza a criação da Comissão Municipal para o Desenvolvimento Sustentável (Agenda 2030) e dá outras providências. Brasília, DF: Presidência da República, 2018. Disponível em: [http://legislacao.prefeitura.sp.gov.br/leis/lei16817de2defevereirode2018#:~:text=Adot a%20a%20Agenda%202030%20para,2030\)%20e%20d%C3%A1%20outras%20prov id%C3%A9ncias](http://legislacao.prefeitura.sp.gov.br/leis/lei16817de2defevereirode2018#:~:text=Adota%20a%20Agenda%202030%20para,2030)%20e%20d%C3%A1%20outras%20provid%C3%A9ncias). Acesso em: 23 ago. 2019.

BRASIL. Ministério da Transparência e Controladoria Geral da União. Modelo de referência para publicação de dados abertos é lançado. **Gov.br**, 2020. Disponível em: <https://www.gov.br/cgu/pt-br/governo-aberto/noticias/2020/10/modelo-de-referencia-para-publicacao-de-dados-abertos-e-lancado>. Acesso em: 20 nov. 2020.

BRASIL. Portaria nº 119, de 21 de agosto de 2019. Institui o Laboratório de Inovação, Inteligência e Objetivos de Desenvolvimento Sustentável (LIOD) e dá outras providências. Brasília, DF: Presidência da República, 2019. Disponível em: https://atos.cnj.jus.br/files//portaria/portaria_119_21082019_22082019154612.pdf. Acesso em: 23 ago. 2019.

BRASIL. Portaria nº 133, de 28 de setembro de 2018. Institui Comitê Interinstitucional destinado a proceder estudos e apresentar proposta de integração das metas Poder Judiciário com as metas e indicadores dos Objetivos de Desenvolvimento Sustentável (ODS), Agenda 2030. Brasília, DF: Presidência da República, 2018. Disponível em: https://atos.cnj.jus.br/files//portaria/portaria_133_28092018_24102018134936.pdf. Acesso em: 23 ago. 2019.

BRASIL. Portaria nº 133, de 28 de setembro de 2018. Institui Comitê Interinstitucional destinado a proceder estudos e apresentar proposta de integração das metas do Poder Judiciário com as metas e indicadores dos OSD, Agenda 2030. Brasília, DF: Presidência da República, 2018. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/atos-normativos?documento=2721>. Acesso em: 23 ago. 2019.

BRASIL. Portaria nº 148, de 20 de novembro de 2018. Designa membros do Comitê Interinstitucional destinado a proceder estudos e apresentar proposta de integração das metas do Poder Judiciário com as metas e indicadores dos Objetivos de Desenvolvimento Sustentável (ODS), Agenda 2030. Brasília, DF: Presidência da República, 2018. Disponível em:

https://atos.cnj.jus.br/files//portaria/portaria_148_20112018_22112018122206.pdf. Acesso em: 23 ago. 2019.

BREITMAN, Karin *et al.* Open government data in Brazil. **IEEE Intelligent Systems**, v. 27, n. 3, p. 45-49, 2012. DOI: <https://doi.org/10.1109/MIS.2012.25>. Disponível em: https://www.researchgate.net/publication/254061491_Open_Government_Dat_in_Brazil. Acesso em: 12 jun. 2020.

CALAZANS, Angélica Toffano Seidel. Qualidade da informação: conceitos e aplicações. **Transinformação**, Campinas, v. 20, n. 1, p. 29-45, jan./abr. 2008. Disponível em: <https://www.scielo.br/pdf/tinf/v20n1/03.pdf>. Acesso em: 12 dez. 2019.

CARDOSO, Paulo Henrique. **Ciência de dados aplicada a dados governamentais abertos sob a ótica da Ciência da Informação**. 2019. Dissertação (Mestrado em Ciência da informação) – Escola de Comunicação, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2019. Disponível em: <https://ridi.ibict.br/bitstream/123456789/1039/1/Ci%C3%A3ncia%20de%20Dados%20aplicada%20a%20Dados%20Governamentais%20Abertos%20sob%20a%20%C3%B3tica%20da%20Ci%C3%A3ncia%20da%20Informa%C3%A7%C3%A3o-V6.pdf>. Acesso em: 20 abr. 2019.

COLLINS, Sandra *et al.* **FAIR data action plan**: interim recommendations and actions from the european commission expert group on FAIR data. 2018. Disponível em: <https://zenodo.org/record/1285290#.X3u-mmhKjIU>. Acesso em: 17 fev. 2020.

CONSELHO NACIONAL DA JUSTIÇA. **Laboratório de inovação, inteligência e ODS**. 2020. Disponível em: <https://www.cnj.jus.br/programas-e-acoes/agenda-2030/liods-cnj-laboratorio-de-inovacao-inteligencia-e-ods/>. Acesso em: 20 ago. 2020.

CONSELHO NACIONAL DE JUSTIÇA. Portaria nº 133, de 28 de setembro de 2018. Institui Comitê Interinstitucional destinado a proceder estudos e apresentar proposta de integração das metas do Poder Judiciário com as metas e indicadores dos Objetivos de Desenvolvimento Sustentável (ODS), Agenda 2030. Brasília, DF, 28 set. 2018. Disponível em: <https://atos.cnj.jus.br/atos/detalhar/2721>. Acesso em: 20 ago. 2020.

CONSELHO NACIONAL DA JUSTIÇA. **Tribunal do DF apresenta painéis e plano para meta desjudicialização no LIODS**. 2020. Disponível em: <https://www.cnj.jus.br/tribunal-do-df-apresenta-paineis-e-plano-para-meta-de-desjudicializacao-no-liods/>. Acesso em: 20 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. DOI: <https://doi.org/10.1016/j.giq.2019.03.004>. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0740624X18305185?via%3Dhub>. Acesso em: 12 jun. 2019.

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, set. 2017. DOI: <https://doi.org/10.1145/3110291>. Disponível em: <https://dl.acm.org/doi/10.1145/3110291>. Acesso em: 12 ago. 2019.

DALENOGARE, Lucas Geraldo Campos; ARAÚJO, Maria Arlete Duarte de. Abordagens teóricas em dados governamentais abertos. **Revista Gestão & Tecnologia**, Pedro Leopoldo, v. 19, n. 5, p. 296-314, out./dez. 2019. DOI: <https://doi.org/10.20397/2177-6652/2019.v19i5.1382>. Disponível em: <http://revistagt.fpl.edu.br/get/article/view/1382>. Acesso em: 28 maio 2020.

DATAHUB. **cnj_ods.xlsx download**. 2018. Disponível em: https://datahub.io/melffagundes/cnj_ods.xlsx/v/1. Acesso em: 29 jul. 2020.

DATAHUB. **Data Factory**. 2018. Disponível em: <https://datahub.io/data-factory>. Acesso em: 29 jul. 2020.

DATAHUB. **Ferramenta DataHub**. 2020. Disponível em: <https://datahub.io>. Acesso em: 29 jul. 2020.

DATA PACKAGE CREATOR. **Ferramenta Data Package Creator**. 2020. Disponível em: <https://create.frictionlessdata.io/>. Acesso em: 20 ago. 2020.

DAVENPORT, Thomas H. **Ecologia da informação**: por que só a tecnologia não basta para o sucesso na era da informação. 6.ed. São Paulo: Futura, 1998.

DAWES, Sharon; HELBIG, Natalie. Information strategies for open government: challenges and prospects for deriving public value from government transparency. **Electronic Government**, v. 62, n. 28, p. 50-60, 2010. DOI: https://doi.org/10.1007/978-3-642-14799-9_5. Disponível em: https://link.springer.com/chapter/10.1007/978-3-642-14799-9_5. Acesso em: 20 out. 2019.

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. DOI: <http://dx.doi.org/10.31763/businta.v1i2.26>. Disponível em: https://www.researchgate.net/publication/311379248_A_Framework_Design_to_Develop_Integrated_Data_System_for_Smart_eGovernment_Based_on_Big_Data_Technology. Acesso em: 27 set. 2019.

EPPLER, Martin.; WITTIG, Dörte. Conceptualizing information quality: a review of information quality frameworks from the last ten years. **Proceedings of the 2000 Conference on Information Quality**, v. 10, n. 4, p. 1-14, 2000. Disponível em: <http://mitiq.mit.edu/iciq/Documents/IQ%20Conference%202000/Papers/ConceptIQAReviewofIQFramework.pdf>. Acesso em: 10 set. 2019.

ESCIENCE VIDENPORTAL. **FAIR Across Disciplines**. 2020. Disponível em: <https://vidensportal.deic.dk/en/FAIRAcross>. Acesso em: 10 maio 2020.

FIGUEIREDO, Ana Sofia. Data sharing: convert challenges into opportunities. **Front Public Health**, v. 5, n. 327, p. 1-6, dez. 2017. DOI:

<https://doi.org/10.3389/fpubh.2017.00327>. Disponível em:
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5723929/>. Acesso em: 12 dez. 2019.

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. DOI: <https://doi.org/10.1080/1369118X.2017.1330418>. Disponível em:
<https://www.tandfonline.com/doi/abs/10.1080/1369118X.2017.1330418>. Acesso em: 20 out. 2019.

FOWLER, Dan; BARRATT, Jo; WALSH, Paul. Frictionless data: making research data quality visible. **International Journal of Digital Curation**, v. 2, n. 2, p. 274-285, 2017. DOI: <https://doi.org/10.2218/ijdc.v12i2.577>. Disponível em:
<http://www.ijdc.net/article/view/577>. Acesso em: 12 mar. 2020.

FRICTIONLESS DATA SPECS. **Data Package**. 2017. Disponível em:
<https://specs.frictionlessdata.io/data-package/#language>. Acesso em: 16 ago. 2020.

FRICTIONLESS DATA SPECS. **Metadata**. [2020]. Disponível em:
<https://specs.frictionlessdata.io/data-package/#metadata>. Acesso em: 03 nov. 2020.

FRICTIONLESS DATA. **Data publication workflow example**. 2018. Disponível em:
<https://frictionlessdata.io/blog/2018/03/12/data-publication-workflow-example/#step-1-package-our-data-as-a-data-package>. Acesso em: 16 ago. 2020.

FRICTIONLESS DATA. **Well packaged datasets**. 2018. Disponível em:
<https://frictionlessdata.io/blog/2018/03/07/well-packaged-datasets/#write-a-table-schema>. Acesso em: 16 ago. 2020.

GALLOPÍN, Gilberto; VESSURI, Hebe. Science for sustainable development. **Interfaces between Science and Society**, v. 336, p. 35-51, jun. 2019. Disponível em:
<https://www.taylorfrancis.com/books/e/9781351280440/chapters/10.4324/9781351280440-3>. Acesso em: 10 abr. 2020.

GERMANO, Edson Carlos. **Modelos de negócios adotados para o uso de dados governamentais abertos**: estudo exploratório de prestadores de serviços na cadeia de valor dos dados governamentais abertos. 2013. Dissertação (Mestrado em Ciências) – Faculdade de Economia, Administração e Contabilidade, Universidade de São Paulo, São Paulo, 2013. Disponível em:
<https://teses.usp.br/teses/disponiveis/12/12139/tde-10012014-155226/publico/EdsonCarlosGermanoVC.pdf>. Acesso em: 10 out. 2019.

GOFAIR. **FAIR Principles**. 2020. Disponível em: <https://www.go-fair.org/fair-principles/>. Acesso em: 9 de maio 2020.

GONZALEZ-ZAPATA, Felipe; HEEKS, Richard. The multiple meanings of open government data: Understanding different stakeholders and their perspectives. **Government Information Quarterly**, v. 32, n. 4, p. 1-40, 2015. DOI: <https://doi.org/10.1016/j.giq.2015.09.001>. Disponível em: https://www.researchgate.net/profile/Richard_Heeks/publication/282433691_The_multiple_meanings_of_open_government_data_Understanding_different_stakeholders_and_their_perspectives/links/56799e6008ae40c0e27dd06a/The-multiple-meanings-of-open-government-data-Understanding-different-stakeholders-and-their-perspectives.pdf. Acesso em: 20 out. 2019.

GOODTABLES. Ferramenta Goodtables. 2020. Disponível em: <http://goodtables.io/>. Acesso em: 10 ago. 2020.

GOVERNO DIGITAL. **Dados abertos governamentais**. 2019. Disponível em: <https://www.gov.br/governodigital/pt-br/dados-abertos/dados-abertos-governamentais#:~:text=Os%20Dados%20Abertos%20Governamentais%20s%C3%A3o,aplica%C3%A7%C3%A7%C3%B5es%20desenvolvidas%20colaborativamente%20pela%20sociedade. Acesso em: 20 dez. 2019.>

HASNAIN, Ali; REBOLZ-SCHUHMANN, Dietrich. Assessing FAIR data principles against the 5-Star open data principles: ESWC 2018 satellite events, heraklion, crete, greece, june 3-7, 2018, revised selected papers. **The Semantic Web**, p. 469-477, 2018. DOI: https://doi.org/10.1007/978-3-319-98192-5_60. Disponível em: https://www.researchgate.net/publication/326757298_Assessing_FAIR_Data_Principles_Against_the_5Star_Open_Data_Principles_ESWC_2018_Satellite_Events_Heraklion_Crete_Greece_June_3-7_2018_Revised_Selected_Papers. Acesso em: 12 ago. 2019.

HENNING, Patrícia Corrêa et al. GO FAIR e os princípios FAIR: o que representam para a expansão dos dados de pesquisa no âmbito da ciência aberta. **Em Questão**, Porto Alegre, v. 25, n. 2, p. 389-412, maio/ago. 2019. DOI: <http://dx.doi.org/10.19132/1808-5245252.389-412>. Disponível em: <https://seer.ufrgs.br/EmQuestao/article/view/84753>. Acesso em: 14 ago. 2019.

HIGMAN, Rosie; BANGERT, Daniel; JONES, Sarah. Three camps , one destination : the intersections of research data management , FAIR and open. **Three camps, one destination: RDM**, v. 32, p. 1-9, 2019. Disponível em: <https://goedoc.uni-goettingen.de/bitstream/handle/1/16174/468-4518-1-PB.pdf?sequence=1&isAllowed=y>. Acesso em: 12 ago. 2019.

INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA.
Serviço de Informação ao Cidadão (SIC). 2018. Disponível em:
<https://ibict.br/acesso-a-informacao/servico-de-informacao-ao-cidadao-sic>. Acesso em: 09 ago. 2020.

INSTITUTO DE PESQUISA ECONÔMICA APLICADA. **Objetivos de desenvolvimento sustentável.** 2019. Disponível em: <https://www.ipea.gov.br/ods/ods16.html>. Acesso em: 10 jan. 2020.

IZDEBSKI, Krzysztof. Transparency and open data principles: why they are

important and how they increase public participation and tackle corruption. 2015. Disponível em: <https://transparencee.org/wpcontent/uploads/2015/12/open-data-principles-by-krzysztof-izdebski.pdf>. Acesso em: 20 out. 2019.

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. DOI: <https://doi.org/10.1080/10580530.2012.716740>. Disponível em: <https://www.tandfonline.com/doi/abs/10.1080/10580530.2012.716740>. Acesso em: 12 set. 2019.

KAHN, Becerly; STRONG, Diane; WANG, Richard. Information quality benchmarks: product and service performance. **Communications of the ACM**, v. 45, n. 4, p. 184-192, 2002. DOI: <https://doi.org/10.1145/505999.506007>. Disponível em: https://www.researchgate.net/publication/220422696_Information_Quality_Benchmarks_Product_and_Service_Performance. Acesso em: 14 ago. 2019.

KALAMPOKIS, Evangelos; TAMBOURIS, Efthimios; TARABANIS, Konstantinos. Open government data: a stage model. **Electronic Government**, v. 68, n. 46, p. 235-246, 2011. Disponível em: https://link.springer.com/chapter/10.1007/978-3-642-22878-0_20. Acesso em: 20 nov. 2019.

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, out. 2013. DOI: <http://dx.doi.org/10.1016/j.giq.2013.05.012>. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0740624X13000683>. Acesso em: 12 ago. 2019.

KIM, Yong Jin; KISHORE, Rajv; SANDERS, Lawrence. From DQ to EQ: understanding data quality in the context of e-business systems. **Communications of the ACM**, v. 48, n. 10, p. 75-81, 2005. DOI: <https://doi.org/10.1145/1089107.1089108>. Disponível em: <https://dl.acm.org/doi/10.1145/1089107.1089108>. Acesso em: 18 dez. 2019.

KLEIN, Rodrigo Hickmann; KLEIN, Deisy Barbiero; LUCIANO, Edimara Mezzomo. Dados abertos governamentais: conceitos, abordagens e dimensões ao longo do tempo. **E&G: Economia e Gestão**, Belo Horizonte, v. 18, n. 49, jan./abr. 2018a. DOI: <https://doi.org/10.5752/P.1984-6606.2018v18n49p4-24>. Disponível em: <http://periodicos.pucminas.br/index.php/economiaegestao/article/view/15581>. Acesso em: 20 jun. 2019.

KLEIN, Rodrigo Hickmann; KLEIN, Deisy Cristina Barbiero; LUCIANO, Edimara Mezzomo. Identificação de mecanismos para a ampliação da transparência em portais de dados abertos: uma análise no contexto brasileiro. **Cadernos EBAPE.BR**, Rio de Janeiro, v. 16, n. 4, p. 692-715, out./dez. 2018b. DOI: <https://doi.org/10.1590/1679-395173241>. Disponível em: <https://www.scielo.br/pdf/cebape/v16n4/1679-3951-cebape-16-04-692.pdf>. Acesso em: 20 jun. 2019.

KUMAR, Vinit. Open government data initiatives: a comparison of India and China. In: INTERNATIONAL CONFERENCE ON BIG DATA AND KNOWLEDGE DISCOVERY, 1., 2016, Bangalone. **Anais** [...]. Bangalone, 2016. Disponível em: https://www.researchgate.net/publication/305768855_Open_Government_Data_Initiatives_a_comparison_of_India_and_China. Acesso em: 24 abr. 2019.

LOURENÇO, Rui Pedro. An analysis of open government portals: a perspective of transparency for accountability. **Government Information Quarterly**, v. 32, n. 3, p. 323-332, 2015. DOI: <https://doi.org/10.1016/j.giq.2015.05.006>. Disponível em: <http://www.periodicoscapes.gov.br>. Acesso em: 10 set. 2019.

MACHADO, Osmar Aparecido. **Qualidade da informação**: uma abordagem orientada para o contexto. 2013. Tese (Doutorado em Ciências) – Escola Politécnica, Universidade de São Paulo, São Paulo, 2013. Disponível em: https://www.teses.usp.br/teses/disponiveis/3/3141/tde-23052014-001437/publico/Tese_OsmarAparecidoMachado_unprotected.pdf. Acesso em: 10 out. 2019.

MARCONI, Marina de Andrade; LAKATOS, Eva Maria. **Fundamentos de metodologia científica**. 5.ed. São Paulo: Atlas 2003.

MARCONI, Marina de Andrade; LAKATOS, Eva Maria. **Técnicas de pesquisa**. 5.ed. São Paulo: Atlas, 2002.

MARTIN, Sébastien et al. Risk analysis to overcome barriers to open data. **Electronic Journal of e-Government**, v. 11, n. 2, p. 348-359, 2013. Disponível em: <http://www.periodicoscapes.gov.br>. Acesso em: 10 set. 2019.

NEHMY, Rosa Maria Quadros; PAIM, Isis. A desconstrução do conceito de qualidade da informação. **Ciência da Informação**, Brasília, v. 27, n. 1, p. 36-45, jan./abr. 1998. Disponível em: <https://www.scielo.br/pdf/ci/v27n1/05.pdf>. Acesso em: 20 abr. 2020.

OBSERVATORIO DE ADMINISTRACIÓN ELECTRÔNICA. **La reutilización de la información del Sector Público**. Governo da Espanha: Ministerio de Maciña y Función Pública, 2017. Disponível em:

https://transparencia.gob.es/transparencia/transparencia_Home/Utilidades/Aviso-legal/reutilizacion.html. Acesso em: 29 out. 2019.

OLIVEIRA, Lairson Emanuel 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, out. 2014. DOI: <https://doi.org/10.6008/SPC2179-684X.2014.002.0010>. Disponível em:

https://www.researchgate.net/publication/291099159_Uma_abordagem_para_captura_de_informacoes_sobre_aplicacoes_que_fazem_uso_de_dados_abertos. Acesso em: 10 jan. 2020.

OLIVEIRA, Wagner (coord.). **Modelo de referência de abertura de dados**: documento de referência do marco 5 do compromisso 2 : ecossistema de dados abertos. Rio de Janeiro: Fgv, 2020. 96 p. Disponível em: https://www.gov.br/cgu/pt-br/governo-aberto/a-ogg/planos-de-acao/4o-plano-de-acao-brasileiro/compromisso-2-docs/modelo-de-referencia-de-abertura-de-dados_versao-final-2.pdf. Acesso em: 20 nov. 2020.

OLSON, Jack E. **Data quality**: the accuracy dimension. São Francisco: Morgan Kaufmann Publishers, 2003. E-book. Disponível em: <https://www.amazon.com/Data-Quality-Accuracy-Dimension-Management/dp/1558608915>. Acesso em: 12 abr. 2020.

OPEN DATA HANDBOOK. **The open data handbook**. 2020. Disponível em: <https://opendatahandbook.org/guide/en/>. Acesso em: 12 ago. 2020.

OPEN KNOWLEDGE FOUNDATION. **Frictionless data specifications and software**: data package. 2018. Disponível em: <https://frictionlessdata.io/specs/data-package/>. Acesso em: 10 set. 2019.

OPEN KNOWLEDGE FOUNDATION. **Goodtables.io**. [2020]. Disponível em: <http://goodtables.io/about>. Acesso em: 10 nov. 2020.

OPEN KNOWLEDGE FOUNDATION. **History**. 2019. Disponível em: <http://opendefinition.org/history/>. Acesso em: 12 jun. 2020.

OPEN KNOWLEDGE FOUNDATION. **Mission**. 2019. Disponível em: <https://okfn.org/about/>. Acesso em: 12 jun. 2020.

OPEN KNOWLEDGE FOUNDATION. **Open data quality**: the next shift in open data? 2017. Disponível em: <https://blog.okfn.org/2017/05/31/open-data-quality-the-next-shift-in-open-data/>. Acesso em: 25 abr. 2020.

OPEN KNOWLEDGE FOUNDATION. **Open definition**. 2019. Disponível em: <http://opendefinition.org/>. Acesso em: 12 jun. 2020.

OPEN KNOWLEDGE FOUNDATION. **Updates**. 2019. Disponível: <http://opendefinition.org/update/>. Acesso em: 12 jun. 2020.

OSORIO-SANABRIA, Mariutsi Alexandra *et al.* Hacia un modelo de evaluación de la preparación para la apertura de datos en instituciones de educación superior colombianas. **Revista Ibérica de Sistemas e Tecnologias de Informação**, Lousada, v. 4, n. 28. Abr. 2020. Disponível em: <https://search.proquest.com/openview/e5411d7ab976c02c08455bbd338fe4f8/1?pq-origsite=gscholar&cbl=1006393>. Acesso em: 20 nov. 2019.

PAIM, Isis; NEHMY, Rosa Maria Quadros; GUIMARÃES, César Geraldo. Problematização do conceito qualidade da informação. **Perspectivas em Ciencia da Informacao**, Belo Horizonte, v. 1, n. 1, p. 111-119, jan./jun. 1996. Disponível em: <http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/8/27>. Acesso em: 10 set. 2019.

PEDROSO, Louise; TANAKA, Asterio; CAPELLI, Claudia. A Lei de Acesso à Informação brasileira e os desafios tecnológicos dos dados abertos governamentais. In: SIMPÓSIO BRASILEIRO DE SISTEMAS DE INFORMAÇÃO, 9., João Pessoa. **Anais** [...]. João Pessoa: UFPB, 2013. p. 523-528. Disponível em: <https://sol.sbc.org.br/index.php/sbsi/article/view/5718/5615>. Acesso em: 12 out. 2019.

POLLOCK, Rufus. **Frictionless Data Video**. 2017. Disponível em: <https://rufuspollock.com/2017/09/04/frictionless-data-video/>. Acesso em 25 nov. 2020.

PORTAL BRASILEIRO DE DADOS ABERTOS. **ONU lança guia de planejamento de dados abertos governamentais**. 2017. Disponível em: <https://dados.gov.br/noticia/onu-lanca-guia-de-planejamento-de-dados-abertos-governamentais>. Acesso em: 20 dez. 2018.

PORTAL BRASILEIRO DE DADOS ABERTOS. Sobre os dados.gov.br, [2019]. **Dados.gov.br**. Disponível em: <https://dados.gov.br/pagina/sobre>. Acesso em: 11 nov. 2020.

PORTER, Michael. E. Towards a dynamic theory of strategy. **Strategic Management Journal**, v. 12, p. 95-117, 1991. DOI: <https://doi.org/10.1002/smj.4250121008>. Disponível em: <https://onlinelibrary.wiley.com/doi/10.1002/smj.4250121008>. Acesso em: 12 jun. 2020.

QUARATI, Alfonso; DE MARTINO, Monica. Open government data usage: a brief overview. **ACM Digital Library**, v. 19, n. 28, p. 6-13, jun. 2019. DOI: <https://doi.org/10.1145/3331076.3331115>. Disponível em: <https://dl.acm.org/doi/10.1145/3331076.3331115>. Acesso em: 20 ago. 2019.

RAMÍREZ-ALUJAS, Álvaro; DASSEN, Nicolás. **Vientos de cambio**: el avance de las políticas de gobierno abierto en américa latina y el caribe. Washington: Banco Interamericano de Desarrollo, 2014. *E-book*. Disponível em:

<https://publications.iadb.org/publications/spanish/document/Vientos-de-cambio-El-avance-de-las-pol%C3%ADticas-de-gobierno-aberto-en-Am%C3%A9rica-Latina-y-el-Caribe.pdf>. Acesso em: 10 maio 2020.

RAMOS-SIMÓN, Luis Fernando. El uso de las licencias libres en los datos públicos abiertos. **Revista Espanola de Documentacion Cientifica**, v. 40, n. 3, p. 1-16, jun./set. 2017. DOI: <https://doi.org/10.3989/redc.2017.3.1376>. Disponível em: <http://redc.revistas.csic.es/index.php/redc/article/view/983/1515>. Acesso em: 20 out. 2019.

REN, Guang-Jie; GLISSLAMM, Susanne. Identifying information assets for open data: the role of business architecture and information quality. In: Proceedings of the 2012 IEEE 14th INTERNATIONAL CONFERENCE ON COMMERCE AND ENTERPRISE COMPUTING, 14., 2012, Hangzhou. **Anais** [...]. Hangzhou: CEC, 2012, p. 94-100. Disponível em: <https://ieeexplore.ieee.org/abstract/document/6470782>. Acesso em: 10 abr. 2020.

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

ROCHA, Eliana da Conceição; SOUZA, Márcia de Figueiredo Evaristo de. **Metodologia para avaliação de produtos e serviços de informação**. Brasília: Ibiti, 2010.

ROZADOS, Helen Beatriz Frota. A informação científica e tecnológica e os serviços de informação. **Informação & Sociedade**, João Pessoa, v. 16, n. 1, p. 49-62, jan./jun. 2006. Disponível em: https://brapci.inf.br/_repositorio/2010/11/pdf_3a5bcfea0f_0012842.pdf. Acesso em: 12 jun. 2020.

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. DOI: <http://dx.doi.org/10.1016/j.giq.2017.01.001>. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0740624X17300011>. Acesso em: 12 out. 2019.

SADIQ, Shazia; INDULSKA, Marta. Open data: quality over quantity. **International Journal of Information Management**, v. 37, n. 3, p. 150-154, 2017. DOI: <http://dx.doi.org/10.1016/j.ijinfomgt.2017.01.003>. Disponível em: https://www.researchgate.net/publication/312934839_Open_data_Quality_over_quantity. Acesso em: 12 out. 2019.

SANTARÉM SEGUNDO, José Eduardo. Web semântica: fluxo para publicação de dados abertos e ligados. **Informação em Pauta**, Fortaleza, v. 3, n. Especial, p. 117-140, nov. 2018. DOI: <https://doi.org/10.32810/25253468.ip.v3iEspecial.2018.39721.117-140>. Disponível em: <http://www.periodicos.ufc.br/informacaoempauta/article/view/39721>. Acesso em: 12 maio 2020.

SAXENA, Stuti. Open Government Data (OGD) usage in India: a conceptual framework using TOE & UTAUT frameworks. **Indexed Scholarly journal**, v. 4, n. 3, p. 124-144, set. 2016. Disponível em: <https://www.socratesjournal.com/index.php/SOCRATES/article/view/214>. Acesso em: 12 jun. 2019.

SECRETARIA DE GOVERNO DA PRESIDÊNCIA DA REPÚBLICA (Brasil). **Relatório nacional voluntário sobre os Objetivos de Desenvolvimento Sustentável**. Brasília : Presidência da República, 2017. Disponível em: <http://portalods.com.br/publicacoes/relatorio-nacional-voluntario-sobre-os-objetivos-de-desenvolvimento-sustentavel/>. Acesso em: 20 nov. 2019.

SHANKARANARAYAN, Ganesan; ZIAD, Mostapha; WANG, Richard. Managing data quality in dynamic decision environments: an information product approach. **Journal of Database Management**, v. 14, n. 4, p. 14-32, 2003. Disponível em: https://www.researchgate.net/publication/220373678_Managing_Data_Quality_in_Dynamic_Decision_Environments_An_Information_Product_Approach. Acesso em: 18 dez. 2019.

SILVA, Jonathas Luiz Carvalho. **Fundamentos da informação I**: perspectivas em Ciência da Informação. São Paulo: ABECIN, 2017. E-book. Disponível em: <http://www.repositoriobib.ufc.br/000042/00004231.pdf>. Acesso em: 17 mar. 2020.

STRONG, Diane; LEE, Yang; WANG, Richard. 10 Potholes in the road to information quality. **IEEE Computer**, v. 18, n. 162, p. 38-46, 1997. Disponível em: <http://mitiq.mit.edu/Documents/Publications/TDQMpub/10potholesIEEEComputerAug97.pdf>. Acesso em: 10 set. 2019.

TAMBOURIS, Efthimios. Multidimensional open government data. **Journal of eDemocracy and Open Government**, v. 8, n. 3, p. 1-11, 2016. DOI: <https://doi.org/10.29379/jedem.v8i3.444>. Disponível em: <https://www.jedem.org/index.php/jedem/article/view/444>. Acesso em: 28 ago. 2019.

TAUBERER, Joshua. **The principles, practices, and a history of the open government data movement**. Open Government Data: The Book. London: Second Editio, 2014. E-book. Disponível em: <https://opengovdata.io/>. Acesso em: 27 set. 2019.

TRINDADE, Ana Lucia Batista; OLIVEIRA, Mírian. Atributos para avaliação da qualidade da informação em sistemas de gestão do conhecimento. In: ENCONTRO DE ADMISTRAÇÃO DA INFORMAÇÃO, 1., 2007. Florianópolis. **Anais** [...]. Florianópolis: UFSC, 2007. p. 1-14. Disponível em: <http://www.anpad.org.br/admin/pdf/enadi318.pdf>. Acesso em: 10 fev. 2020.

UNESCO. **Sobre a Liberdade de Informação (FOI)**. 2017. Disponível em: <http://www.unesco.org/new/en/communication-and-information/freedom-of-expression/freedom-of-information/about/>. Acesso em: 12 jun. 2019.

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. DOI: <https://doi.org/10.1016/j.giq.2016.02.001>. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0740624X16300132>. Acesso em: 12 nov. 2019.

VILELA, Allyson Bruno Campos Barros. **OpenData processor**: uma ferramenta para a automatização do processo de extração e publicação de dados abertos. 2018. Dissertação (Mestrado em Engenharia de Software) – Instituto Metrópole Digital, Universidade Federal do Rio Grande do Norte, Natal, 2018. Disponível em: https://repositorio.ufrn.br/jspui/bitstream/123456789/26014/1/OpenDataprocessorferramenta_Vilela_2018.pdf. Acesso em: 12 jan. 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. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0306261918318130>. Acesso em: 20 maio 2020.

WILKINSON, Mark D. et al. The FAIR guiding principles for scientific data management and stewardship. **Scientific Data**, v. 3, p. 1-9, 2016. DOI: <https://doi.org/10.1038/sdata.2016.18>. Disponível em: https://www.researchgate.net/publication/298345883_The_FAIR_Guiding_Principles_for_scientific_data_management_and_stewardship. Acesso em: 10 maio 2020.

WIXOM, Barbara; TODD, Peter. A theoretical integration of user satisfaction and technology acceptance. **Information Systems Research**, v. 16, n. 1, p. 85-102, 2005. Disponível em: <https://www.jstor.org/stable/23015766?seq=1>. Acesso em: 18 dez. 2019.

WORLD WIDE WEB CONSORTIUM. **Dados abertos governamentais**. São Paulo: W3C Brasil, 2009. Disponível em: <https://www.w3c.br/GT/GrupoDadosAbertos>. Acesso em: 10 abr. 2019.

WORLD WIDE WEB CONSORTIUM. **Manual de dados abertos**. Paulo: W3C Brasil, 2011. Disponível em: <https://www.w3c.br/Noticias/W3CBrasilLancaManualDadosAbertosDesenvolvedores>. Acesso em: 10 abr. 2019.

ZUIDERWIJK, Anneke et al. Socio-technical Impediments of open data. **Electronic Journal of e-Government**, v. 10, n. 2, p. 156-172, 2012. Disponível http://pure.tudelft.nl/ws/portalfiles/portal/32002646/ejeg_volume10_issue2_article255.pdf. Acesso em: 10 set. 2019.

ZUIDERWIJK, Anneke; JANSSEN, Marijn; DAVIS, Chris. Innovation with open data: Essential elements of open data ecosystems. **Information Polity**, v. 19, n. 1, p. 17-33, 2014. Disponível em: <https://pdfs.semanticscholar.org/07bc/6a8a6ec78e4311b5b3bf513f1cba585b3d7a.pdf>. Acesso em: 10 set. 2019.

ZUIDERWIJK, Anneke; JANSSEN, Marijn; PARNIA, Armin. The complementarity of open data infrastructures: an analysis of functionalities. In: ACM INTERNATIONAL CONFERENCE PROCEEDING SERIES, 1., 2013. Austria. *Anais* [...]. Austria: SIDER, 2013, p. 166-171. Disponível em:
<https://dl.acm.org/doi/10.1145/2479724.2479749>. Acesso em: 26 set. 2019.

ZUIDERWIJK, Anneke; JEFFERY, Keith; JANSSEN, Marijn. The potential of metadata for linked open data and its value for users and publishers. **Journal of Democracy and Open Government**, v. 4, n. 2, p. 222-244, 2012. DOI:
<https://doi.org/10.29379/jedem.v4i2.138>. Disponível em:
https://www.researchgate.net/publication/262276781_The_Potential_of_Metadata_for_Linked_Open_Data_and_its_Value_for_Users_and_Publishers. Acesso em: 10 set. 2019.

5-STAR OPEN DATA. **5-Star deployment scheme for open data here**. 2012. Disponível em: <https://5stardata.info/en/>. Acesso em: 18 maio 2020.