



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE ARTES E COMUNICAÇÃO
DEPARTAMENTO DE CIÊNCIA DA INFORMAÇÃO
CURSO DE GRADUAÇÃO EM GESTÃO DA INFORMAÇÃO

GETÚLIO VALDEMIR BATISTA

DADOS ABERTOS GOVERNAMENTAIS: avaliação de maturidade e qualidade

Recife
2021

GETÚLIO VALDEMIR BATISTA

DADOS ABERTOS GOVERNAMENTAIS: avaliação de maturidade e qualidade

Trabalho de Conclusão de Curso apresentado ao Departamento de Ciência da Informação da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de bacharel em Gestão da Informação.

Orientadora: Profa. Dra. Sandra de Albuquerque Siebra

Recife

2021

Catálogo na fonte
Biblioteca Joaquim Cardozo – Centro de Artes e Comunicação

B333d Batista, Getúlio Valdemir

Dados abertos governamentais: avaliação de maturidade e qualidade / Getúlio Valdemir Batista. – Recife, 2021. 59f.: il.

Sob orientação de Sandra de Albuquerque Siebra. Trabalho de Conclusão de Curso (Graduação) – Universidade Federal de Pernambuco. Centro de Artes e Comunicação. Departamento de Ciência da Informação. Curso de Gestão da Informação, 2021.

Inclui referências.

1. Gestão da Informação. 2. Dados abertos governamentais. 3. Governo aberto. 4. Qualidade de dados. 5. Maturidade de abertura de dados. 6. Acesso à informação. I. Siebra, Sandra de Albuquerque (Orientação). II. Título.

020 CDD (22. ed.) UFPE (CAC 2021-243)



Serviço Público Federal
Universidade Federal de Pernambuco
Centro de Artes e Comunicação
Departamento de Ciência da Informação

FOLHA DE APROVAÇÃO

Dados abertos governamentais: avaliação de maturidade e qualidade

Getúlio Valdemir Batista

Trabalho de Conclusão de Curso submetido à Banca Examinadora, apresentado e aprovado de modo remoto (online), conforme autorizado pelo PROACAD/UFPE em Ata de Reunião Virtual dos Coordenadores de Graduação do dia 12 de Maio de 2020, pelo Curso de Gestão da Informação, do Departamento de Ciência da Informação, da Universidade Federal de Pernambuco, como requisito parcial para obtenção do título de Bacharel em Gestão da Informação.

TCC aprovado 16 de Dezembro de 2021.

Banca Examinadora:

Orientadora – Profa. Dra. Sandra de Albuquerque Siebra.
DCI/Universidade Federal de Pernambuco.

Examinadora 1 – Profa. Dra. Edilene Maria da Silva.
DCI/Universidade Federal de Pernambuco.

Examinadora 2 – MSc. Faysa de Maria Oliveira e Silva.
PPGCI/UFPE.

AGRADECIMENTOS

Primeiramente, quero agradecer à Deus, por ter chegado até aqui. Sem ele nada disso seria possível.

Aos meus pais, D. **Socorro** e Sr. **Batista** que não tiveram a oportunidade de ter acesso aos estudos, mas se desdobraram para que eu conseguisse ter. Agradeço o amor, o cuidado, o incentivo e o apoio incondicional. Essa conquista é de vocês!

Aos amigos que encontrei nessa trajetória, em especial à **Camila de Sá, Faysa de Oliveira e Rinaldo Ribeiro**, vocês foram essenciais para eu chegar até aqui.

A minha estimada orientadora, **Sandra Siebra**, que desde o início do curso me acompanha, obrigado pelas oportunidades que me foram dadas, pelas suas correções e incentivos.

A esta universidade, seu corpo docente, direção e administração que abriram a porta para que hoje eu vislumbrasse um horizonte maior.

RESUMO

A disponibilização dos denominados dados abertos governamentais fomenta a participação social e amplia a colaboração da sociedade junto às ações desenvolvidas por administradores públicos. Neste cenário, esta pesquisa teve por objetivo avaliar o grau de maturidade e a qualidade dos dados abertos disponibilizados pelo governo federal, com foco na análise dos dados relacionados à previdência social brasileira. Esta pesquisa caracteriza-se quanto aos fins como descritiva e qualitativa, e quanto aos meios como bibliográfica e documental, com realização de estudo de caso em uma amostra de conjuntos de dados da Previdência Social. Observou-se que dos 119 conjuntos de dados recuperados, apenas 13 conjuntos possuem atualização periódica; que a maior parte dos conjuntos é pobremente descrito, possuindo metadados limitados, inconsistentes e/ou incompletos, dificultando a compreensão dos conteúdos dos conjuntos e, conseqüentemente sua interpretação e uso. Também foi detectado que alguns conjuntos de dados disponibilizados são efetivamente inacessíveis, pois seus links de acesso levam à uma página de erro, não sendo possível finalizar o download. Os formatos em que os dados são disponibilizados, em geral, são apropriados e indicados para disponibilização dos dados para processamento por máquinas, porém são insuficientes em termos de facilidade de uso por parte de humanos, sendo a visualização desses dados considerada pobre e difícil, inclusive devido a inexistência ou falta de clareza da descrição dos diversos campos do conjunto de dados. Os conjuntos de dados analisados no estudo de caso, em sua maioria, atenderam apenas parcialmente aos critérios de análise, o que mostra que a qualidade e maturidade dos dados da previdência social ainda carecem de aprimoramento.

Palavras-chave: dados abertos governamentais; governo aberto; qualidade de dados; maturidade de abertura de dados; acesso à informação.

ABSTRACT

The availability of so-called open government data encourages social participation and expands society's collaboration with actions taken by public administrators. In this scenario, this research aimed to assess the degree of maturity and quality of open data made available by the federal government, focusing on the analysis of data related to Brazilian social security. This research is characterized in terms of purposes as descriptive and qualitative, and in terms of means such as bibliographical and documentary, with a case study carried out in a sample of data sets from Social Security. It was observed that of the 119 data sets retrieved, only 13 sets are periodically updated; that most of the sets are poorly described, having limited, inconsistent and/or incomplete metadata, making it difficult to understand the contents of the sets and, consequently, their interpretation and use. It was also detected that some datasets made available are effectively inaccessible, as their access links lead to an error page, making it impossible to complete the download. The formats in which the data are made available, in general, are appropriate and indicated for making the data available for processing by machines, but they are insufficient in terms of ease of use by humans, and the visualization of this data is considered poor and difficult, including due to the lack or lack of clarity in the description of the various fields in the dataset. The datasets analyzed in the case study, for the most part, only partially met the analysis criteria, which shows that the quality and maturity of social security data still need to be improved.

Keywords: open government data; open government; data quality; data open maturity; access to information.

LISTA DE FIGURAS

| | | |
|-------------|---|----|
| Figura 1 – | Principais Marcos relacionados ao Governo Eletrônico | 18 |
| Figura 2 – | Principais legislações brasileiras relacionadas à transparência pública e dados abertos | 20 |
| Figura 3 – | Modelo para avaliar o grau de estruturação dos dados de Berners-Lee | 31 |
| Figura 4 – | Classificação por estrelas do modelo proposto por Berners-Lee | 32 |
| Figura 5 – | Mapa Conceitual que resume os principais conceitos trabalhados na pesquisa | 41 |
| Figura 6 – | Percentual de Conjunto de Dados com Atualização Periódica | 42 |
| Figura 7 – | Ausência de PDA Válido da previdência social brasileira | 44 |
| Figura 8 – | Preenchimento não padronizado dos metadados | 45 |
| Figura 9 – | Preenchimento não padronizado dos metadados | 46 |
| Figura 10 – | Campo última atualização duplicado e inconsistente | 46 |
| Figura 11 – | Inconsistência nos metadados e atendimento à INDA | 47 |
| Figura 12 – | Inconsistência no preenchimento dos Metadados | 48 |
| Figura 13 – | Quantidade de conjuntos de dados disponíveis por formato | 48 |
| Figura 14 – | Alerta ao abrir o arquivo de dados | 49 |

LISTA DE QUADROS

| | | |
|-------------|--|----|
| Quadro 1 – | Conceito de Governo Aberto segundo Organismos Internacionais | 23 |
| Quadro 2 – | Os oito princípios para Dados Abertos Governamentais | 27 |
| Quadro 3 – | Previsão dos Princípios para Dados Abertos Governamentais na LAI | 28 |
| Quadro 4 – | Documentos que orientam a publicação de DAGs | 29 |
| Quadro 5 – | Benefícios para consumidores e publicadores de dados segundo o esquema cinco estrelas | 32 |
| Quadro 6 – | Síntese dos Princípios FAIR | 34 |
| Quadro 7 – | Documentos analisados na pesquisa | 35 |
| Quadro 8 – | Protocolo de avaliação da pesquisa | 38 |
| Quadro 9 – | Metadados dos conjuntos de dados da Previdência | 43 |
| Quadro 10 – | Sugestão de Metadados obrigatórios | 44 |
| Quadro 11 – | Resultado da avaliação dos conjuntos de dados com base na junção dos critérios de avaliação | 50 |

LISTA DE ABREVIATURAS E SIGLAS

| | |
|------|---|
| CGU | CONTROLADORIA-GERAL DA UNIÃO |
| CSV | COMMA-SEPARATED VALUES |
| DA | DADOS ABERTOS |
| DAG | DADOS ABERTOS GOVERNAMENTAIS |
| IBGE | INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA |
| INDA | INFRAESTRUTURA NACIONAL DE DADOS ABERTOS |
| JSON | JAVASCRIPT OBJECT NATATION |
| LAI | LEI DE ACESSO À INFORMAÇÃO |
| LGPD | LEI GERAL DE PROTEÇÃO DE DADOS PESSOAIS |
| LRF | LEI DE RESPONSABILIDADE FISCAL |
| PNAD | PESQUISA NACIONAL POR AMOSTRA DE DOMICÍLIOS |
| PDA | PLANO DE DADOS ABERTOS |
| RDF | RESOURCE DESCRIPTION FRAMEWORK |
| TIC | TECNOLOGIA DA INFORMAÇÃO E COMUNICAÇÃO |
| XML | EXTENSIBLE MARKUP LANGUAGE |

SUMÁRIO

| | | |
|------------|---|-----------|
| 1 | INTRODUÇÃO | 12 |
| 2 | FUNDAMENTAÇÃO TEÓRICA | 17 |
| 2.1 | Governo eletrônico | 17 |
| 2.2 | Transparência pública | 20 |
| 2.3 | Governo aberto | 22 |
| 2.4 | Dados abertos governamentais | 24 |
| 2.5 | Princípios ou diretrizes para dados abertos governamentais | 27 |
| 2.6 | Maturidade da abertura de dados | 30 |
| 2.7 | Princípios <i>FAIR</i> | 33 |
| 3 | PROCEDIMENTOS METODOLÓGICOS | 35 |
| 4 | RESULTADOS E DISCUSSÕES | 40 |
| 4.1 | Atualidade, uso de metadados, formato de dados e licenças de uso | 42 |
| 4.2 | Análise de qualidade dos dados e da maturidade de abertura dos dados | 50 |
| 5 | CONSIDERAÇÕES FINAIS | 53 |
| | REFERÊNCIAS | 55 |

1 INTRODUÇÃO

No decorrer do tempo, as práticas de democratização do acesso à dados e informações, no contexto de governo aberto¹, tem permeado a comunicação entre o Estado e a sociedade e vêm ganhando cada vez mais notoriedade na gestão pública, especialmente depois da Lei de Acesso à Informação - LAI (BRASIL, 2011). Nesse contexto, foram criadas várias iniciativas de Dados Abertos Governamentais (DAG), no contexto internacional e nacional (CHUN *et al.*, 2010). Segundo Attard *et al.* (2015), Lourenço (2015) e Linders (2013), essas iniciativas visam a reutilização dos dados, a transparência², a melhoria da participação dos cidadãos e da prestação pública de contas (*accountability*³), com a possibilidade de responsabilização dos agentes públicos e dos representantes eleitos por suas ações e decisões tomadas. Ressalta-se que, nesta pesquisa, adotou-se o conceito de Dado Aberto Governamental (DAG), definido por Galiotou e Fragkou (2013), que o define como dado aberto⁴ disponibilizado por órgãos públicos, que fornecem informações úteis para os cidadãos, instituições, organizações e empresas em suas transações com o setor público. Estes dados podem ser exportados, reutilizados e, inclusive, integrados em novos produtos, tais como sistemas, aplicativos ou outros serviços, ganhando valor social e econômico (LOURENÇO, 2015; UBALDI, 2013).

No que tange às iniciativas relacionadas aos DAG, destaca-se a disponibilização dos mesmos, que pode ser feita via solicitação ao setor público de direito (transparência passiva); ou disponibilizado em portais (transparência ativa) (BRASIL, 2011; CORRÊA, 2017). Sendo esta disponibilização uma prática iniciada com a ampliação dos canais de comunicação, a exemplo das ouvidorias e com a criação de portais de transparência e serviços de informação ao cidadão (SICs). De fato, os portais de transparência e as informações disponibilizadas no contexto da transparência ativa, pregada pela Lei nº 12.527 - Lei de Acesso à Informação (BRASIL, 2011), nos sites das instituições, aproximou a relação entre o governo e o cidadão. E estas práticas se tornaram um meio de divulgação, avaliação e controle das informações governamentais para o livre uso pela sociedade e uma forma de prestar contas das ações desenvolvidas pelo Estado, promovendo a *accountability*. O que endossa o afirmado por Dahl

¹ O termo Governo Aberto agrupa um conjunto de diversas iniciativas objetivando conferir maior transparência, *accountability* e permeabilidade ao Estado.

² A transparência pode ser entendida como o esclarecimento prestado ao cidadão pelo Estado, referente ao que ocorre na sua esfera de competência, disponibilizando informações públicas de vários tipos com presteza, rapidez e correteza (LOURENÇO, 2015).

³ Nesta pesquisa será utilizado doravante o termo em inglês *accountability*.

⁴ Dado Aberto (DA) que é aquele que pode ser livremente utilizado, reutilizado e redistribuído por qualquer pessoa que deve, no máximo, atribuir à fonte original ou fazer o seu compartilhamento utilizando a mesma licença original (ATTARD *et al.*, 2015).

(1997), que a liberdade de produção, circulação e acesso à informação são necessidades mínimas à concretização da efetiva oposição e participação dos cidadãos no processo político de manifestação das preferências.

No contexto desta pesquisa, a transparência ativa foi a trabalhada, atentando para os portais por meio dos quais os DAG são publicados na forma de conjuntos de dados (*datasets*), classificados, por vezes, em catálogos (categorias) de dados. De acordo com Attard *et al.* (2015), cada catálogo corresponde a uma lista de fontes de dados (*datasources*) e cada fonte corresponde a um conjunto de dados, constituído por um link para download de um arquivo de DAG, cujo conteúdo está estruturado em campos e valores. Diante de um cenário informacional cada vez mais democrático, a disponibilização dos DAG, fomenta a participação social e amplia a colaboração da sociedade junto às ações desenvolvidas por administradores públicos (CETIC, 2013; PARYCEK *et al.*, 2010; SHERIDAN *et al.*, 2010). Por isso, a forma que os DAG são disponibilizados é relevante e precisa atender a diversos princípios e recomendações, de forma a estes poderem ser efetivamente descobertos, usados e reusados.

Dessa forma, a disponibilização dos DAG abrange a publicação e disseminação de informações do setor público, publicizadas em formatos compreensíveis logicamente por computadores e por seres humanos, não proprietários e livres de licenças que restrinjam sua reutilização (WORLD WIDE WEB CONSORTIUM, 2009). Em outras palavras, estes dados devem ser aqueles que qualquer indivíduo pode utilizar, reutilizar e redistribuir livremente, tendo que, no máximo, atentar para creditar a sua autoria e fazer o compartilhamento deles sob a mesma licença (OPEN GOVERNMENT WORKING GROUP, 2007; PIRES, 2015). Adicionalmente, devem estar disponíveis e indexados na Web, em formato compreensível por máquina, para que se viabilize sua reprodução e reutilização (EAVES, 2010). Assim, os DAG podem viabilizar o acesso rápido a dados essenciais à gestão e facilitar a tomada de decisão. Além disso, eles têm a probabilidade de afetar positivamente a democracia, ao proporcionar uma maior transparência, facilitar o controle social e contribuir para a gestão pública. Isso devido a eles darem acesso às informações necessárias à participação e à realização do controle social.

Entretanto, do ponto de vista prático, a realidade que se apresenta é de uma grande massa de dados e informações públicos não compartilhados, não tratados e/ou não disseminados (FRANZOSI *et al.*, 2011). De fato, são variadas as formas como os distintos órgãos lidam com a publicização de dados abertos (DA) e, salvo algumas exceções, apesar de existirem orientações disponíveis, percebe-se a ausência de uma estratégia comum para dar

tratamento à questão. Logo, a produção e publicização de dados e informações é realizada, muitas vezes, de maneira isolada e não interoperável pelas mais diversas instituições, públicas e privadas. Além disso, há casos em que as informações não estão facilmente disponíveis para o uso pelos tomadores de decisão, tampouco para a população (FRANZOSI *et al.*, 2011). Adicionalmente, torna-se difícil utilizar responsavelmente dados sem informações que os descrevam tais como: de que forma e quando foram criados; onde e como foram gerados; detalhes sobre sua estrutura e significado; sobre como coletá-los, integrá-los e analisá-los. Por isso, além dos dados, é necessária uma descrição abrangente e significativa deles, o que pode ser feito por meio de metadados ⁵(SHANKARANARAYANAN; EVEN, 2006). Sem tudo isso, torna-se complexo para a sociedade civil realizar o pleno aproveitamento destes conjuntos de dados, pois alguns dos dados disponíveis para visualização ainda oferecem barreiras técnicas e/ou políticas para serem reutilizados pelos cidadãos para criação de novos projetos, aplicativos e serviços (OPEN KNOWLEDGE INTERNATIONAL, 2011).

Várias iniciativas foram desenvolvidas nos últimos anos para tentar sanar estes problemas identificados, tais como: os princípios dos dados governamentais abertos (OPEN GOVERNMENT WORKING GROUP, 2007); a Política de Dados Abertos do Poder Executivo Federal, (BRASIL, 2016); a Parceria para Governo Aberto (*Open Government Partnership* – OGP) (OGP, 2018)⁶; a criação da Infraestrutura Nacional de Dados Abertos (INDA, 2019)⁷; a tradução do Manual dos Dados Abertos: governo (OPEN KNOWLEDGE INTERNATIONAL, 2011); os dez princípios para melhorar a transparência federal (UBALDI, 2013), a Cartilha Técnica para Publicação de Dados Abertos no Brasil⁸; além dos princípios FAIR que estabelecem que os dados devem ser Encontráveis (*Findable*), Acessíveis (*Accessible*), Interoperáveis (*Interoperable*) e Reutilizáveis (*Reusable*), delineando as características, ferramentas, vocabulários e infraestruturas necessários para descoberta e reutilização de dados por terceiros (GO FAIR, 2018; KLEIN, 2017; WILKINSON *et al.*, 2016). Nesse cenário, faz-se os seguintes questionamentos:

⁵ Pode ser definido como dado ou informação sobre o dado. Normalmente, é utilizado para armazenar informações úteis à recuperação ou acesso à informação, devendo ser capaz de descrever ou servir de sumário para o conteúdo de determinada informação. (VICTORINO *et al.*, 2020)

⁶ A Parceria para Governo Aberto ou OGP (do inglês *Open Government Partnership*) é uma iniciativa internacional que pretende difundir e incentivar globalmente práticas governamentais relacionadas à transparência dos governos, ao acesso à informação pública e à participação social (OGP, 2018)

⁷ É um conjunto de padrões, tecnologias, procedimentos e mecanismos de controle necessários para atender às condições de disseminação e compartilhamento de dados e informações públicas no modelo de Dados Abertos (INDA, 2019)

⁸ Produzida pelo Ministério do Planejamento Orçamento e Gestão –MP, inspirada em uma norma técnica do governo do Chile. Disponível no endereço: <http://dados.gov.br/pagina/cartilha-publicacao-dados-abertos>. Acesso em: 13 dez. 2021.

1. Os DAGs da previdência social disponibilizados pelo governo federal seguem os critérios, princípios e padrões existentes para DAGs?

2. Qual é o nível de maturidade e qualidade dos DAGs da previdência social brasileira?

Diante dessa perspectiva, a pesquisa tem como objetivo geral: avaliar o grau de maturidade e a qualidade dos dados abertos governamentais disponibilizados pelo governo federal, com foco nos dados relacionados à previdência social. E tem como objetivos específicos:

- Caracterizar os dados abertos governamentais no contexto do Governo Aberto, identificando possibilidades de sua aplicação na promoção da transparência e da democracia e no fortalecimento da gestão pública;
- Mapear as legislações, padronizações recomendações e critérios de qualidade e maturidade atualmente existentes para dados abertos;
- Analisar a maturidade e qualidade de uma amostra de dados abertos relacionados à previdência social;
- Pontuar o papel do Gestor da Informação na garantia de maturidade e qualidade de dados abertos.

Dessa forma, este trabalho busca compreender a dinâmica de disponibilização dos DAGs com foco nos dados abertos da previdência social brasileira. Vale salientar que parte desta pesquisa se iniciou no ano de 2019, com o projeto de iniciação científica intitulado “Dados Abertos governamentais: Avaliação de Maturidade e Qualidade”. O contexto da previdência social brasileira foi escolhido como foco da pesquisa pelo fato de os dados deste segmento estarem no alvo de interesse da sociedade na época do início da pesquisa, visto que a chamada “reforma da previdência” estava sendo discutida, tornando esses dados de interesse social e com potencial para informar a sociedade a respeito da realidade em que se encontravam os principais conjuntos de dados do órgão. Como grande esforço já havia sido despendido neste conjunto de dados, optou-se por continuar a utilizá-los, pela familiaridade adquirida.

Esta pesquisa se faz relevante pois, como afirma Janssen (2011) e Klein (2017), há uma importância crescente da transparência para os governos e, neste cenário, os DAG são indispensáveis para a participação cidadã, elaboração de políticas públicas e prestação de

serviços, sendo valiosos para os cidadãos, organizações, instituições e empresas públicas ou privadas, na tomada de decisão e na criação de produtos e serviços inovadores. Logo, estes deveriam ser dados disponibilizados com qualidade. Por isso, as sistemáticas nacionais de avaliação da transparência, do grau de maturidade, da qualidade e do acesso e uso dos dados, em portais que disponibilizam DAG, vem sendo ampliadas e aprimoradas. Ressalta-se que, no contexto da Ciência da Informação, muito do que vem sendo estudado em relação à dados de pesquisa (SAYÃO; SALES, 2020) também pode ser aplicado aos DAG, tais como os princípios FAIR (GO FAIR, 2018).

Como contribuição espera-se tanto delinear uma sistemática de avaliação de maturidade e qualidade que poderá ser aplicada à conjuntos de dados (*datasets*), como oferecer um cenário de como se encontram os DAG relacionados à previdência social, verificando se os problemas identificados por Franzosi *et al.* (2011) ainda persistem no contexto dos DAG brasileiros.

Continuando a escrita deste trabalho, a seção 2 apresenta a fundamentação teórica deste trabalho e busca apresentar a compreensão dos seis temas que serviram de base para o desenvolvimento da pesquisa: Governo eletrônico, Transparência pública, Governo Aberto, Dados Abertos Governamentais, Maturidade e Qualidade de Dados e Princípios FAIR. Na seção 3 são apresentados os procedimentos metodológicos utilizados no desenvolvimento da pesquisa. Na seção 4, são apresentados e discutidos os resultados da pesquisa, a partir da análise dos dados coletados. A seção 5 aponta algumas considerações finais. E, por fim, constam as referências bibliográficas que forma utilizadas para o embasamento teórico da pesquisa.

2 FUNDAMENTAÇÃO TEÓRICA

Esta seção abrange a fundamentação necessária para a compreensão e realização deste trabalho, englobando nas subseções, as temáticas governo eletrônico, transparência pública, governo aberto, dados abertos governamentais, diretrizes para dados abertos governamentais, maturidade da abertura de dados e princípios *FAIR*.

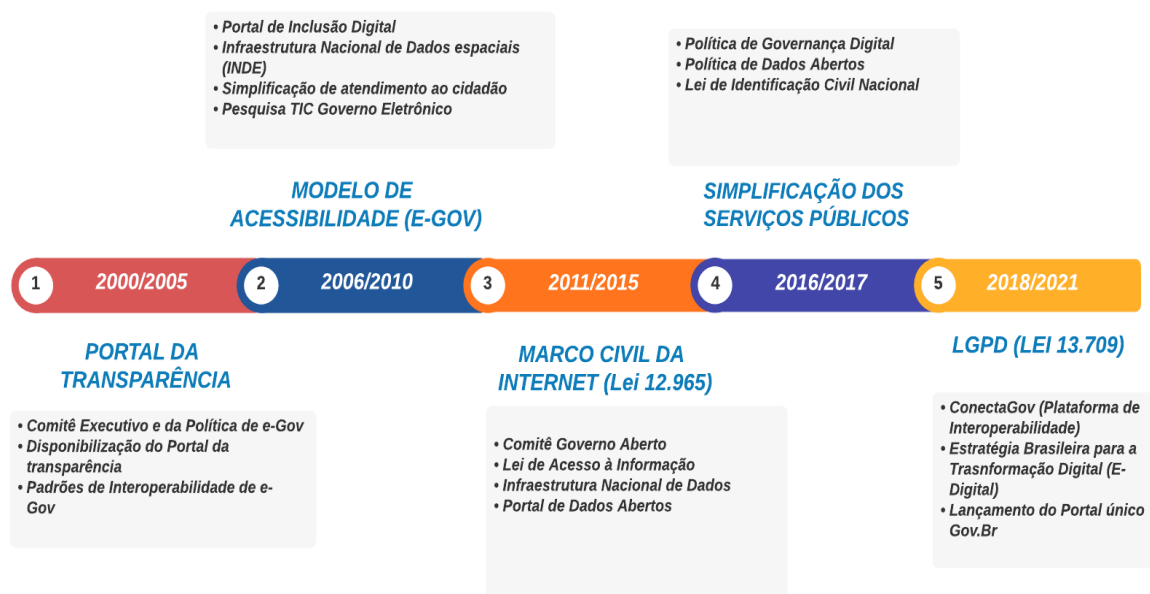
2.1 Governo Eletrônico

A evolução e popularização das tecnologias da informação e comunicação (TIC) possibilitaram mudanças significativas tanto nas telecomunicações, quanto na forma de intercâmbio de informações, modificando os processos produtivos dos mais variados setores da sociedade e facilitando as formas de disseminação de informação e comunicação entre pessoas, pessoas e empresas, pessoas e governo (em ambos os sentidos). Isso fez surgir a necessidade de inovação e adaptação de todas as organizações e instituições, inclusive do governo.

A introdução da tecnologia digital na administração pública não é recente, remota as décadas de 1960 e 1970, quando o governo passou pela informatização de seus processos administrativos e de seu cotidiano de trabalho, principalmente com o objetivo de tornar a gestão pública mais eficaz, ágil e flexível. E a partir dos anos 2000, começou a haver a implantação do governo eletrônico na administração pública brasileira, que se tornou uma importante ferramenta, capaz de fortalecer os processos produtivos dos mais diversos órgãos e entidades do governo, possibilitando automatizar, facilitar e diminuir os erros na rotina diária.

Pimenta e Canabarro (2014) definem governo eletrônico como a efetiva oferta e entrega de serviços, nos diferentes canais utilizados pela administração pública (e.g. portais *web*, *e-mails*, redes sociais, *e-sic*, etc.), junto com a aproximação e interação entre o Estado e o cidadão por meio desses canais. Alguns dos principais marcos da adoção do governo eletrônico na administração pública brasileira são destacados na linha do tempo representada na Figura 1.

Figura 1 – Principais Marcos relacionados ao Governo Eletrônico



Fonte: Baseado em Gov.br (2021).

Em busca da eficiência dos processos de trabalhos e da prestação de serviços públicos, o Estado brasileiro passa a guiar suas ações com foco na desburocratização, modernização e melhoria no acesso à informação pública. Nesse sentido, as primeiras iniciativas de governo eletrônico, no início dos anos 2000, são os sites do governo, que geralmente se limitavam a fornecer informações sobre gestão, informar os cidadãos sobre os planos do governo e informar sobre seus direitos e serviços ofertados pela administração pública (BRASIL, 2019). Almarabeh e AbuAli (2010), delinearão ações esperadas do governo eletrônico, tais como que ele: (1) fizesse uso de redes internas e configurasse sistemas de e-mail; (2) permitisse que organizações governamentais e o público acessassem informações; (3) estabelecesse comunicação bidirecional entre governo e os cidadãos; (4) valorizasse o intercâmbio; (5) implementasse a democracia digital; e (6) criasse um governo em conjunto com os cidadãos.

Nos anos posteriores, além da implementação do portal da transparência, outras políticas e serviços foram incrementados à rotina da administração pública, sobretudo a prestação de atendimento ao cidadão por meio da rede mundial de computadores, reduzindo ou eliminando a necessidade de presença física. Para Prado (2009), as TIC possibilitaram o governo eletrônico, que permite uma maior qualidade nas relações entre o governo e o cidadão em prol da construção de uma democracia mais participativa. Nas palavras do autor o governo eletrônico possibilita “o uso das aplicações das TICs para melhoria dos processos da administração pública, aumentando a eficiência, buscando melhorar a governança, auxiliando

a elaboração e o monitoramento das políticas públicas e integrando os diversos níveis de governo, dentre outros” (PRADO, 2019, p. 32).

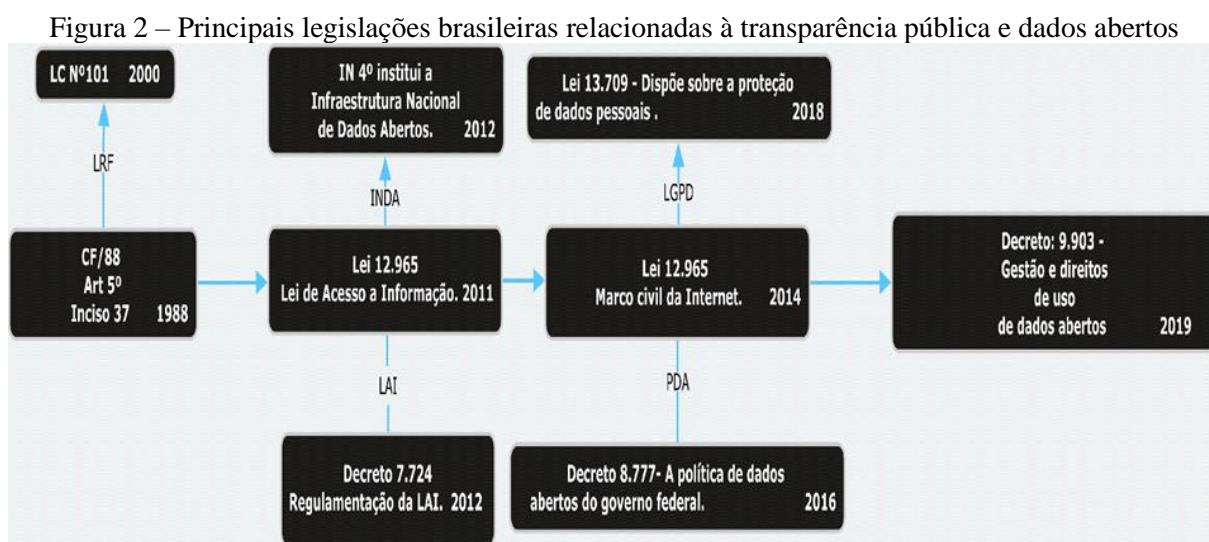
Nos últimos 10 anos, o aprimoramento das tecnologias móveis, como smartphones e tablets, trouxe novas possibilidades para o governo eletrônico. Assim, a partir de 2019, a transformação digital passou a ocorrer em um ritmo mais acelerado e mais de 500 serviços públicos digitais foram entregues à população brasileira (BRASIL, 2019). É importante enfatizar que isso foi possível graças a popularização dos dispositivos móveis, pois foram posicionados como o principal meio de usar e acessar o ambiente digital (especialmente a Internet), conforme destacado pela Pesquisa Nacional por Amostra de Domicílios (PNAD) de 2019, realizada pelo IBGE. De acordo com o levantamento, o telefone móvel vem sendo a principal ferramenta utilizada pelos brasileiros para acessar os serviços ofertados (BRASIL, 2021).

Além do crescimento (número e importância) dos dispositivos de comunicação móvel, a gestão pública também introduziu ferramentas que permitem a produção, armazenamento e compartilhamento de grandes quantidades de informação, além do processamento de dados. Segundo Possamai (2016), os dados e as informações públicas são particularmente adequados para: o controle do território e da população e como subsídio para a tomada de decisões no setor governamental, bem como para promover a transparência, o controle social e a participação dos cidadãos. Portanto, considera-se que a informação governamental aberta representa uma série de conquistas democráticas desejáveis. É notório que a informatização dos serviços prestados pela administração pública levou à implementação de práticas que ajudam no processo de ampliar o acesso à informação, o debate e a participação popular na construção de políticas públicas. As ferramentas de governo eletrônico favorecem a transparência na gestão pública e implicam, segundo explicam Culau e Fortis (2006), na disseminação periódica de informações pormenorizadas através de relatórios, da realização de audiências públicas e da prestação de contas à sociedade. Adicionalmente, destacam os autores, quando se amplia o acesso às informações sobre a gestão pública se fortalece a ligação entre o Estado e a sociedade civil. E, por meio da interação com os cidadãos, é possível verificar suas necessidades e os rumos mais adequados que as políticas públicas devem seguir.

2.2 Transparência pública

As discussões em torno do tema transparência pública ocupam espaço na agenda nacional e internacional (RAUPP; PINHO 2013). No Brasil, essas discussões decorrem de iniciativas da sociedade civil, instituições e organizações não-governamentais, que se sentem inconformadas com os sucessivos casos de corrupção; com a má gestão dos recursos públicos; e com a cultura de que a informação pertence à instituição, órgão ou agente público e não ao cidadão. Ao distanciar a informação do cidadão, o gestor fortalece seu poder e confirma o seu autoritarismo. Logo, a transparência é a forma de evitar esse tipo de conduta (FIGUEIREDO; SANTOS, 2013).

No processo de redemocratização, na década de 80, a sociedade passou a ter reconhecido o direito de acesso à informação, previsto na carta magna de 1988 (BRASIL, 1988) e, nos anos posteriores, há a construção de uma agenda pautada na transparência como regra e no sigilo como exceção. A transparência está associada à divulgação de informações que permitem que sejam averiguadas as ações dos gestores e a consequente responsabilização por seus atos (FIGUEIREDO; SANTOS, 2013). A garantia da transparência é um desafio, mas, também, é um pilar necessário à administração pública. Assim, desde a Constituição de 1988 (BRASIL, 1988), o país avançou na promoção de um conjunto de normas que dão respaldo à promoção da transparência (SANTANA; SILVA; SILVA, 2022), como pode ser visualizado na Figura 2.



Fonte: O autor (2021).

Além do previsto na Constituição Federal acerca da disponibilização de informações públicas, outros dispositivos legais surgiram com o passar dos anos. Por exemplo, a Lei de Responsabilidade Fiscal de 2000 (Lei Complementar nº 101, de 2000) (BRASIL, 2000), que

dispõe sobre a disponibilização, em tempo real, de informações sobre as receitas e despesas da União, Estados e Municípios. Isso corrobora com o compromisso da República Federativa do Brasil em fomentar um ambiente transparente, cujo objetivo é alcançar a participação da sociedade no desenvolvimento do Estado, possibilitando a fiscalização do governo, o empoderamento do cidadão e a participação da sociedade no desenvolvimento de soluções inovadoras para a administração pública (OGP, 2011).

Na sequência, em 2011, é aprovada a Lei de Acesso à Informação (LAI), Lei Federal nº 12.527, em novembro de 2011, que dispõe sobre os procedimentos a serem observados pela União, Estados, Distrito Federal e Municípios, com a finalidade de garantir o acesso à informação previsto no inciso XXXIII do art. 5º, no inciso II do § 3º do art. 37 e no § 2º do art. 216 da Constituição Federal (BRASIL, 2011). Nesse sentido, a transparência passa a ser fortemente associada às boas práticas de gestão por meio da disponibilização de informações públicas. A partir dela, torna-se possível saber o universo de ações governamentais de forma detalhada, assim a publicação de dados e informações em formato aberto ou sob uma licença aberta, atualizada, completa e com a presença de bons metadados pode contribuir para o controle das ações do estado, resolução de problemas, a inovação e participação social. Logo, é papel do estado permitir que informações sob sua guarda, que não sejam classificadas como sigilosas, sejam disponibilizadas. Reforçando a ideia de que a regra é a publicação, como previsto desde a Constituição Brasileira, por meio do Artigo 5º, Inciso XXXIII:

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, art. 5, inc. XXXIII).

A disponibilização de dados e informações no setor público pode acontecer de duas formas: por meio da transparência ativa ou passiva. A primeira diz respeito à disponibilização espontânea, sistemática e periódica de dados e informações sobre a administração, independente de solicitações. A segunda diz respeito à disponibilização via solicitação do cidadão. Nesta última, a administração pública é acionada por meio dos sistemas de informação ao cidadão (E-SIC, FalaBR). Ressalta-se que, independentemente da forma de disponibilização, a transparência é o cerne para o exercício da *accountability* entre Estado e cidadãos. Assim, de acordo com Pinho e Raupp (2013), os gestores e agentes públicos devem enfrentar e superar a opacidade que historicamente reveste a atuação do estado, o que exige mudanças na cultura e um compromisso com a transparência.

Dando continuidade ao desenvolvimento de uma política de disseminação de dados informações governamentais para o livre uso pela sociedade, o antigo Ministério do Planejamento, Orçamento e Gestão disponibilizou a Cartilha técnica para Publicação de DAGs no Brasil. Este documento é inspirado em uma norma técnica do governo do Chile para publicação de dados (BRASIL 2020). Posteriormente surgiu a Política de Dados Abertos do Poder Executivo Federal, cuja responsabilidade pela gestão e monitoramento está a cargo da Controladoria-Geral da União (CGU), por meio da Infraestrutura Nacional de Dados Abertos (INDA). Por meio da INDA é possível identificar padrões, tecnologias e orientações para disseminação e compartilhamento de dados em formato aberto (BRASIL, 2021). O principal instrumento que regula essa política é o decreto Nº 8.777/2016 (BRASIL, 2016), que dispõe sobre a publicação de dados contidos em bases de dados de órgãos e/ou entidades da administração pública direta, autárquica e fundacional. A partir deste instrumento normativo surgiu o Plano de Dados Abertos (PDA) que é um instrumento que orienta os órgãos da administração pública do Poder Executivo Federal a implementar e promover a abertura de dados, contribuindo para um governo aberto.

2.3 Governo aberto

As iniciativas de abertura do governo surgem em meio a um movimento internacional, iniciado na segunda metade do século XXVIII, ganhando força em 2009, quando foi publicada a Declaração de Governo Aberto (OGP, 2018). Os oito países fundadores, entre eles o Brasil, reconheceram as exigências da sociedade por mais participação e abertura nos governos, defendendo que a participação popular possibilita a melhoria tanto nos serviços, quanto na gestão responsável dos recursos públicos (SANCHEZ, 2020). Assim, em setembro de 2011, ocorreu a fundação da *Open Government Partnership* (OGP), nascida com o intuito de promover uma parceria voltada ao comprometimento com transparência, empoderamento do cidadão, combate à corrupção e desenvolvimento de novas tecnologias (SANCHEZ, 2020). No Brasil, o direito de acesso à informação previsto na Constituição Federal (BRASIL, 1988) e foi instituído em 2011, por meio da Lei de Acesso à Informação (Lei nº 12.527 / 2011), a LAI (BRASIL, 2011), que estabelece padrões de transparência pública nos governos federal, estadual e municipal. Portanto, no início, a ideia de governo aberto estava principalmente relacionada à geração de transparência governamental, e essa expressão foi usada como sinônimo de governo que proporcionava acesso à informação, que tradicionalmente era classificada como restrita. Na verdade, essa expressão em si parece ter

sido usada pela primeira vez na década de 1950, justamente durante o debate anterior à aprovação da Lei de Liberdade de Informação (WEINSTEIN; GOLDSTEIN, 2012). Ainda hoje, para alguns autores, a transparência ainda é a característica mais importante (LORENZO; CÉSAR, 2010) e condição básica para outras exigências do chamado governo aberto (POSSAMAI, 2016).

No Quadro 1 são apresentados alguns conceitos para Governo Aberto.

Quadro 1 – Conceito de Governo Aberto segundo Organismos Internacionais

| ORGANISMO | CONCEITO | ENFOQUE |
|---|---|---|
| Parceria para Governo Aberto - OGP | Governo que promove transparência, empoderamento da sociedade, combate à corrupção e incorpora novas tecnologias para melhoria da governança. (OGP, 2018). | Transparência, empoderamento, sociedade, combate à corrupção e adoção de TICs pelo governo. |
| Banco Interamericano de Desenvolvimento - BID | Não apresenta um consenso claro, mas governo aberto se insere na área de modernização/reforma do Estado, cujo foco é na abertura (<i>openness</i>) e <i>accountability</i> ” (BID, 2016). | Transparência, inovação e <i>accountability</i> |

Fonte: Baseado em Bellix, Guimarães e Machado (2017).

Destaca-se que, embora a terminologia contemporânea possa remontar a menos de uma década, dois dos três elementos básicos do governo aberto não são novos. Na verdade, os temas da transparência e participação social vêm sendo discutidos há muito tempo. Por exemplo, no século XXVIII, a ideia de que os cidadãos deveriam acessar as informações do governo já era discutida e, posteriormente se tornou um movimento difundido mundialmente.

Fica evidente, no enfoque das definições apresentadas no Quadro 1, que o tema transparência e seus derivados, como acesso à informação, participação social, inovação e prestação de contas (*accountability*) merecem destaque no contexto do governo aberto. Nesse sentido, no Brasil, conforme a Controladoria Geral da União (BRASIL, 2020), esses pontos convergentes acabaram por tornar-se os princípios do governo aberto.

1- Transparência

As informações sobre as atividades de governo são abertas, compreensíveis, tempestivas livremente acessíveis e atendem ao padrão básico de dados abertos.

2- Prestação de Contas e Responsabilização

Regras e mecanismos que estabelecem como os atores justificar suas ações, atuam sobre críticas e exigências e aceitam as responsabilidades que lhe são incumbidas

3- Participação cidadã

O governo procurar mobilizar a sociedade para debater, colaborar e propor contribuições que levam a um governo mais efetivo e responsivo.

4- Tecnologia e Inovação

O governo reconhece a importância das TIC's no fomento à inovação, ampliando a capacidade da sociedade de utilizá-la.

Assim, para o Brasil, um governo é considerado aberto se sua gestão, iniciativas e projetos e programas refletem esses quatro princípios (BRASIL, 2020). Neste contexto, a ideia base é que os dados e informações das instituições possam ser acessados, usados e reutilizados livremente pela sociedade. O que já era endossado por Agune, Gregório Filho e Bolliger (2010), quando afirmavam que o governo aberto deve primar pela:

Disponibilização, através da Internet, de informações e dados governamentais de domínio público para a livre utilização pela sociedade. É parte integrante do conceito que, à sociedade, seja garantido acesso aos dados primários, de forma que o interessado possa combiná-los, cruzá-los e, enfim, produzir novas informações e aplicações, colaborando com o governo na geração de conhecimento social a partir das bases governamentais. Este sistema segue o princípio que assegura o direito de acesso irrestrito aos documentos, dados e informações no âmbito da administração pública, salvo nas hipóteses de sigilo previstas na Constituição e legislação em vigor, e ainda ressalvadas as exceções disciplinadas por regulamentação específica do órgão ou entidade da administração pública. (AGUNE; GREGORIO FILHO; BOLLIGER, 2010, p.3)

Entre as informações e dados que a administração pública deve disponibilizar estão os dados abertos governamentais (DAG), que serão definidos, a seguir.

2.4 Dados abertos governamentais

O termo Dados Abertos Governamentais (DAG) pode ser definido, como dados abertos que qualquer pessoa pode livremente usar, reutilizar e redistribuir, estando sujeito a, no máximo, à exigência de creditar a sua autoria e compartilhá-los via a mesma licença.

(OPEN KNOWLEDGE FOUNDATION, 2014). O movimento em prol dos dados abertos foi discutido após uma reunião de líderes governamentais, programadores e ativistas de Dados Abertos em 2007 e começou a ganhar atenção pública, resultando no acesso irrestrito a um grande volume de conjuntos de dados, que antes não eram acessíveis. Por meio de entrevistas com funcionários públicos e pesquisadores do Canadá, do Reino Unido e dos Estados Unidos foi possível perceber que tais iniciativas tornaram-se um dos principais elementos relacionados às recentes reformas da administração pública, naqueles países, destinadas a aumentar a transparência e melhorar a prestação de serviços públicos (CLARKE; MARGETTS, 2014). O que vai ao encontro da maior parte da literatura sobre dados abertos governamentais, que os define, principalmente, como promotores de mais transparência pública e *accountability* (GEIGER; LUCKE, 2012; JANSSEN, 2011; LOURENÇO, 2015) e de envolvimento e participação dos cidadãos nos negócios públicos (CONRADIE; CHOENNI, 2014).

As instituições públicas são os maiores produtores, coletores e depositantes de dados públicos em diversos campos. Porém, cabe ressaltar que a obrigatoriedade oriunda das legislações que orientam a questão dos dados e das informações públicas não é suficiente para alcançar a filosofia por trás do movimento de dados abertos. A partir de então, surgem propostas de definição sobre o que são os DAG que, em sua maioria, possuem como ponto de convergência as ideias relacionadas às atividades de uso e reuso dos dados por qualquer pessoa, envolvendo o universo das informações de propriedade pública que devem ser disponibilizadas para a sociedade visando à participação e colaboração de todos. Inclusive, no contexto do governo, os dados abertos criam oportunidades para indivíduos, empresas do setor privado ou organizações sem fins lucrativos apresentarem novas ideias e criarem produtos com base nesses dados, tais como jogos, aplicativos, portais, plataforma, entre outros (BERTOT *et al*, 2014).

Porém, apesar do Brasil prever e aprovar leis de acesso à informação, a sociedade enfrenta dificuldades de efetivamente concretizar esse direito. Pois a realidade que se apresenta é a de que os órgãos do governo acabam divulgando apenas as informações exigidas por lei e, na maioria dos casos, não as mantém atualizadas, nem colaboram com a compreensão das mesmas por meio do fornecimento de metadados e formatos adequados, de forma a atender aos requisitos necessários para promover, acesso, uso e reuso. De fato, os DAG devem ser disponíveis em formatos compreensíveis logicamente por computadores e por seres humanos, devem ser não proprietários e livres de licenças que restrinjam sua reutilização (WORLD WIDE WEB CONSORTIUM, 2009). Apenas dessa forma, os DAG

podem viabilizar o acesso rápido à dados essenciais à gestão e facilitar a tomada de decisão, além de terem a probabilidade de afetar positivamente a democracia, ao proporcionar uma maior transparência, facilitar o controle social e contribuir para a gestão pública.

Porém ressalta-se que a demanda por informação adequada e transparente, clara e congruente, trouxe novos desafios para a administração pública, para estar alinhada com o planejamento estratégico do poder executivo, que menciona a garantia no “acesso à informação de forma transparente e tempestiva e a comunicação efetiva com a sociedade.” (BRASIL, 2016, p. 3). Pois, observa-se que, na prática, são variadas as formas como os órgãos lidam com a disponibilização de dados abertos e, salvo algumas exceções, apesar de existirem orientações disponíveis, percebe-se a ausência de uma estratégia comum para dar tratamento à questão. Logo, a produção e publicação de dados e informações são realizadas, muitas vezes, de maneira isolada, confusa e/ou não interoperável pelas diversas instituições, públicas e privadas. Além disso, há casos em que as informações não estão facilmente disponíveis para o uso pelos tomadores de decisão, tampouco para a população (FRANZOSI *et al.*, 2011). O que é endossado pela Escola de Políticas Públicas - EPP (2016) quando afirma que grande parte dos dados governamentais

Não está disponível ao público. E tampouco esses dados estão estruturados para facilitar sua compreensão mesmo por aqueles que podem acessá-los e manipulá-los. Como resultado, a extração de informações e produção de conhecimentos, que poderiam ser úteis para a sociedade, não acontece com a agilidade e eficácia necessárias para lidar com questões sociais e econômicas do século 21 (EPP, 2016).

Assim, torna-se necessário estabelecer padrões de qualidade, delimitando objetivos, diretrizes e metas, que garantam que o dado seja acessível, interoperável, encontrável e possa ser reutilizado por qualquer pessoa. Ainda assim, não basta apenas disponibilizar, é preciso dar acesso aos conjuntos de dados, de modo que o usuário consiga extrair o máximo de informação e compreender com o que está lidando. E assim, estando na Internet, os dados abertos devem:

- (A) estar disponibilizados em formato adequado e passível de ser trabalhado, de acordo com a conveniência de cada usuário;
- (B) possuir a qualidade de serem reutilizáveis e redistribuídos, se necessário;
- (C) estar inseridos em bases que interajam umas com as outras, no que diz respeito à tecnologia de configuração;
- (D) garantir acesso universal, permitindo uso por qualquer pessoa; e

(E) ser livres de restrições ou discriminações.

É nesse aspecto que o desenvolvimento dos denominados padrões Web se apresentam necessários para que estes conjuntos de dados possam ser apresentados em um formato, ou por meio de mecanismos, que estejam ao alcance da população e que possam ser acessados tanto por máquinas quanto por pessoas. Logo, as organizações públicas, não obstante o conhecimento e tecnologia de que disponham internamente, e o imenso acervo de dados e informações de seus bancos de dados, precisam identificar e perceber que é preciso fortalecer o contexto dos dados abertos, como uma forma de atender necessidades dos cidadãos. Porém, esses dados não podem, nem devem ser disponibilizados de qualquer maneira, requerendo que sejam seguidas diretrizes, como será discutido na seção, a seguir.

2.5 Princípios ou diretrizes para dados abertos governamentais

A disponibilização dos DAG deve ser baseada em diretrizes que orientam a forma como eles devem ser publicados. Essas diretrizes são frutos de discussões iniciadas desde 2007, quando um grupo de especialistas do *Open Government Data* definiram os oito princípios para dados abertos, e, posteriormente, definiram as três “Leis” dos Dados Abertos. Os oito princípios para DAG são aceitos em todo o mundo e um dado é considerado aberto quando é publicado segundo tais princípios, que estão definidos no Quadro 2 (DAVIES; BAWA, 2012; OPEN GOVERNMENT WORKING GROUP, 2015; UBALDI, 2013).

Quadro 2 – Os oito princípios para Dados Abertos Governamentais

| PRINCÍPIO | DEFINIÇÃO |
|-----------------------------------|--|
| COMPLETOS | Todos os dados públicos devem ser disponibilizados. Dado público é aquele que não está sujeito a restrições de privacidade, segurança ou outros privilégios; |
| PRIMÁRIOS | O dado é apresentado tal como colhido da fonte, com o maior nível possível de granularidade, sem agregação ou modificação. Por exemplo, um gráfico não é fornecido aberto, mas os dados utilizados para construir a planilha que deu origem a ele podem ser abertos; |
| ATUAIS | O dado deve ser publicado o mais rápido possível, para preservar seu valor. Em geral, têm periodicidade, quanto mais recentes e atuais, mais úteis para seus usuários; |
| ACESSÍVEIS | Os dados são disponibilizados para a maior quantidade possível de pessoas, atendendo, assim, aos mais diferentes propósitos; |
| COMPREENSÍVEIS POR MÁQUINA | Devem estar estruturados de modo razoável, possibilitando que sejam processados automaticamente. Por exemplo, uma tabela em PDF é muito bem compreendida por pessoas, mas para um computador é |

| | |
|-----------------------------|---|
| | apenas uma imagem, já uma tabela em formato estruturado, como CSV ou XML, é processada facilmente por softwares e sistemas; |
| NÃO DISCRIMINATÓRIOS | Devem estar disponíveis para qualquer pessoa, sem necessidade de cadastro ou qualquer outro procedimento que impeça o acesso; |
| NÃO PROPRIETÁRIOS | Nenhuma entidade ou organização deve ter controle exclusivo sobre os dados disponibilizados; |
| LIVRE DE LICENÇAS | Não devem estar submetidos a copyrights, patentes, marcas registradas ou regulações de segredo industrial. |

Fonte: Baseado em Davies e Bawa (2012); Open Government Working Group (2015); Ubaldi (2013).

Por meio desses princípios é possível estabelecer a maneira como as instituições poderão disponibilizar seus dados à sociedade. Como focam em possibilitar o acesso à informação, é possível estabelecer uma relação entre eles e a LAI (BRASIL, 2011), que orienta como dados e informações devem ser disponibilizados à sociedade. Dois dos oito princípios não foram identificados na LAI, são eles: Completos e Livre de Licenças. Com relação ao primeiro, não foi encontrado elementos na LAI que explicitem a necessidade de completude das informações. Com relação ao segundo princípio, não foi indicado na LAI se o dado/informação deve ou não estar Livre de Licenças. Os outros princípios estão previstos na LAI, conforme pode ser visto no Quadro 3.

Quadro 3 – Previsão dos Princípios para Dados Abertos Governamentais na LAI

| PRINCÍPIO | PREVISÃO NA LAI |
|-----------------------------------|---|
| COMPLETOS | Não foram encontrados elementos que explicitem a necessidade de completude das informações. |
| PRIMÁRIOS | Está previsto no Art. 7º inciso IV – o direito de obter informação primária , íntegra, autêntica e atualizada; [...] |
| ATUAIS | Está previsto no Art. 7º inciso IV – o direito de obter informação primária, íntegra, autêntica e atualizada ; [...] No Art. 8º inciso VI – Manter atualizadas as informações disponíveis para acesso. [...] |
| ACESSÍVEIS | No Art. 6º cabe aos órgãos e entidades do poder público, assegurar a: I- Gestão transparente da informação, propiciando amplo acesso a ela e sua divulgação. |
| COMPREENSÍVEIS POR MÁQUINA | No Art. 8º é dever dos órgãos e entidades promover a divulgação em local de fácil acesso [...] § 2º para cumprimento do disposto no caput, os órgãos e entidades deverão divulgar em sítios oficiais da rede mundial de computadores (internet). Esses sites deverão atender, entre outros, aos seguintes requisitos: [...] III – Possibilitar o acesso automatizado por sistemas externos em formatos abertos, estruturados e legíveis por máquina; [...] |
| | Art. 10. Qualquer interessado poderá apresentar pedido de acesso a |

| | |
|-----------------------------|---|
| NÃO DISCRIMINATÓRIOS | informação. § 1º Para o acesso a informações de interesse público, a identificação do requerente não pode conter exigências que inviabilizem a solicitação. |
| NÃO PROPRIETÁRIOS | Art. 8º é dever dos órgãos e entidades promover a divulgação em local de fácil acesso [...] § 2º para cumprimento do disposto no caput, os órgãos e entidades deverão divulgar em sítios oficiais da rede mundial de computadores (internet). Esses sites deverão atender, entre outros, aos seguintes requisitos: [...] II- Possibilitar a gravação de relatórios em diversos formatos eletrônicos, inclusive abertos e não proprietários, tais como planilhas e texto, de modo a facilitar a análise das informações. III- Possibilitar o acesso automatizado por sistemas externos em formatos abertos, estruturados e legíveis por máquina; [...] |
| LIVRE DE LICENÇAS | Não foi indicado se o dado/informação deve ou não estar Livre de Licenças. |

Fonte: O autor (2021).

Os pontos convergentes entre os oito princípios para DAGs e a LAI corroboram para a definição dos procedimentos necessários para disponibilização de dados e informações contidos em registros e/ou documentos produzidos e/ou acumulados pelos órgãos públicos. Porém, os dados precisam obedecer a padrões mínimos de qualidade, para que seja possível entendê-los e reutilizá-los. Nesse sentido, o decreto Nº 8.777/2016 (BRASIL, 2016) dispõe sobre a publicação de dados contidos em bases de dados de órgãos e/ou entidades da administração pública direta, autárquica e fundacional.

A partir do decreto, surgiu o Plano de Dados Abertos (PDA) que é um instrumento que orienta os órgãos da administração pública do Poder Executivo Federal a implementar e promover a abertura de dados. O PDA tem como objetivo organizar e padronizar o processo de publicação de dados abertos, resultando em maior disponibilidade, acesso, qualidade e ampla reutilização. Dessa forma, cada órgão/ entidade possui a obrigação de elaborar um PDA com vigência de dois anos (CGU, 2020). Para ajudar nesse propósito, em 2020, a CGU lançou o Manual de elaboração de PDAs, que passou a ser um importante instrumento de operacionalização da política de dados abertos, pois apresenta os aspectos gerais, o passo-a-passo para elaborar um PDA, modelo da estrutura formal de um PDA, pontos obrigatórios, sugestão de bases para abertura, além de recomendações e boas práticas para promoção, fomento, uso e reuso de DAG.

Outros documentos que merecem destaque com relação à orientação para publicação de DAG estão descritos no Quadro 4.

Quadro 4 – Documentos que orientam a publicação de DAGs

| DOCUMENTO | ESPECIFICAÇÃO |
|---|--|
| Padrão de Metadados | Conjunto padrão de campos obrigatórios e opcionais que descrevem os conjuntos de dados no portal dados.gov.br. |
| Cartilha Técnica para Publicação de Dados v1.0 | Inspirado em uma norma do governo do Chile para publicação de Dados Abertos o documento dispõe sobre a INDA, premissas, publicação, formatos, metadados, entre outros. |
| Manual de Catalogação no Portal dados.gov.br | Dispõe sobre todo passo-a-passo para catalogação no portal dados.gov.br. |
| Vocabulário Controlado de Governo Eletrônico do Governo Federal - VCGE | O VCGE dispõe sobre o vocabulário controlado utilizado pelo governo federal para indexar informações (documentos, bases de dados, sites, entre outros) |

Fonte: O Autor (2021).

No Brasil, o ponto central para publicação de Dados dos órgãos e/ou entidades do Poder Executivo Federal é o Portal Brasileiro de Dados Abertos⁹, constituindo-se como a porta de acesso ao repositório de dados abertos disponibilizados pelas várias instituições da Administração Pública. Este portal disponibiliza vários conjuntos de dados, para que qualquer cidadão, empresa ou a própria administração acesse, use e reutilize para diversos fins. Porém, apesar de ter vários conjuntos de dados disponibilizados, fica ainda o questionamento de se eles atendem aos critérios de maturidade e qualidade definidos para dados abertos.

2.6 Maturidade da abertura de dados

A *Open Knowledge International* (2021) define as três principais características da abertura de dados:

- **Disponibilidade e acesso:** os dados devem estar disponíveis por completo, com custo mínimo de reprodução, em formato conveniente e modificável por humanos e por máquinas.
- **Reutilização e redistribuição:** os termos legais que regem os dados devem permitir a sua reutilização, redistribuição e o intercâmbio com outros dados, ou seja, deve ser

⁹ <https://dados.gov.br/>

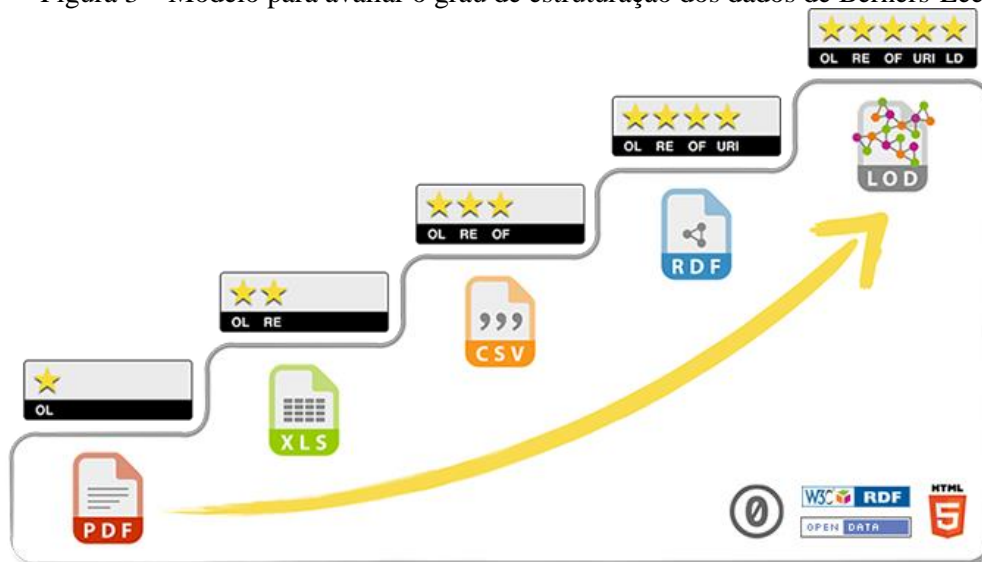
possível promover interoperabilidade.

- **Participação universal:** o uso de dados deve estar universalizado, pois todos devem ter a capacidade de usar, reutilizar e distribuir, não podendo haver nenhum tipo de restrição, seja ele utilizado para pesquisa ou exploração econômica.

Assim, os dados do governo precisam ser compartilhados, descobertos, acessíveis e manipuláveis por aqueles que desejam bem aproveitar as vantagens da Web e o acervo de informações das organizações (WORLD WIDE WEB CONSORTIUM, 2009).

Nesse sentido, no ano de 2010, o fundador da Web, Tim Berners-Lee, propôs um modelo para analisar a maturidade de abertura dos dados que ficou conhecido como esquema das cinco estrelas para Dados Abertos (Figura 3). Rocha e Vasconcelos (2004, p. 93) consideram que “Modelos de Maturidade fornecem aos gestores das organizações um poderoso instrumento para determinarem em que estágio de maturidade se encontram e planearem as ações necessárias para progredirem em direção a uma maturidade superior e, por consequência, alcançarem os objetivos desejados.”

Figura 3 – Modelo para avaliar o grau de estruturação dos dados de Berners-Lee



Fonte: <https://5stardata.info/en/>.

O modelo proposto possui uma classificação em estrelas onde a quantidade de estrelas tem um significado relacionado ao grau de estruturação dos dados (como pode ser visualizado na Figura 4), em direção aos dados abertos ligados, que são considerados a estrutura de abertura de dados mais completa e adequada.

Figura 4 – Classificação por estrelas do modelo proposto por Berners-Lee

| | |
|-------|---|
| ★ | Disponível na internet e com uma licença aberta – OL (<i>Open Licence</i>). |
| ★★ | Disponível como dados estruturados e legíveis por máquina – RE (<i>Readable</i>) |
| ★★★ | Como o anterior, mas disponível em formato aberto – OF (<i>Open Format</i>) |
| ★★★★ | Como todos os anteriores, e ainda usando os padrões aberto do W3C para identificar as coisas – URI (<i>Uniform Resource Identifier</i>) |
| ★★★★★ | Como todos os anteriores, e ainda referenciar dados de terceiros para prover contexto – LD (<i>Linked Data</i>) |

Fonte: Beners-Lee (2010).

Isotani e Betencourt (2015) recomendam que a abertura de dados deve ser feita trabalhando no contexto de pelo menos três estrelas, pois consideram que estar disponível na internet, em formato aberto, sob uma licença aberta e em formato estruturado e legível também por máquinas é a base para a abertura de dados, para que eles possam ser efetivamente utilizados. Os referidos autores também apontam benefícios para os consumidores e os publicadores dos dados, sob o ponto de vista de cada nível (quantidade de estrelas) de abertura de dados, como descrito no Quadro 5.

Quadro 5 – Benefícios para consumidores e publicadores de dados segundo o esquema cinco estrelas

| ESTRELAS | PARA CONSUMIDORES | PARA PUBLICADORES |
|-----------|--|--|
| ★ | Ver os dados; Imprimi-los; Guardá-los (no disco rígido ou em um pen-drive, por exemplo); Modificar os dados como queira; Acessar o dado de qualquer sistema; Compartilhar o dado com qualquer pessoa. | E simples de publicar; Não precisa explicar repetitivamente que as pessoas podem fazer uso dos dados |
| ★ ★ | Os mesmos benefícios de quem usa 1 estrela; Usar softwares proprietários para processar, agregar, calcular e visualizar os dados; Exportá-los em qualquer formato estruturado. | É fácil de publicar. |
| ★ ★ ★ | Os mesmos benefícios de quem usa 2 estrelas; Manipular os dados da forma que lhe agrada, sem estar refém de algum software em particular. | E ainda fácil de publicar. Obs.: Podem ser necessários conversores ou <i>plugins</i> para exportar os dados do formato proprietário. |
| ★ ★ ★ ★ | Os mesmos benefícios de quem usa 3 estrelas; Fazer marcações; Reutilizar parte dos dados Reutilizar ferramentas e bibliotecas de dados existentes, mesmo que elas entendam apenas parte dos padrões usados por quem publicou; Combinar os dados com outros | Há controle dos itens dos dados e pode melhorar seu acesso; Outros publicadores podem conectar seus dados, promovendo-os as 5 estrelas. |
| ★ ★ ★ ★ ★ | Descobrir mais dados vinculados enquanto consome dados; Aprender sobre a classificação das 5 estrelas. | Torna o dado mais fácil de ser descoberto; Aumenta o valor do dado; A organização ganha os |

| | | |
|--|--|--|
| | | mesmos benefícios com a vinculação de dados que os consumidores. |
|--|--|--|

Fonte: Isotani e Bittencourt (2015).

Apesar desses benefícios expressos no Quadro 5, segundo Beghin e Zigone (2014), a maioria dos portais brasileiros que disponibilizam DAG não atendem a esse requisito de três estrelas, pois a maioria fornece dados em formato PDF e, portanto, são classificados com uma estrela. Villazón-Terrazas *et al.* (2011) aponta que o esquema cinco estrelas é mais uma alternativa para guiar as instituições publicadoras para que disponibilizem dados com mais benefícios para os consumidores de dados, mesmo que isto possa implicar em maior trabalho para os publicadores. Mas o autor ressalta que, na maioria das vezes, as orientações existentes são destinadas a desenvolvedores de software, não para governos e sua administração em geral.

2.7 Princípios FAIR

Fruto de uma mudança no contexto científico, em que o acesso, uso e reuso de dados enfrentava dificuldades, surge no ano de 2014, as primeiras discussões sobre os princípios FAIR. (HENNING *et al.* 2019). Como resultado disso, a comunidade global entra em consenso quanto à necessidade de criar uma infraestrutura global que suportasse a abertura de dados de pesquisa voltados para publicação, compartilhamento e reutilização de dados. (DATA FAIRPORT, 2014). Apesar de terem sido definidos em um contexto de dados de pesquisa, a adoção dos princípios FAIR pode contribuir para que os DAG passem a ser descobertos, acessados, reutilizados e, se forem devidamente adotados, viabilizam a interoperabilidade entre as diferentes bases de conjuntos de dados.

O acrônimo FAIR estabelece que os dados devem ser Encontráveis (*Findable*), Acessíveis (*Accessible*), Interoperáveis (*Interoperable*) e Reutilizáveis (*Reusable*), com destaque para a adoção de metadados bem definidos, a fim de facilitar o uso dos dados por máquinas, assim como por pessoas. Metadados podem ser entendidos como informações estruturadas que descrevem, explanam, localizam ou facilitam a recuperação, o uso e o gerenciamento de um dado ou recurso informacional (HODGE, 2001).

Uma síntese dos princípios FAIR, de acordo com o *FAIR Guiding Principles* (WILKINSON *et al.*, 2016) é apresentada no Quadro 6.

Quadro 6 – Síntese dos Princípios FAIR

| <i>FINDABLE</i> (Encontráveis) | <i>Accessible</i> (Acessíveis) | <i>Interoperable</i> (Interoperáveis) | <i>Reusable</i> (Reutilizáveis) |
|---|---|--|--|
| F.1 Metadados devem ter identificadores globais, | A.1 Metadados devem ser recuperáveis pelos seus identificadores usando protocolo de comunicação padronizado. | I.1 Metadados devem ser representados por meio de uma linguagem formal, acessível, compartilhada e amplamente aplicável para representação do conhecimento. | R.1 Metadados são descritos com uma pluralidade de atributos precisos e relevantes. |
| F.2 Dados devem ser descritos utilizando metadados ricos. | A.1.1 O protocolo deve ser aberto, gratuito e universalmente implementável | I.2 Metadados devem usar vocabulários de acordo com os princípios <i>FAIR</i> . | R.1.1 Metadados devem ser disponibilizados com licenças de uso claras e assecíveis |
| F.3 Metadados devem incluir clara e explicitamente os identificadores dos dados que o descrevem. | A.1.2 O protocolo deve permitir procedimentos de autenticação e autorização, quando necessário. | I.3 Metadados devem incluir referência qualificadas para outros metadados. | R.1.2 Metadados devem estar associados a sua proveniência. |
| F.4 Metadados devem ser registrados ou indexados em mecanismos de busca | A.2 Metadados devem ser acessíveis, mesmo quando os dados não estiverem mais disponíveis. | | R.1.3 Metadados devem estar alinhados com padrões relevantes do seu domínio. |

Fonte: Baseado em Wilkinson *et al.* (2016).

Ressalta-se o quanto a adoção de metadados bem definidos é um elemento primordial para gestão de dados, uma vez que o processo de acesso, uso e reuso dos dados requer a utilização de padrões de metadados.

3 PROCEDIMENTOS METODOLÓGICOS

Esta é uma pesquisa quanto aos fins descritiva e qualitativa; e quanto aos meios é bibliográfica e documental (MICHEL, 2009), que realiza um estudo de caso com alguns conjuntos de dados da previdência social, extraídos do portal brasileiro de dados abertos¹⁰ da previdência social. As atividades desenvolvidas foram divididas em etapas que estão descritas a seguir.

1) **Revisão bibliográfica** – foi realizada para buscar entender a origem, o conceito de governo aberto e de dados abertos governamentais; os princípios e critérios aplicáveis a dados abertos em geral e, especificamente, a dados abertos governamentais. Para isso, foram realizados levantamentos de artigos (em periódicos e em eventos), livros, teses e dissertações em bases de dados como: BRAPCI, *Google Scholar* e Biblioteca Digital Brasileira de Teses e Dissertações (BDTD). Foram utilizados os termos de busca “dados abertos governamentais”; “dados abertos + governo aberto”; “qualidade de dados abertos”; “dados + governo aberto”; “princípios + dados abertos”; “critérios + dados abertos”; “disponibilização + dados abertos governamentais”, tanto em português, quanto em inglês. Estabeleceu-se o período de dez anos e o critério de inclusão foi a publicação estar disponível na íntegra e ser um texto completo. Essa etapa serviu para o embasamento teórico da pesquisa e deu subsídios para a análise e discussão dos resultados. A estruturação e conexão dos principais conceitos estudados e utilizados na pesquisa foram organizados por meio de um mapa conceitual, que será apresentado na seção 4.

Análise Documental – analisou-se algumas legislações e materiais não convencionais, como conteúdo de sites, que foram relevantes para complementar o embasamento sobre a temática sendo pesquisada. Alguns desses documentos estão especificados no Quadro 7.

Quadro 7 – Documentos analisados na pesquisa

| |
|---|
| <p>LEI COMPLEMENTAR Nº 131, DE 27 DE MAIO DE 2009 – que acrescenta dispositivos à Lei Complementar Nº 101, de 4 de maio de 2000, que estabelece normas de finanças públicas voltadas para a responsabilidade na gestão fiscal e dá outras providências, a fim de determinar a disponibilização, em tempo real, de informações pormenorizadas sobre a execução orçamentária e financeira da União, dos Estados, do Distrito Federal e dos Municípios;</p> |
| <p>Lei No. 12.527/2011, conhecida como Lei de acesso à informação (LAI), que dispõe sobre os procedimentos a serem observados pela União, Estados, Distrito Federal e Municípios, com o fim de</p> |

¹⁰ <https://dados.gov.br/>

| |
|--|
| garantir o acesso a informações; |
| Instrução Normativa, Nº 4 de 13 de abril de 2012 - Institui a Infraestrutura Nacional de Dados Abertos - INDA. |
| Decreto Nº 7.724, de 16 de maio de 2012 - Regulamenta a Lei Nº 12.527, de 18 de novembro de 2011; |
| Lei nº 12.965, de 23 de abril de 2014 , conhecida como Marco Civil da Internet, que estabelece princípios, garantias, direitos e deveres para o uso da Internet no Brasil; |
| Decreto N 8.771, de 11 de maio de 2016 , que regulamenta o Marco Civil da Internet; |
| Decreto N 8.777, de 11 de maio de 2016 , que institui a Política de Dados Abertos do Governo Federal; |
| Resolução Nº 2, de 24 de março de 2017 - Aprova os Termos de Uso do Portal Brasileiro de Dados Abertos; |
| Resolução Nº 3, de 13 de outubro de 2017 - Aprova as normas sobre elaboração e publicação de Planos de Dados Abertos, conforme disposto no Decreto nº 8.777, de 11 de maio de 2016. |
| Lei nº 13.709, de 14 de agosto de 2018 , conhecida como Lei Geral de Proteção de Dados Pessoais, que dispõe sobre a proteção de dados pessoais e altera a Lei nº 12.965, de 23 de abril de 2014 (Marco Civil da Internet); |
| Medida Provisória n. 869 de 2018 . Altera a Lei nº 13.709, de 14 de agosto de 2018, para dispor sobre a proteção de dados pessoais e para criar a Autoridade Nacional de Proteção de Dados, e dá outras providências; |
| Lei nº 13.853, de 8 de julho de 2019 , conhecida como Lei Geral de Proteção de Dados Pessoais (LGPD), que dispõe sobre a proteção de dados pessoais e da criação da Autoridade Nacional de Proteção de Dados e altera a Lei nº 13.709, de 14 de agosto de 2018; |
| Decreto Nº 9.903, de 8 de julho de 2019 - Altera o Decreto nº 8.777, de 11 de maio de 2016, que institui a Política de Dados Abertos do Poder Executivo federal, para dispor sobre a gestão e os direitos de uso de dados abertos; |
| Manual de Dados Abertos: governo ¹¹ |
| Cartilha Técnica para Publicação de Dados Abertos no Brasil ¹² |

Fonte: O Autor (2021).

Além de terem sido consultadas e lidas as informações do Portal Brasileiro de Dados Abertos¹³ e do Grupo da W3C Dados Abertos: Governo e Sociedade da W3C¹⁴.

¹¹ https://www.w3c.br/pub/Materiais/PublicacoesW3C/Manual_Dados_Abertos_WEB.pdf

¹² <https://dados.gov.br/pagina/cartilha-publicacao-dados-abertos>

¹³ <https://dados.gov.br/>

¹⁴ <https://www.w3c.br/GT/GrupoDadosAbertos>

Realização do Estudo de Caso - Levantamento de dados no Portal Brasileiro de Dados Abertos - Foi realizada a coleta dos dados necessários para realização do estudo de caso pretendido. O critério de busca foi o grau de relevância utilizando o termo “conjunto de dados previdência social” no campo de busca do portal. A partir dos resultados, foram selecionados os conjuntos publicados na forma coletada na fonte, com a mais fina granularidade possível. Com os dados recuperados foram criadas duas planilhas no software Excel, a fim de organizar os dados de acordo com: **a) grau de relevância; b) anos de publicação dos dados; c) data de atualização d) título; e) formato; f) Autores; g) Licenças e h) Observações Gerais.** A navegação por esses dados permitiu conhecer um pouco a estruturação dos mesmos e os tipos e formatos dos dados disponibilizados. Como resultado foram coletados e organizados 119 conjuntos de dados da previdência social brasileira.

Análise dos dados recuperados - Para a avaliação de qualidade e maturidade dos dados recuperados foi realizada uma análise nos conjuntos de dados coletados, no conteúdo das páginas de disponibilização e nos formatos de publicação para identificar os dados que, de fato, atendessem aos critérios de qualidade em sua abordagem. Os registros que não possuíam a atualização periódica por ano ou por mês foram descartados das planilhas. Todas as dificuldades relativas à recuperação e uso destes conjuntos de dados também foram registradas. Para a realização da análise, tomou-se como base os 8 (oito) princípios de qualidade que norteiam a publicação de Dados Abertos Governamentais (DAG) (MALAMUD *et al.*, 2007), que estabelecem que os dados precisam ser:

- **Completo** - todos os dados que não possuam limitações de privacidade, segurança e que não requerem privilégios de acesso, devem ser disponibilizados de forma pública.
- **Primários** - os dados devem ser disponibilizados na maior granularidade possível, sem agregação ou modificação, tal qual foram produzidos na fonte.
- **Atuais** - os dados devem estar disponíveis o mais breve possível após produzidos e devem ser mantidos atualizados.
- **Acessíveis** - os dados devem estar disponíveis para o maior escopo possível de usuários e poderem ser utilizados para as mais diferentes finalidades.
- **Processáveis por máquinas** – os dados devem ser estruturados de forma a permitir processamento automatizado.
- **Não-discriminatórios** - os dados devem estar disponíveis para todos, sem necessidade de cadastro.

- **Não-proprietários** – os dados devem estar em formatos não-proprietários, ou seja, que não sejam de propriedade exclusiva de alguma entidade.
- **Livres de licenças** - os dados não devem estar sujeitos a nenhuma regulação de direitos autorais, patentes, propriedade intelectual ou segredo industrial. Podem ser admitidas restrições sensatas relacionadas à privacidade, segurança e privilégios de acesso.

Assim como os princípios FAIR, que estabelecem que os dados devem ser **Encontráveis** (*Findable*), **Acessíveis** (*Accessible*), **Interoperáveis** (*Interoperable*) e **Reutilizáveis** (*Reusable*) (GO FAIR, 2018). Já para a análise da maturidade de abertura dos dados foi utilizado o modelo cinco estrelas de Tim Berners-Lee (2010), onde a quantidade de estrelas tem um significado relacionado ao grau de estruturação dos dados em direção aos dados abertos ligados, levando em conta características da disponibilização dos dados.

Durante as primeiras análises, verificou-se a necessidade de incluir na análise mais três critérios a respeito da qualidade dos dados, são eles: **Credibilidade** (o quão um dado é considerado verdadeiro e confiável); **Consistência** (o quanto existe uma padronização entre os conjuntos de dados e nos metadados adotados) e **Facilidade de Uso por Humanos** (o quanto a visualização dos dados pode ser compreendida, da forma que está, por humanos) (SANTOS, 2009). Vale salientar que alguns critérios foram unificados devido ao fato deles estarem presentes em ambas as categorias de avaliação.

Os critérios escolhidos foram usados para compor um protocolo de avaliação, que está resumido no Quadro 8.

Quadro 8 – Protocolo de avaliação da pesquisa

| ITEM | CRITÉRIO | ORIGEM |
|------|----------------------------|---|
| 01 | PRIMÁRIOS. | Princípios para DAGs. |
| 02 | ACESSÍVEL | Princípios para DAGs e Princípios FAIR. |
| 03 | CREDIBILIDADE | Qualidade dos dados. |
| 04 | COMPLETOS | Princípios para DAGs. |
| 05 | NÃO-DISCRIMINATÓRIOS | Princípios para DAGs. |
| 06 | CONCISÃO | Princípios FAIR. |
| 07 | PROCESSAMENTO POR MÁQUINAS | Princípios para DAGs e Cinco estrelas para Dados Abertos. |

| | | |
|----|--------------------------------|---|
| 08 | INTERPRETABILIDADE | Cinco estrelas para Dados Abertos. |
| 09 | LIVRE DE LICENÇAS | Princípios para DAGs e Cinco estrelas para Dados Abertos. |
| 10 | NÃO PROPRIETÁRIOS | Princípios para DAGs. |
| 11 | ATUALIDADE | Princípios para DAGs. |
| 12 | ENCONTRABILIDADE | Princípios <i>FAIR</i> . |
| 13 | REUSABILIDADE | Princípios <i>FAIR</i> . |
| 14 | INTEROPERABILIDADE | Princípios <i>FAIR</i> . |
| 15 | MATURIDADE | Cinco estrelas para Dados Abertos. |
| 16 | CONSISTÊNCIA. | Qualidade dos dados. |
| 17 | FACILIDADE DE USO POR HUMANOS. | Qualidade dos dados. |

Fonte: O autor (2021).

O uso do protocolo de avaliação visou padronizar e facilitar a análise dos conjuntos de dados mapeados que integraram o estudo de caso.

4 RESULTADOS E DISCUSSÕES

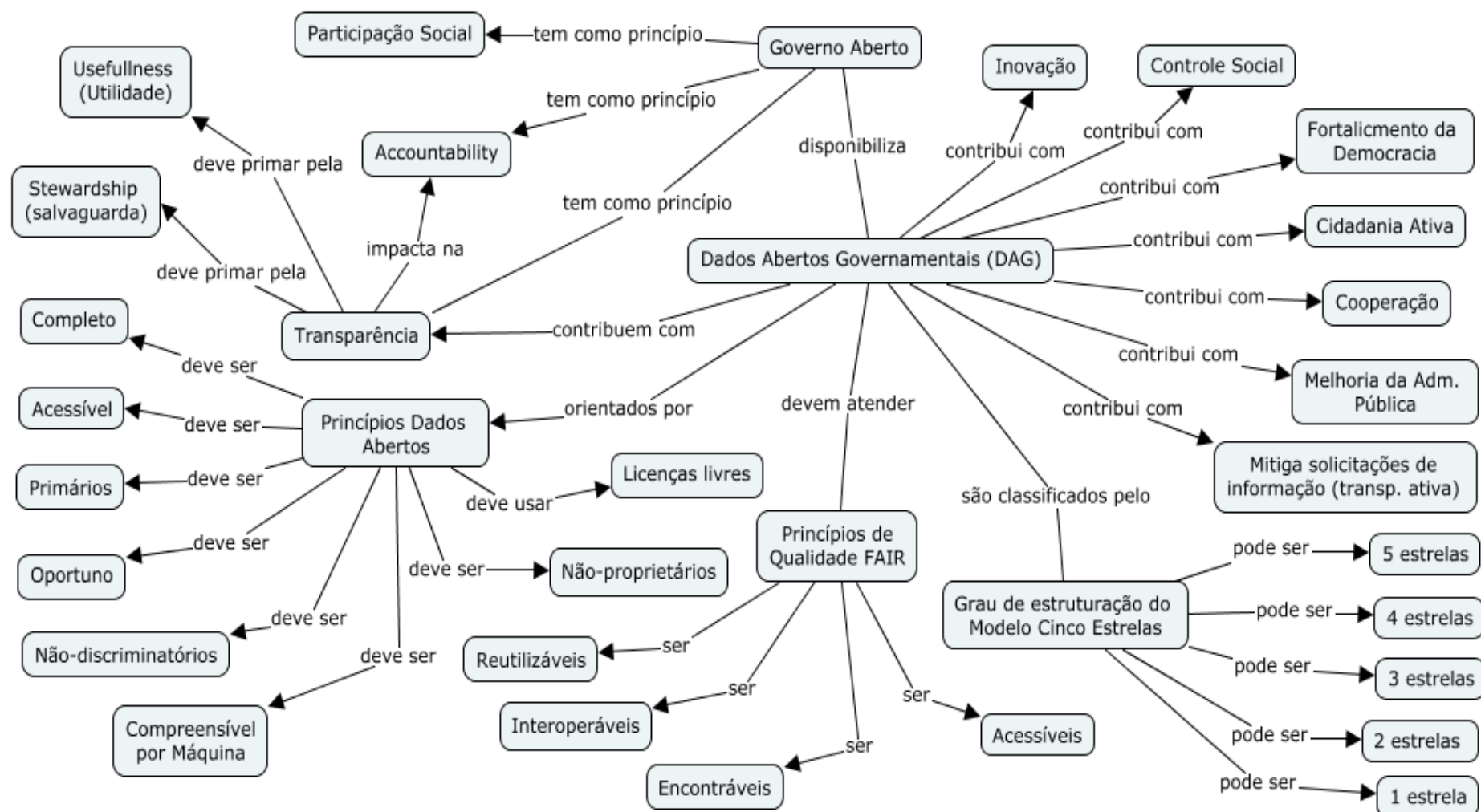
Antes de iniciar o estudo de caso, achou-se por bem revisar o referencial teórico e estruturar todos os conceitos estudados na etapa de levantamento bibliográfico, o que deu origem ao mapa conceitual da Figura 5, que também pode ser visto como uma contribuição da pesquisa.

O conceito inicial estudado foi o de governo aberto, que representa um método de governança que oferece novos espaços de abertura, transparência e diálogo contínuo entre o governo e seus cidadãos (VELJKOVIĆ; BOGDANOVIĆ-DINIĆ; STOIMENOV, 2014). Para a implementação de um Governo Aberto existem três princípios-chave: *accountability*, transparência e participação social (UBALDI, 2013). Por sua vez, a transparência possui dois princípios: o *stewardship* (salvaguarda) que se concentra em garantir a precisão, validade, segurança, gestão e preservação das informações sobre a tutela de alguém, no caso o governo. Por esse princípio, de acordo com Dawes (2010), todos os agentes públicos e organizações governamentais são responsáveis pelo tratamento e manuseio de informações com cuidado e integridade, independentemente da sua finalidade ou fonte; e a utilidade (*usefulness*) reconhece que a informação precisa estar adequada ao propósito e ao uso (KLEIN, 2017).

No contexto do governo aberto os DAG são disponibilizados e devem atender os princípios dos dados abertos, os princípios FAIR e, dependendo de como estão estruturados e conectados, são enquadrados em algum dos níveis de maturidade de abertura de dados do modelo de 5 estrelas (que vai de 1 a 5 estrelas) (BENERS-LEE, 2009).

Os DAG tem potencial para contribuir, segundo Klein (2017), para o controle social, o fortalecimento da democracia, a cidadania ativa, as melhorias na administração pública, a inovação, a cooperação e a transparência. Além de contribuírem para a mitigação das solicitações de informações ao governo (transparência passiva), possibilitando que as solicitações sejam acessadas diretamente pela população e trocada entre os entes governamentais (transparência ativa).

Figura 5 – Mapa Conceitual que resume os principais conceitos trabalhados na pesquisa

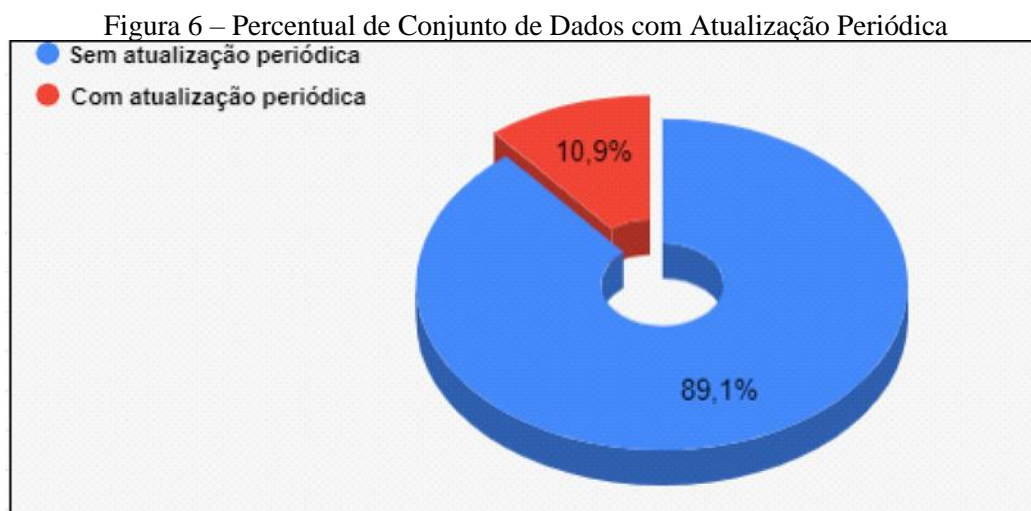


Fonte: Baseado em Klein (2017).

Dando sequência ao desenvolvimento da pesquisa, foi realizada a análise dos 119 conjuntos de dados coletados. A primeira parte da análise buscou informações acerca de: quais conjuntos de dados constam com atualização periódica; quais deles dispõem de bons metadados, capazes de fornecer informação sobre o conteúdo do conjunto de dados, sem dúvida para o usuário; além de informação sobre em quais formatos e sob quais licenças de publicação esses dados estão sendo disponibilizados. Posteriormente, é apresentada a análise e discussão dos conjuntos de dados de acordo com os critérios de análise descritos nos procedimentos metodológicos.

4.1 Atualidade, uso de metadados, formato de dados e licenças de uso

Dos 119 conjuntos de dados recuperados e analisados apenas 13 conjuntos possuem atualização periódica (10,9% dos conjuntos), como pode ser observado na Figura 6. Também, em alguns conjuntos de dados disponibilizados, colocados como sumarização anual, constam apenas alguns meses dos anos indicados (quando deveriam estar todos os meses do referido ano). Isto viola o princípio da **atualidade dos dados**.



Fonte: Dados da Pesquisa (2021).

Observou-se que há conjuntos de dados que estão sem atualização desde janeiro de 2018, o que é uma lacuna muito grande, que se traduz em informações disponibilizadas à sociedade que podem não corresponder à realidade atual da previdência social. Ressalta-se que a previdência social brasileira, disponibiliza o seu plano de dados abertos (PDA), válido para o biênio julho/2016 a julho/2018 (versão mais atual encontrada) ou seja, o PDA encontra-se também desatualizado e seria ele que apresentaria a prioridade dos dados que

serão disponibilizados para a sociedade, e especificação de que não recaia nenhuma vedação de acesso sobre os dados.

Quanto ao uso de metadados, há conjuntos de dados sem especificação de metadado algum ou que em seus metadados há nomes conflitantes, como mês de atualização e data de atualização, que causa dúvidas ao usuário sobre o que está sendo representado nos conjuntos de dados. No PDA de julho/2018 disponibilizado, são estabelecidas descrições obrigatórias para os dados, denominadas de premissas, que devem ser fornecidas junto com os conjuntos de dados. Analisando os conjuntos de dados recuperados no contexto dessas premissas foi obtido o resultado apresentado no Quadro 9.

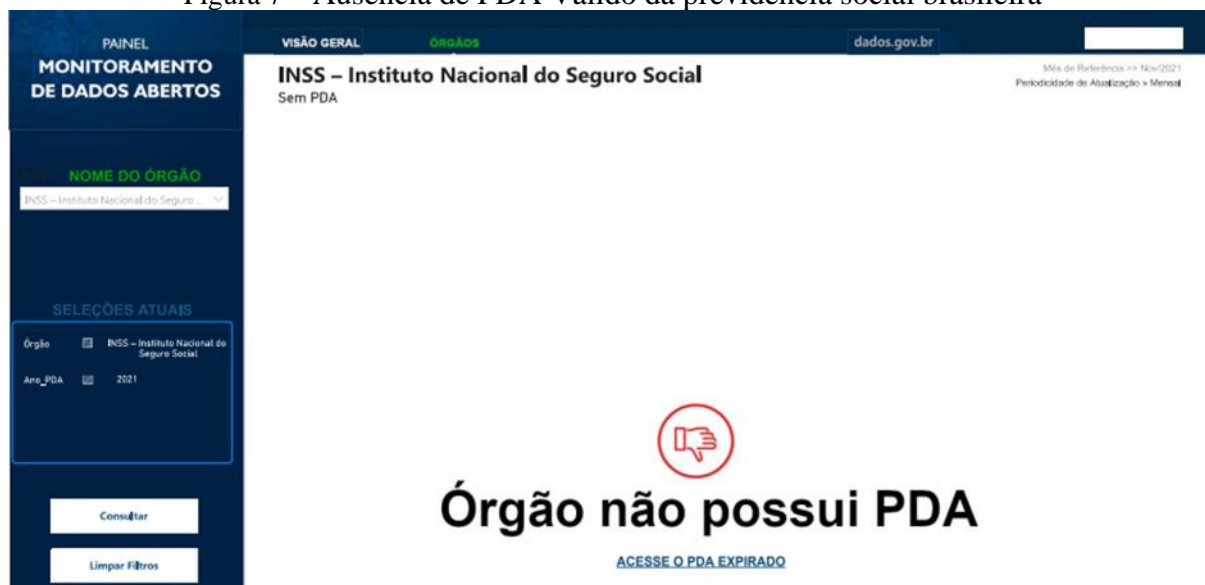
Quadro 9 – Metadados dos conjuntos de dados da Previdência

| PREMISSAS | STATUS |
|---|--|
| Nome ou título do conjunto de dados | Presente |
| Descrição sucinta | Presente |
| Palavras-chave (etiquetas) | Presente em quantidades variáveis (variando de 1 a 3) |
| Assuntos relacionados do VCGE ¹ - Vocabulário Controlado do Governo Eletrônico | Presente apenas em alguns conjuntos em formato de link, porém, o link não funciona, ou seja, não leva a lugar algum. |
| Nome e e-mail do setor responsável pelos dados | Consta apenas o nome do setor responsável em alguns e em outros conjuntos nada consta. |
| Periodicidade de atualização | Ausente em diversos conjuntos |
| Escopo temporal (anual, mensal, diário, bimestral) | Presente |
| Escopo geo-político (por cidade, por estado, por região). | Presente, mas limitado ao valor “Brasil”, não detalhando região, estado ou cidade. |
| Outros Metadados | Mantenedor (que é copiado o valor do campo Nome do responsável), data da última atualização e data da criação. |

Fonte: Dados da Pesquisa (2021).

Vale ressaltar, que a disponibilização de um PDA, conforme indicado pela INDA deve ser realizada a cada dois anos. Realizada consulta ao painel de monitoramento de Dados Abertos do Poder Executivo Federal, constatou-se que a Previdência Social brasileira não possui PDA válido, conforme pode ser visto na Figura 7. A ausência de um PDA válido é um fator crítico, pois traz impactos na implementação e promoção da abertura de dados pelo órgão, uma vez que os padrões mínimos de qualidade não estão sendo atendidos. Dessa forma, trazendo dificuldades para entender e reutilizar os conjuntos de dados disponibilizados.

Figura 7 – Ausência de PDA Válido da previdência social brasileira



Fonte: Painel de Monitoramento de Dados Abertos (2021).

Além disso, os conjuntos de dados analisados acabam por violar os critérios de qualidade que são impactados pela falta de metadados, tais como: **Acessibilidade**, **interpretabilidade** e **completude**. Por consequência, dificultando o **uso e reuso** dos dados. Além de também haver a violação do critério **consistência**, devido ao preenchimento nem sempre consistente de alguns dos campos e não padronizados entre os conjuntos de dados. De fato, apesar de existirem normativas que orientam de que forma devem ser publicados os conjuntos de dados, e que informação sobre os mesmos deve estar disponível, são variadas as formas de disponibilização, não havendo uma uniformidade.

Como contribuição desta pesquisa, no Quadro 10 são sugeridos metadados mínimos e obrigatórios para os conjuntos de dados, com base no INDA (2019).

Quadro 10 – Sugestão de Metadados obrigatórios

| METADADOS OBRIGATÓRIOS |
|---|
| Título: Nome do conjunto de dados. |
| Descrição: Uma breve explicação sobre os dados. |
| Catálogo origem: Página (URL) do órgão onde está publicado o conjunto de dados. |
| Órgão responsável: Nome e sigla do órgão ou entidade responsável pela publicação do conjunto de dados. |
| Categorias no VCGE: O Vocabulário Controlado de Governo Eletrônico é uma lista hierarquizada de assuntos do governo que utiliza termos comuns e é voltada para a sociedade. Para navegar e escolher as categorias acesse o VCGE em https://vocab.e.gov.br/2011/03/vcge . |

| |
|---|
| Recursos: Um conjunto de dados pode ser composto por mais de um arquivo de dados. O critério básico para separar vários recursos em mais de um conjunto de dados é a constatação de que eles divergem em vários metadados. |
| Identificador: URL persistente que aponta para o recurso na Web. |
| Título: Nome do recurso. |

Fonte: Baseado nas recomendações do Inda (2019).

A adoção de metadados descritivos na disponibilização dos dados é de suma importância, a fim de proporcionar informação relevante para quem deseje usar os conjuntos de dados. Porém, os conjuntos de dados da previdência social brasileira recuperados mal atendem aos metadados básicos mínimos, especificados no próprio PDA (ressalta-se que fora de validade) da instituição. Além disso, não existe uma padronização de preenchimento entre os diversos conjuntos, como pode ser visualizado nas Figuras 8 e 9. Nestas figuras, tanto os metadados utilizados não estão padronizados, como o preenchimento do campo última atualização é diferente entre os conjuntos de dados.

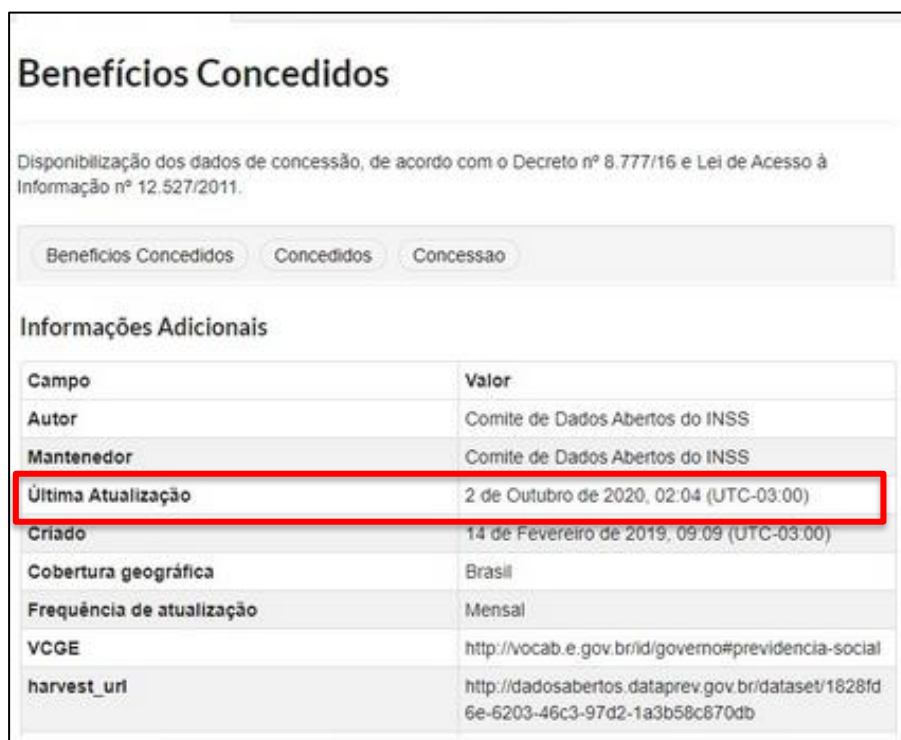
Figura 9 – Preenchimento não padronizado dos metadados



| Perfil das unidades do INSS | |
|--|---|
| Agências da Previdência Social com os respectivos horários de atendimento bem como os emails e seus municípios. | |
| Dados e recursos | |
| <div> <div>Agências da Previdência Social setembro 2020</div> <div>Agências da Previdência Social setembro 2020</div> <div>Explorar</div> </div> | |
| <div> <div>APS</div> <div>INSS</div> <div>agências</div> </div> | |
| Informações Adicionais | |
| Campo | Valor |
| Autor | Comite de Dados Abertos do INSS |
| Mantenedor | Comite de Dados Abertos do INSS |
| Cobertura geográfica | Brasil |
| Frequência de atualização | Mensal |
| VCGE | http://vocab.e.gov.br/id/governo#previdencia-social |
| Última Atualização | outubro de 2020 |

Fonte: dados.gov.br (2021).

Figura 9 – Preenchimento não padronizado dos metadados



Benefícios Concedidos

Disponibilização dos dados de concessão, de acordo com o Decreto nº 8.777/16 e Lei de Acesso à Informação nº 12.527/2011.

Benefícios Concedidos Concedidos Concessao

Informações Adicionais

| Campo | Valor |
|---------------------------|---|
| Autor | Comite de Dados Abertos do INSS |
| Mantenedor | Comite de Dados Abertos do INSS |
| Última Atualização | 2 de Outubro de 2020, 02:04 (UTC-03:00) |
| Criado | 14 de Fevereiro de 2019, 09:09 (UTC-03:00) |
| Cobertura geográfica | Brasil |
| Frequência de atualização | Mensal |
| VCGE | http://vocab.e.gov.br/id/governo#previdencia-social |
| harvest_url | http://dadosabertos.dataprev.gov.br/dataset/1628fd6e-6203-46c3-97d2-1a3b58c870db |

Fonte: dados.gov.br (2021).

Além disso, todos os 119 conjuntos de dados apresentam em seus metadados alguma informação incompleta ou confusa. Por exemplo, o campo "última atualização", como ilustra a Figura 10, aparece duas vezes em alguns conjuntos de dados e com preenchimento diferente em cada campo.

Figura 10 – Campo última atualização duplicado e inconsistente

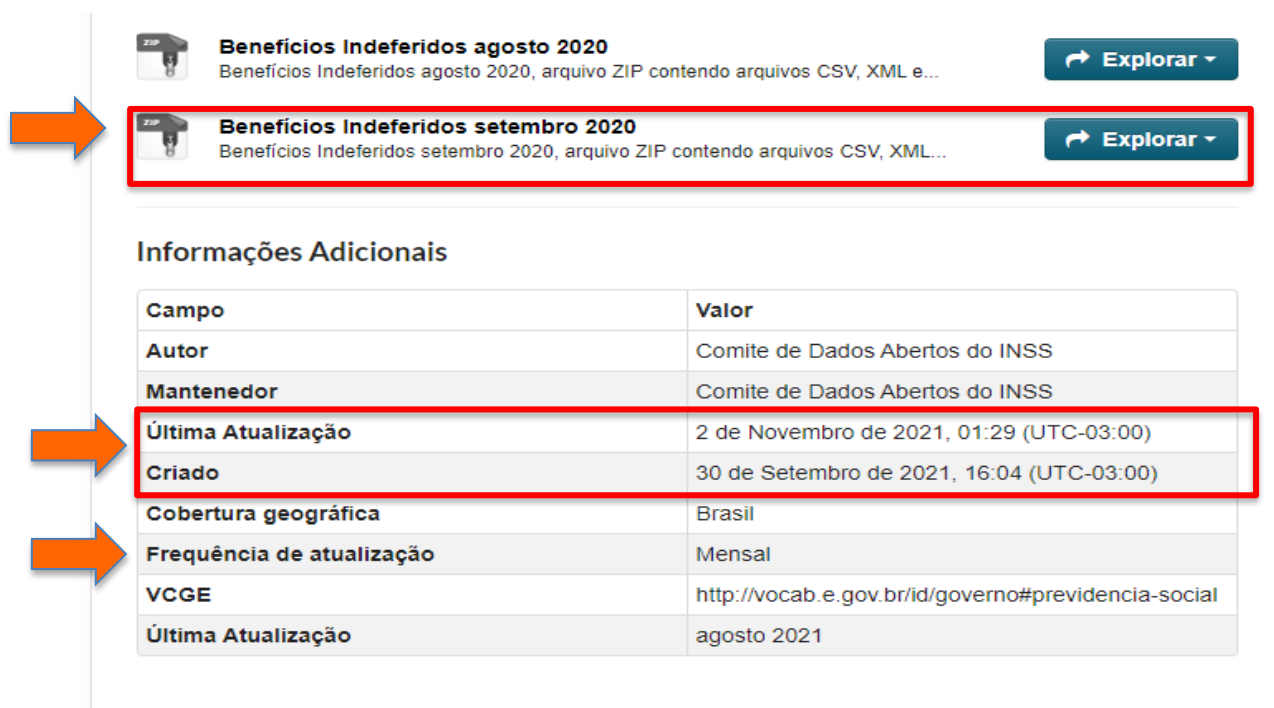
Informações Adicionais

| Campo | Valor |
|---------------------------|---|
| Autor | Comite de Dados Abertos do INSS |
| Mantenedor | Comite de Dados Abertos do INSS |
| Última Atualização | 2 de Novembro de 2021, 01:28 (UTC-03:00) |
| Criado | 30 de Setembro de 2021, 16:03 (UTC-03:00) |
| Cobertura geográfica | Brasil |
| Frequência de atualização | Trimestral |
| VCGE | http://vocab.e.gov.br/id/governo#previdencia-social |
| Última Atualização | agosto 2021 |

Fonte: dados.gov.br (2021).

Outro ponto que pode trazer dificuldades aos usuários é o fato de que o conjunto de dados ao ser explorado, por exemplo o arquivo de dados “Benefícios Indeferidos”, em seu metadados, além de não esclarecer quando de fato ocorreu a última atualização (datas conflitantes entre o arquivo e os campos de metadados), a disponibilização consta até o mês de setembro de 2020, quando é especificado que os dados deveriam ser atualizados mensalmente, conforme dispõe a INDA (Figura 11).

Figura 3 - Inconsistência nos metadados e atendimento à INDA.



Benefícios Indeferidos agosto 2020
Benefícios Indeferidos agosto 2020, arquivo ZIP contendo arquivos CSV, XML e...

Benefícios Indeferidos setembro 2020
Benefícios Indeferidos setembro 2020, arquivo ZIP contendo arquivos CSV, XML...

Informações Adicionais

| Campo | Valor |
|---------------------------|---|
| Autor | Comite de Dados Abertos do INSS |
| Mantenedor | Comite de Dados Abertos do INSS |
| Última Atualização | 2 de Novembro de 2021, 01:29 (UTC-03:00) |
| Criado | 30 de Setembro de 2021, 16:04 (UTC-03:00) |
| Cobertura geográfica | Brasil |
| Frequência de atualização | Mensal |
| VCGE | http://vocab.e.gov.br/id/governo#previdencia-social |
| Última Atualização | agosto 2021 |

Fonte: Dados da pesquisa (2021) – pesquisa realizada em 15 de novembro/2021.

Ao clicar no link “Explorar” dos Benefícios Indeferidos Setembro 2020 (Figura 11) surgem outras inconsistências na descrição dos metadados (Figura 12), uma vez que a data do conjunto difere completamente do preenchimento dos metadados, além de divergir dos que é estabelecido pela INDA.

Destaca-se que não foi encontrado dicionário de dados para os conjuntos de dados, o que dificulta, por exemplo, a compreensão dos nomes das colunas da extensa planilha em excel que é produzida a partir dos conjuntos de dados publicados no formato CSV.

Figura 12 – Inconsistência no preenchimento dos Metadados

Benefícios Indeferidos setembro 2020

URL: https://armazenamento-dadosabertos.s3.sa-east-1.amazonaws.com/Plano+2016_2018_Grupos+de+dados/INSS+-+Benefícios+indeferidos+setembro+2020.zip

Benefícios Indeferidos setembro 2020, arquivo ZIP contendo arquivos CSV, XML e JSON

Não há visões criadas para este recurso ainda.

| Recursos | Informações Adicionais | | | | | | | | | | |
|--------------------|---|-------|-------|--------------------|-----------------|--------|-----------------|---------|-----------------|---------|------------------------------|
| Benefícios ... | <table border="1"> <thead> <tr> <th>Campo</th> <th>Valor</th> </tr> </thead> <tbody> <tr> <td>Ultima atualização</td> <td>21/Outubro/2021</td> </tr> <tr> <td>Criado</td> <td>21/Outubro/2021</td> </tr> <tr> <td>Formato</td> <td>application/zip</td> </tr> <tr> <td>Licença</td> <td>Creative Commons Attribution</td> </tr> </tbody> </table> | Campo | Valor | Ultima atualização | 21/Outubro/2021 | Criado | 21/Outubro/2021 | Formato | application/zip | Licença | Creative Commons Attribution |
| Campo | Valor | | | | | | | | | | |
| Ultima atualização | 21/Outubro/2021 | | | | | | | | | | |
| Criado | 21/Outubro/2021 | | | | | | | | | | |
| Formato | application/zip | | | | | | | | | | |
| Licença | Creative Commons Attribution | | | | | | | | | | |
| Benefícios ... | | | | | | | | | | | |
| Benefícios ... | | | | | | | | | | | |
| Benefícios ... | | | | | | | | | | | |
| Benefícios ... | | | | | | | | | | | |

Fonte: dados.gov.br (2021).

Com relação ao formato de publicação dos dados, apesar de existirem normativas que orientam de que forma devem ser publicados os conjuntos de dados, são variadas as formas de disponibilização, não havendo uma padronização. Podem ser encontrados arquivos nos seguintes formatos: RDF, JSON, XML, CSV, ZIP e PDF, conforme pode ser visualizado na Figura 13. Observa-se que o descrito viola o critério da **consistência**, uma vez que não há um padrão de formato(s) de publicação e alguns dos formatos adotados podem trazer dificuldade para a **leitura por máquinas**, como os formatos PDF e ZIP.

Figura 13 – Quantidade de conjuntos de dados disponíveis por formato

| | | |
|---|--------------|-----|
| 1 | Formato CSV | 112 |
| 2 | Formato XML | 101 |
| 3 | Formato JSON | 102 |
| 4 | Formato RDF | 22 |
| 5 | Formato PDF | 1 |
| 6 | Formato OWL | 1 |
| 7 | Formato ZIP | 3 |
| 8 | Formato HTML | 1 |

Fonte: Dados da pesquisa (2021).

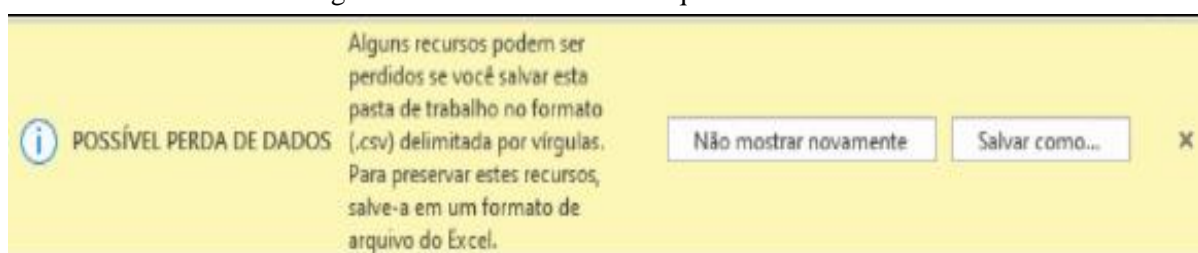
É importante registrar que esses números acima fazem parte do universo dos 119 conjuntos de dados e que a maior parte está disponibilizado em mais de um formato, o que justifica a variação da quantidade ser maior do que o total de 119 conjuntos. Observa-se que

os formatos mais reconhecidos para publicação de dados abertos são utilizados na previdência social brasileira, são eles:

- *CSV - Comma-Separated Values*, formato para armazenamento de dados tabulares em texto, cada linha do arquivo representa uma linha na tabela, e as colunas são separadas por vírgula;
- *XML - Extensible Markup Language*, conjunto de regras que codificam documentos com estrutura hierárquica e em um formato legível por máquina;
- *JSON - acrônimo para JavaScript Object Notation*, padrão aberto de estruturação de dados baseado em texto e legível por humano;
- *RDF - Resource Description Framework*, modelo de dados estruturado em gráfos que possuem diversos formatos de serialização.

Estes formatos são um aspecto de suma importância para tornar os dados reusáveis, por possibilitarem o processamento por sistemas computacionais. Além disso, disponibilizar dados em mais de um formato possibilita atender a uma maior gama de usuários e tornar os dados, possivelmente, mais inteligíveis por homens e máquinas. Outro problema identificado com relação aos formatos foi que alguns conjuntos de dados ao serem abertos em determinados formatos emitiam alertas sobre possível perda de dados conforme pode ser visto na Figura 14. O que viola o critério da **acessibilidade**.

Figura 14 – Alerta ao abrir o arquivo de dados.



Fonte: Dados da pesquisa (2021).

Finalmente, **com relação às licenças de uso**, diversos conjuntos de dados não as especificam (Figuras 8, 9, 10 e 11). Quando especificada, verifica-se que se optou por uma licença adequada (*Creative Commons*), que permite o acesso, uso e derivação, contanto que se faça menção à fonte original.

Ressalta-se que estes problemas identificados prejudicam a transparência das informações da administração pública e a oferta de dados que sejam acessíveis, gratuitos, completos e compreensíveis.

4.2 Análise de qualidade dos dados e da maturidade de abertura dos dados

Um resumo da avaliação utilizando o protocolo de avaliação definido nos procedimentos metodológicos pode ser visualizado no Quadro 11. Neste quadro a primeira coluna apresenta os critérios estabelecidos nos procedimentos metodológicos. Cada critério é categorizado com ATENDE, NÃO ATENDE, ATENDE PARCIALMENTE. A exceção é o critério Maturidade cujo conteúdo é expresso pela quantidade de estrelas. E, na última coluna encontram-se observações sobre a análise realizada.

Quadro 11 – Resultado da avaliação dos conjuntos de dados com base na junção dos critérios de avaliação

| CRITÉRIOS | AVALIAÇÃO | OBSERVAÇÕES |
|-----------------------------------|------------------------------|--|
| Primários | ATENDE | Em uma primeira análise os dados são disponibilizados com alta granularidade, sem agregação ou modificação. |
| Acessível | ATENDE PARCIALMENTE | Os dados disponibilizados nem sempre podem ser acessados, há diversos links que levam à uma tela de erro. |
| Credibilidade | NÃO PÔDE SER AVALIADO | Não foi possível fazer uma validação dos dados para verificar o quanto eles são verdadeiros. |
| Completos | ATENDE PARCIALMENTE | Alguns conjuntos de dados são incompletos, e deficientes na representação por metadados. Ao abrir alguns conjuntos de dados é informado que há perda de dados. |
| Não-discriminatórios | ATENDE | Os dados estão disponíveis online sem exigência de solicitação ou cadastro prévio. |
| Concisão | NÃO ATENDE | Os dados não possuem metadados suficientes para sua representação de forma clara, inequívoca e congruente. |
| Processamento por Máquinas | ATENDE PARCIALMENTE | Alguns conjuntos de dados são disponibilizados em formatos não legíveis por máquinas. |
| Interpretabilidade | ATENDE | Os dados estão em linguagem clara, porém, faltam |

| | | |
|---------------------------|----------------------------|--|
| | PARCIALMENTE | metadados que possam contribuir com a sua compreensão de forma a poder ser usado, interpretado e reutilizado. Não foi identificado dicionário de dados para os conjuntos de dados. |
| Livre de licenças | ATENDE | Não há restrição de direito autoral, patente, propriedade intelectual ou segredo industrial. Porém, ressalta-se que muitos conjuntos de dados não tem explicitado o tipo de licença. |
| Não Proprietários | ATENDE | Os conjuntos de dados são apresentados em formato aberto. |
| Atualidade | ATENDE PARCIALMENTE | Em alguns conjuntos de dados não é possível verificar se há atualização periódica. E outros conjuntos de dados estão desatualizados. Inclusive a instituição encontra-se sem um PDA válido. |
| Encontrabilidade | ATENDE PARCIALMENTE | Nem todos os conjuntos de dados possuem identificadores persistentes que permitam identificar inequivocamente cada um deles. |
| Reusabilidade | ATENDE PARCIALMENTE | Os dados estão em linguagem clara, porém, faltam metadados que possam contribuir com a sua compreensão de forma a poder ser usado, interpretado e reutilizado. E, apesar de estarem em formatos abertos e livres de licenças, a definição do tipo de licença não é explicitada em todos os conjuntos de dados, o que pode deixar dúvidas com relação ao uso dos dados. |
| Interoperabilidade | ATENDE PARCIALMENTE | Alguns formatos utilizados trazem dificuldade para interoperar entre bases de dados. O fato de não haver dicionário de dados ou metadados elucidando os campos também dificultam a interoperabilidade. |
| Maturidade | Entre ** e *** | Os dados estão disponíveis na internet, sob licença aberta, em formato aberto e não proprietário, porém, nem sempre estruturados e nem sempre em formatos legíveis por máquinas. Os conjuntos de dados, em sua maioria, carecem de identificação persistente e não há interligação entre os conjuntos de dados ou deles com outros conjuntos de outras instituições. Logo, em termos de maturidade de abertura de dados, os dados da previdência podem ser enquadrados entre 2 e 3 estrelas. |
| Consistência | ATENDE PARCIALMENTE | Há diversidade de formatos de arquivos, assim como também há diversidade de campos de metadados |

| | | |
|--------------------------------------|-------------------|---|
| | | utilizados entre os conjuntos e ainda campos de mesmo nome com informações diferentes e conflitantes. |
| Facilidade de Uso por humanos | NÃO ATENDE | Devido a forma de apresentação dos dados e à falta do dicionário de dados e incompletude dos metadados, não é trivial para os humanos compreenderem o que está sendo disponibilizado. |

Fonte: O autor (2021).

Ressaltam-se que dois dos principais problemas encontrados e que têm impacto em vários critérios é o problema da descrição limitada, inconsistente ou ausente dos conjuntos de dados e a falta de acesso à alguns dos conjuntos de dados, cujos links levam a páginas de erro.

Os resultados apresentados no Quadro 11 apontam para a inconformidade dos conjuntos de dados da previdência social com relação aos padrões existentes para publicação de dados, visto que a maioria dos critérios são atendidos apenas parcialmente. Isto pode trazer prejuízo para o cidadão que deseje se manter informado sobre o conteúdo desses conjuntos de dados ou deseje fazer uso deles para as mais diversas finalidades. Também dificulta para empresas, instituições e ONGs fazerem uso destes dados, tanto para fins informativos como para a construção de produtos sobre eles, tais como jogos, aplicativos, plataformas de visualização de dados, entre outros.

5 CONSIDERAÇÕES FINAIS

A disponibilização dos dados governamentais e seu acesso facilitado, além de atender a uma questão legal, é essencial para uma sociedade mais democrática, inclusiva e participativa. Porém, está pode ainda ser considerada uma nova e desafiadora tarefa para grande parte dos órgãos governamentais (SILVA, 2013), especialmente considerando a falta de pessoal treinado para viabilizá-la e, muitas vezes, também de infraestrutura. Talvez, por isso, por meio dos resultados obtidos pela presente pesquisa, foi possível verificar que os conjuntos de dados da previdência social, publicados no portal brasileiro de Dados Abertos, entre os anos de 2019 e 2021, em sua maioria, atendem parcialmente aos critérios de qualidade e maturidade estabelecidos.

A maturidade de abertura dos dados fica entre 2 e 3 estrelas, ficando abaixo do recomendado por Isotani e Betencourt (2015) de que a abertura de dados deve ser feita trabalhando no contexto de pelo menos três estrelas.

Também, verifica-se que, com relação ao trabalho de FRANZOSI *et al.* (2011), houve uma evolução na quantidade de dados sendo compartilhados/disseminados. Porém, apesar de existirem critérios, princípios e recomendações para a publicação de dados abertos, ainda é fato que não existe uma forma padrão e consistente para publicação dos DAG pelos órgãos. Assim, a produção e publicação de dados e informações é realizada, ainda, de maneira isolada e não interoperável pelas mais diversas instituições, públicas e privadas. Adicionalmente, verificam-se que o problema de uso correto e eficiente de metadados, de estabelecimento de formatos adequados de publicação e do preenchimento correto das informações sobre os dados (metadados) é ainda um problema em boa parte dos conjuntos de dados analisados.

Dessa forma, a maneira como os dados são disponibilizados, mostra-se ainda inadequada para sanar as possíveis demandas da sociedade, podendo gerar dúvidas a quem deles necessite. As barreiras de disponibilização, acesso, uso e reuso vão desde os formatos, linguagem até a maneira de apresentar a informação, que, teoricamente, deveria atender a filosofia de que os dados e/ou informações no contexto do Governo aberto devem ser acessíveis, para o maior número de pessoas.

Adicionalmente, além das questões de qualidade e maturidade que implicam diretamente no uso e reuso desses dados, foram observadas dificuldades relativas a quem é o mantenedor dos conjuntos de dados, quando foi criado e qual foi a última data de atualização, sendo essa última um dado crítico para saber se os valores obtidos com os dados são atuais ou

não. Os problemas encontrados requerem atenção da administração pública, pois violam princípios chave do governo aberto: a transparência, a *accountability* e a participação social.

Neste ponto identifica-se a possibilidade de atuação do profissional Gestor da Informação, pois o mesmo pode atuar na identificação dos problemas e apresentação de possíveis alternativas de resolução para a publicação de dados, em especial podendo atuar na descrição dos dados a serem publicizados. Outra possibilidade é a atuação na definição de ferramentas, sites e aplicativos para apresentar a visualização dos dados de forma mais amigável à população em geral. Visto que os conjuntos de dados disponibilizados são grandes agrupamentos numéricos de difícil compreensão e manipulação, se a pessoa não possuir experiência com recursos tecnológicos, em especial com o manuseio de planilhas de dados.

Ademais, em virtude dos desafios associados à efetivação do direito de acesso à informação, há que se considerar que o acesso à rede mundial de computadores ainda não alcançou toda a população brasileira. Por exemplo, pessoas que vivem nas áreas mais remotas do país, não dispõem de uma infraestrutura necessária para acessar os benefícios advindos das novas tecnologias do governo eletrônico, bem como participar das ações do estado. Assim, por um lado é inegável o papel das TIC no processo de fortalecimento da transparência e da participação social. Mas, por outro lado, apesar de o número de pessoas com acesso à internet ter crescido, ano a ano, as desigualdades de acesso ao mundo digital persistem no país, tendo reflexo direto no processo de participação social (BRASIL, 2021)

De fato, o movimento de abertura de dados e acesso à informação somente será completo, quando os dados e/ou conjuntos de dados cumprirem com sua finalidade e estiverem em conformidade com as legislações e políticas de dados abertos às quais estão submetidos. Ademais, é necessária uma infraestrutura técnica e de pessoal para que um DAG consiga alcançar as cinco estrelas. As instituições que não dispõem desses elementos e que não possuem experiência deparam-se como uma tarefa complexa a cumprir para publicar dados desta forma (NASCIMENTO-SILVA; PINHEIRO, 2019). Além disso, nossas legislações que surgem podem trazer novos desafios, ainda não muito bem definidos, para a publicação dados em formato aberto, como, por exemplo, o impacto que a Lei nº 13.709 - Lei Brasileira de Proteção de Dados Pessoais (LGPD) (BRASIL, 2018), uma legislação recente, que impacta sobre a publicação de dados.

Considera-se que os objetivos da pesquisa foram alcançados e espera-se como trabalho futuro aplicar o protocolo de avaliação criado a outros conjuntos de dados de forma a melhor validá-lo, assim como incorporar ao mesmo elementos referentes à LGPD, tal como a privacidade de dados.

REFERÊNCIAS

AGUNE, R. M.; GREGORIO FILHO, A. S.; BOLLIGER, S. P. Governo aberto SP: disponibilização de bases de dados e informações em formato aberto. *In: CONGRESSO CONSAD DE GESTÃO PÚBLICA*, 3., 2010, Brasília. **Anais** [...]. Brasília: [s. n.], 2010.

ALMARABEH, T.; ABUALI, A. A general framework for e-government: definition maturity challenges, opportunities, and success. **European Journal of Scientific Research**, [s. l.], v. 39, n. 1, p. 29-42, 2010.

ATTARD, J.; ORLANDI, F.; SCERRI, S.; AUER, S. A systematic review of open government data initiatives. **Government Information Quarterly**, [s. l.], v. 32, n. 4, p. 399-418, 2015. Disponível em: <https://doi.org/10.1016/j.giq.2015.07.006>. Acesso em: 13 dez. 2021.

BELLIX, L.; BURLE S. GUIMARÃES, C.; MACHADO, J. Qual conceito de Governo Aberto? Uma aproximação aos seus princípios. **GIGAPP Estudos Working Papers**, [s. l.], v. 4, n. 55-59, p. 78-91, 22 abr. 2017. Disponível em: <http://www.gigapp.org/ewp/index.php/GIGAPP-EWP/article/view/43>. Acesso em: 10 out. 2021.

BEGHIN, N., ZIGONE, C. (org.). **Avaliando os websites de transparência orçamentária nacionais e sub-nacionais e medindo impactos de dados abertos sobre direitos humanos no Brasil**. Brasília: Instituto de Estudos Socioeconômicos, 2014.

BID. 2016. **Reform / modernization of the state. Effective, efficient and open governments for the region**. BID's Website. Disponível em: <http://www.iadb.org/en/sector/reform-modernizationof-the-state/overview,18347.html> Acesso em: 17 out. 2021.

BERNERS-LEE, T. **Is your Linked Open Data 5 Star?** 2010. Disponível em: <http://www.w3.org/DesignIssues/LinkedData>. Acesso em: 13 out. 2021.

BERTOT, J. C. *et al.* Big Data, Open Government and E-Government: Issues, Policies and Recommendations. **Information Polity**, [s. l.], v. 19, n. 1-2, p. 5-16, 2014.

BRASIL. **Acesso à internet cresceu nos lares brasileiros**. 2021. Disponível em: <https://www.gov.br/pt-br/noticias/transito-e-transportes/2021/04/acesso-a-internet-cresceu-nos-lares-brasileiros>. Acesso em: 18 out. 2021.

BRASIL. **Cartilha Técnica para Publicação de Dados Abertos no Brasil v1.0**. 2012. Disponível em: <https://dados.gov.br/pagina/cartilha-publicacao-dados-abertos>. Acesso em: 18 dez. 2021.

BRASIL. **Constituição (1988)**. Constituição da República Federativa do Brasil. Texto constitucional promulgado em 5 de outubro de 1988, com as alterações adotadas pelas Emendas Constitucionais no 1/1992 a 92/2016, pelo Decreto Legislativo no 186/2008 e pelas Emendas Constitucionais de Revisão no 1 a 6/1994. Disponível em: http://www.planalto.gov.br/ccivil_03/Constituicao/Constituicao.htm . Acesso em 07 nov. 2021.

BRASIL. Ministério da Economia. **Do Eletrônico ao Digital**. Brasília, DF: Ministério da Economia, 2019. Disponível em: <https://www.gov.br/governodigital/pt-br/estrategia-de-governanca-digital/do-eletronico-ao-digital>. Acesso em: 25 out. 2021.

BRASIL. Controladoria- Geral da União. **O que é Governo Aberto**: Saiba o conceito de governo aberto. 2020. Disponível em: <https://www.gov.br/cgu/pt-br/governo-aberto/governo-aberto-no-brasil/o-que-e-governo-aberto>. Acesso em: 11 nov. 2021.

BRASIL. Decreto nº 8.777, de 11 de maio de 2016. Institui a Política de Dados Abertos do Poder Executivo federal. **Diário Oficial da União**, Atos do Poder Executivo, Brasília, DF, 12 mai. 2016. Seção 1, p. 21-22. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2016/decreto/d8777.htm. Acesso em: 7 out. 2021.

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. **Diário Oficial da União**, Atos do Poder Legislativo, Brasília, DF, 18 nov. 2011. Seção 1, Edição Extra, 264p. Disponível em: http://legislacao.planalto.gov.br/legisla/legislacao.nsf/Viw_Identificacao/lei%2012.5272011?OpenDocument. Acesso em: 5 out. 2021.

BRASIL. Lei nº 13.709, de 14 de agosto de 2018. Lei Geral de Proteção de Dados Pessoais: (LGPD). Brasília, Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/113709.htm. Acesso em: 20 dez. 2021

BRASIL. **Decreto 7.724/2012, de 16 de maio de 2012**. Regulamenta a Lei no 12.527, de 18 de novembro de 2011, que dispõe sobre o acesso a informações previsto no inciso XXXIII do caput do art. 5º, no inciso II do § 3º do art. 37 e no § 2º do art. 216 da Constituição. **Diário Oficial da União**, Brasília, DF, 16 de maio 2012.

CETIC. Pesquisa sobre o uso das Tecnologias de Informação e Comunicação - TIC Governo Eletrônico 2013. Cetic.br - Centro Regional para o Desenvolvimento da Sociedade da Informação, 26 Fev. 2015. Disponível em: . Acesso em: 07 out. 2021.

CHUN, S. A.; SHULMAN, S.; SANDOVAL, R.; HOVV, E. Government 2.0: Making connections between citizens, data and government. **Information Polity**, [s. l.], v. 15, n. 1-2, p. 1-9, 2010.

CLARKE, A.; MARGETTS, H. Governments and Citizens Getting to Know Each Other? Open, Closed, and Big Data in Public Management Reform. **Policy & Internet**, [s. l.], v. 6, n. 4, 393–417, 2014. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1002/1944-2866.POI377>. Acesso em: 13 dez. 2021.

CONRADIE, P.; CHOENNI, S. On the barriers for local government releasing open data. **Government Information Quarterly**, [s. l.], v. 31, p. s10-s17, June, 2014. Suplemento. Disponível em: <http://dx.doi.org/10.1016/j.giq.2014.01.003>. Acesso em 22 de nov. 2021.

CORRÊA, A. S. et al. Transparency and open government data: a wide national assessment of data openness in Brazilian local governments. *Transforming Government: People, Process and Policy*, v. 11, n. 1, p. 58-78, 2017.

CULAU, A. A.; FORTIS, A. F. M. Transparência e controle social na administração pública brasileira: avaliação das principais inovações introduzidas pela Lei de Responsabilidade Fiscal. *In: CONGRESO INTERNACIONAL DEL CLAD SOBRE LA REFORMA DEL ESTADO Y DE LA ADMINISTRACIÓN PÚBLICA, CIUDAD DE GUATEMALA, 11., 2006, Guatemala. Anais [...]. Guatemala: [s. n.], 2006, p. 16.*

DAHL, R. A. **Poliarquia**: participação e oposição. São Paulo: EDUSP, 1997.

DATA FAIRPORT. Jointly designing a data FAIRPORT. 2014. Disponível em: <https://www.lorentzcenter.nl/lc/web/2014/602/info.php3?wsid=602>. Acesso em: 8 nov. 2021.

DAVIES, T., BAWA, Z. The promises and perils of open government data (OGD). **The Journal of Community Informatics**, Vancouver, v. 8, n. 2, p. 1-13, abr. 2012.

DAWES, S. S. Stewardship and usefulness: Policy principles for information-based transparency. **Government Information Quarterly**, [s. l.], v. 27, n. 4, p. 377-383, 2010.

EAVES, D. **The Three Laws of Open Government Data**. 2010. Disponível em: <http://eaves.ca/2009/09/30/three-law-of-open-government-data/>. Acesso em: 19 out. 2020.

EPP – Escola de Políticas Públicas. Dados abertos conectados. Disponível em: <http://www.ep.org.br/dados-abertos-conectados> Acesso em: 07 out. 2021.

FIGUEIREDO, Vanuza da Silva; SANTOS, Waldir Jorge Ladeira dos. **Transparência e controle social na administração pública**. 2013. Disponível em: <https://www.fclar.unesp.br/Home/Departamentos/AdministracaoPublica/RevistaTemasdeAdministracaoPublica/vanuza-da-silva-figueiredo.pdf>. Acesso em: 20 nov. 2021.

FRANZOSI, E. M. *et al.* Rede de Informações Pública e Aberta - DadosGov - Um modelo aberto para a troca de informações. *In: DADOS abertos para a Democracia na Era Digital*. Brasília: Fundação Alexandre de Gusmão, 2011. p. 21-51.

GALIOTOU, E.; FRAGKOU, P. Applying Linked Data Technologies to Greek Open Government Data: A Case Study. *Procedia - Social and Behavioral Sciences*, [s. l.], v. 73, p. 479-486, 2013.

GEIGER, C. P.; LUCKE, J. V. Open government and (linked) (open) (government) (data). **JeDEM-eJournal of e Democracy and Open Government**, [s. l.], v. 4, n. 2, p. 265-278, 2012.

GO FAIR. **Global Open, Findable, Accessible, Interoperable, Reusable**. 2018. Disponível em: <https://www.go-fair.org/>. Acesso em: 9 out. 2021.

BRASIL. **Acesso à internet cresceu nos lares brasileiros**. 2021. Disponível em: <https://www.gov.br/pt-br/noticias/transito-e-transportes/2021/04/acesso-a-internet-cresceu-nos-lares-brasileiros>. Acesso em: 18 out. 2021.

HENNING, P. C.; RIBEIRO, C. J. S.; SANTOS, L. O. B. S.; SANTOS, P. X. D. Go fair e os princípios fair: o que representam para a expansão dos dados de pesquisa no âmbito da ciência aberta. **Em Questão**, v. 25, n. 2, p. 389-412, 2019. DOI: 10.19132/1808-5245252.389-412 Acesso em: 20 nov. 2021.

HODGE, G. **Metadata made simpler**: a guide for libraries. Bethesda: National Information Standards Organizations, 2001.

INDA - **Infraestrutura Nacional de Dados Abertos**. 2019. Disponível em: <http://wiki.dados.gov.br/>. Acesso em 07 out. 2021.

ISOTANI, S.; BITTENCOURT, I. I. **Dados abertos conectados**. São Paulo: Novatec, 2015.

JANSSEN, K. The influence of the PSI directive on open government data: An overview of recent developments. **Government Information Quarterly**, [s. l.], v. 28, n. 4, p. 446–456, 2011.

KLEIN, R H. **Mecanismos de Ampliação da Transparência em Portais de Dados Abertos Governamentais Brasileiros à Luz da Accountability Theory**. 2017. Tese (Doutorado em Administração) – Escola de Negócios, PUCRS, Rio Grande do Sul, 2017. Disponível em: <http://tede2.pucrs.br/tede2/handle/tede/7724>. Acesso em: 13 dez. 2021.

LINDERS, D. Towards open development: Leveraging open data to improve the planning and coordination of international aid. **Government Information Quarterly**, [s. l.], v. 30, n. 4, p. 426– 434, 2013.

LOURENÇO, R. P. An analysis of open government portals: A perspective of transparency for accountability. **Government Information Quarterly**, [s. l.], v. 32, n. 3, p. 323–332, 2015.

LORENZO, S.; CÉSAR, C. **Open government**: gobierno abierto. Jaén: Algón Editores MMX, 2010.

MALAMUD, C. *et al.* **8 Principles of Open GovernmentData**. December 2007. Disponível em: https://public.resource.org/8_principles.html. Acesso em: 9 out. 2021.

OGP - **Parceria para Governo Aberto**. Governo Federal. 2018. Disponível em: <http://www.governoaberto.cgu.gov.br/a-ogp/o-que-e-a-iniciativa> . Acesso em 07 out. 2020.

OPEN GOVERNMENT WORKING GROUP. Eight principles of open government data. [S. l.]: **Open Government Working Group**, 2007. Disponível em: <https://opengovdata.org/>. Acesso em: 19 out. 2020.

OPEN KNOWLEDGE FOUNDATION. Open data handbook. [on line]. Cambridge: [2014]. Disponível em: <http://opendatahandbook.org/guide/en/>. Acesso em: 07 out. 2021.

OPEN KNOWLEDGE INTERNATIONAL. **What is Open Data?** 2011. Disponível em: <https://okfn.org/opendata/>. Acesso em: 9 out. 2021.

PARYCEK, Peter; SACHS, Michael. Open government–information flow in Web 2.0. *European Journal of ePractice*, v. 9, n. 1, p. 1-70, 2010.

PRADO, O. **Governo eletrônico, reforma do Estado e transparência: o Programa de Governo Eletrônico no Brasil**. São Paulo: FGV, 2009. Tese (Doutorado em Administração Pública e Governo) – Escola de Administração de Empresas de São Paulo, FGV, São Paulo, 2009.

PIMENTA, M. S.; CANABARRO, D. R. (org.). **Governança Digital**. Porto Alegre: Ed. da UFRGS, 2014.

PIRES, Marco Túlio. Guia de Dados Abertos. São Paulo, 2015. Disponível em: <http://www.governoaberto.sp.gov.br/wp-content/uploads/2016/02/Book-Web-Guia-de-Dados-Abertos.pdf>. Acesso em: 08 set. 2017

POSSAMAI, A. J. **Dados abertos no governo federal brasileiro: desafios de transparência e interoperabilidade**. 2016. Tese (Doutorado em Desenvolvimento Rural) – Faculdade de Ciências Econômicas, UFRGS, Porto Alegre, 2016. Disponível em: <https://www.lume.ufrgs.br/handle/10183/15636>. Acesso em: 29 nov. 2021.

RAUPP, F. M.; PINHO, J. A. G. Accountability em câmaras municipais: uma investigação em portais eletrônicos. **Revista de Administração**, São Paulo, v. 48, n. 4, p. 770-782, out./nov./dez. 2013.

ROCHA, A.; VASCONCELOS, J. Os modelos de maturidade na gestão de sistemas de informação. **Edições Universidade Fernando Pessoa**, 2004.

SANCHEZ, C. S. Comportamento informacional e governo aberto: o “horizonte de informação” dos agentes da sociedade civil organizada brasileira. **Revista Brasileira de Biblioteconomia e Documentação**, [s. l.], v. 16, p. 1-26, 2020. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/150404>. Acesso em: 03 nov. 2021.

SANTANA, D. B.; SILVA, E. M.; SILVA, M. L. Transparência da informação e dados abertos dos governos municipais da região metropolitana do recife. **Em Questão**, Porto Alegre, v. 28, n. 1, jan./mar. 2022. DOI: <https://doi.org/10.19132/1808-5245281.128-158>. Disponível em: <https://www.seer.ufrgs.br/EmQuestao/article/view/114169>. Acesso em: 05 dez. 2021.

SANTOS, M. G. D. Estudo Empírico da relação entre Qualidade da Informação e Impacto Individual no Contexto Organizacional, 2009. 258p. Tese (Mestrado em Administração) - Faculdade de Economia, Administração e Contabilidade, Universidade de São Paulo, São Paulo, 2009.

SHANKARANARAYANAN, G.; EVEN, A. The metadata enigma. *Communications of the ACM*, ACM, v. 49, n. 2, p. 88–94, 2006.

SAYÃO, L. F.; SALES, L. F. **Afinal, o que é dado de pesquisa?. BIBLOS** - Revista do Instituto de Ciências Humanas e da Informação, v. 34, n. contexto, 2020. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/162776>. Acesso em: 20 nov. 2021.

SHERIDAN, John; TENNISON, Jeni. Linking UK Government Data. In: LDOW. 2010.

SILVA, N. Maturidade em Dados Abertos: entenda as 5 estrelas. **Open Knowledge Brasil**, 2013. Disponível em: <https://br.okfn.org/2013/01/17/maturidade-em-dados-abertos-entenda-as-5-estrelas/>. Acesso em: 10 out. 2021.

NASCIMENTO-SILVA, P.; KERR-PINHEIRO, M. M. Métrica alternativa para dados governamentais abertos na América Latina. **Transinformação**, [s. l.], v. 31, e190009, 2019. Disponível em: <http://dx.doi.org/10.1590/2318-0889201931e190009>. Acesso em: 10 out. 2021.

UBALDI, B. Open government data: Towards empirical analysis of open government data initiatives. **OECD Working Papers on Public Governance**, n. 22, 2013.

VELJKOVIĆ, N.; BOGDANOVIĆ-DINIĆ, S.; STOIMENOV, L. Benchmarking open government: an open data perspective. **Government Information Quarterly**, [s. l.], v. 31, n. 2, p. 278–290, 2014. DOI: <http://dx.doi.org/10.1016/j.giq.2013.10.011>. Disponível em: <https://isiarticles.com/bundles/Article/pre/pdf/44375.pdf>. Acesso em: 13 dez. 2021.

VILLAZON-TERRAZAS, B.; VILVHES-BLÁZQUEZ, L. M.; CORCHO-PEREZ, A. Methodological guidelines for publishing government linked data. In: **Linking Government Data**. [S.l.: s.n.], 2011. Cap.2, p. 27-49.

VICTORINO, M. C.; MARTINS, L.; HOLANDA, M.; FONSECA, R. Arquitetura de publicação de dados abertos conectados governamentais da universidade de Brasília. **Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação**, v. 25, p. 1-25, 2020. DOI: 10.5007/1518-2924.2020.e67665 Acesso em: 20 dez. 2021.

WILKINSON, M.D. et al. The FAIR Guiding Principles for scientific data management and stewardship. **Scientific Data**. Disponível em: <https://www.nature.com/articles/sdata201618> . Acesso em 07 nov. 2021.

WEINSTEIN, J.; GOLDSTEIN, J. The Benefits of a Big Tent: Opening Up Government in Developing Countries A Response to Yu & Robinson's The New Ambiguity of "Open Government". **UCLA Law Review Disclosure**, [s. l.], v. 60, n. 38, 2012 p. 38-48. Disponível em: <http://www.uclalawreview.org/pdf/discourse/60-3.pdf>. Acesso em: 20 nov. 2021.

WORLD WIDE WEB CONSORTIUM. Dados abertos governamentais. São Paulo: **W3C Brasil**, 2009. Disponível em: <http://www.w3c.br/divulgacao/pdf/dados-abertos-governamentais.pdf>. Acesso em: 9 out. 2021.