

DADOS ABERTOS GOVERNAMENTAIS avaliação de maturidade e qualidade no contexto da covid-19 em Pernambuco

Getúlio Valdemir Batista¹

Universidade Federal de Pernambuco
getulio.batista@ufpe.br

Sandra de Albuquerque Siebra²

Universidade Federal de Pernambuco
sandra.siebra@ufpe.br

Resumo

O objetivo geral deste artigo foi avaliar o grau de maturidade e a qualidade dos dados abertos disponibilizados pelo governo do Estado de Pernambuco, com foco na análise dos dados relacionados às compras e contratações decorrentes da Covid-19. Esta foi uma pesquisa descritiva e qualitativa, bibliográfica e documental, que fez uso do método estudo de caso nos conjuntos de dados relativos às compras e contratações decorrentes da Covid-19. Aponta como principais resultados que a situação dos conjuntos de dados atende apenas parcialmente aos critérios e padrões de dados abertos. Apresentando como principais problemas: a descrição limitada ou ausência de descrição dos conjuntos de dados; possíveis perdas de dados e a falta de consistência de algumas das informações apresentadas, o que pode causar problemas no acesso, interpretação, uso e reuso dos conjuntos de dados por quem deles necessite.

Palavras-chave: Acesso à Informação. Dados abertos. Dados Abertos Governamentais. Qualidade da Informação. Covid-19.

OPEN GOVERNMENT DATA

maturity and quality assessment in the context of covid-19 in Pernambuco

Abstract

The general objective of this article was to evaluate the degree of maturity and the quality of open data made available by the government of the State of Pernambuco, focusing on the analysis of data related to the procurement and contracting arising from Covid-19. This was a descriptive and qualitative, bibliographic and documental research, which made use of the case study method in the data sets related to the purchases and contracting arising from Covid-19. It points out as main results that the situation of the datasets only partially meets the criteria and standards of open data. Presenting as main problems: the limited or no description of the datasets; possible data losses and the lack of consistency of some of the presented information, which may cause problems in the access, interpretation, use and reuse of the datasets by those who need them.

Keywords: Access to Information; Open Data. Open Government Data. Quality of Information. Covid-19.

¹Mestrando do Pós-Graduação Programa em Ciência da Informação – PPGCI/UFPE. Recife, PE, Brasil.

²Pesquisadora e Docente do Programa de Pós-graduação em Ciência da Informação – PPGCI/UFPE. Recife, PE, Brasil.



1 INTRODUÇÃO

No decorrer do tempo, as práticas de democratização do acesso à dados e informações, no contexto de governo aberto³, tem permeado a comunicação entre o Estado e a sociedade e vêm ganhando cada vez mais notoriedade na gestão pública, especialmente depois da Lei Federal nº 12.557 de novembro de 2011, conhecida como Lei de Acesso à Informação - LAI (BRASIL, 2011). De fato, no contexto brasileiro, a LAI tornou-se um importante instrumento, uma vez que regulamentou o acesso à informação previsto no inciso XXXIII do art. 5º, no inciso II do § 3º art. 37 e no § 2º do art. 216 da Constituição Federal (BRASIL, 1988).

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 transparência⁴, a melhoria da participação dos cidadãos e da prestação pública de contas (*accountability*⁵), 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) de Galiotou e Fragkou (2013), que o definem como dado aberto⁶ 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). Considera-se que os portais de transparência e as informações disponibilizadas no contexto da transparência ativa, pregada pela LAI (BRASIL, 2011), nos sites das instituições, aproximou a relação entre o governo e o cidadão. E estas práticas se tornaram um meio de

³ O termo Governo Aberto agrupa um conjunto de diversas iniciativas objetivando conferir maior transparência, *accountability* e permeabilidade ao Estado

⁴ 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).

⁵ Nesta pesquisa será utilizado doravante o termo em inglês *accountability*.

⁶ 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).

divulgação, avaliação e controle das informações governamentais para o livre uso pela sociedade, além de uma forma de prestar contas das ações desenvolvidas pelo Estado, promovendo a *accountability*. O que endossa o afirmado por Dahl (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.

Nesse cenário, destaca-se que a disponibilização de dados abertos governamentais, deve abranger a publicação e disseminação de informações do setor público, em formatos compreensíveis logicamente por computadores e por seres humanos, não proprietários e livres de licenças que restrinjam sua reutilização (WORLD..., 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, facilitar a tomada de decisão e estimular e promover a participação social. 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. No entanto, são variadas as formas como os distintos órgãos lidam com a publicização 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 é realizada, muitas vezes, de maneira isolada e não interoperável pelas mais diversas instituições, públicas e privadas. Adicionalmente, há casos em que as informações não estão facilmente disponíveis para o uso pelos tomadores de decisão, tampouco para a população. O que já era pontuado por Franzosi *et al.* (2011) quando enfatizou que a realidade que se apresentava, na época, era de uma grande massa de dados e informações públicos não compartilhados, não tratados e/ou não disseminados. 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).

Dessa forma, a ausência de informação clara, atualizada, disponível em tempo real, é uma questão que preocupa não só os cientistas, mas também os políticos, empresários e a sociedade em geral. Essa preocupação ganhou destaque, a partir de meados de dezembro de

2019, após a China ter detectado um novo coronavírus (SARS-COV-2⁷), que ultrapassou fronteiras e tornou-se a maior crise de saúde no mundo em um século (LIMA *et al.*, 2020). Diante do cenário de emergência global em que o mundo se encontrou, uma das alternativas de enfrentamento requeria o acesso a dados e informações. Assim, vários governos lançaram portais de dados abertos sobre a Covid-19 (SILVA, 2020), a fim de informar a população sobre o avanço da doença; de promover transparência sobre o quantitativo de infectados, recuperados, mortos e vacinados; além de sobre o quantitativo de recursos públicos que estavam sendo direcionados para o enfrentamento da questão.

Desse modo, este trabalho tem como objetivo avaliar o grau de maturidade e a qualidade dos dados abertos disponibilizados pelo governo de Pernambuco, com foco nos dados relativos às despesas relacionadas à Covid-19, entre os anos 2020 e 2021. Especificamente, buscou-se identificar e mapear as bases de dados portais e/ou repositórios oficiais do estado de Pernambuco, limitando a análise, em sete conjuntos de dados, a fim de entender a dinâmica de disponibilização, problemas e lacunas existentes. Destaca-se que 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.

Esta pesquisa se faz relevante porque, 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 (PARYCEK; SACHS, 2010). Por isso, a forma que os DAG são disponibilizados é importante e precisa atender a diversos princípios e recomendações, de forma a estes poderem ser efetivamente descobertos, usados e reusados.

Como uma das contribuições da pesquisa foi delineada um protocolo de avaliação de maturidade e qualidade que poderá ser aplicada à conjuntos de dados (*datasets*), assim como foi realizado com os DAG relacionados às despesas realizadas no contexto da pandemia da Covid-19, nesta pesquisa.

⁷ O coronavírus da síndrome respiratória aguda grave 2 - (SARS-COV-2) causador da atual pandemia de Covid-19. É uma família de vírus que causam infecções respiratórias e apresenta um quadro clínico que varia de infecções assintomáticas a quadros respiratórios graves.

2 PROCEDIMENTOS METODOLÓGICOS

Trata-se de uma pesquisa de abordagem qualitativa; que é bibliográfica e documental quanto os procedimentos técnicos e descritiva quanto aos objetivos (MICHEL, 2009). A pesquisa fez uso do método do estudo de caso (MICHEL, 2009), onde os conjuntos de dados trabalhados no contexto da pesquisa foram os dados das Despesas, Contratações e Aquisições decorrentes da COVID-19 no Estado de Pernambuco.

Para a revisão bibliográfica foram realizados levantamentos de artigos, livros, teses e dissertações em bases de dados como: BRAPCI, *Google Scholar* e *SciELO*, para entender a origem, o conceito de governo aberto e de dados abertos governamentais. Posteriormente, realizou-se a análise documental de algumas legislações e materiais não convencionais que foram relevantes para complementar o embasamento sobre a temática sendo pesquisada.

Os dados analisados foram recuperados por meio do portal transparencia.pe.gov.br, totalizando sete conjuntos de dados como amostra. O portal da transparência fornece acesso aos painéis de despesas da COVID-19, os processos de contratações e demais informações relacionadas ao coronavírus e ao combate à Covid19 no estado de Pernambuco. Assim o universo dos conjuntos de dados compreende: Contratações emergenciais, despesas detalhadas, despesas por item de material, remuneração e encargos custeados pela União, demonstrativos da aplicação de recursos recebidos da União.

Foram coletados, de forma *ad-hoc*, 7 (sete) conjuntos de dados publicados na forma coletada na fonte, e, posteriormente, foi criada uma planilha no software Excel para organizar os conjuntos de dados de acordo com: A) Título dos conjuntos de dados; B) Autores; C) Formatos; D) Metadados; E) Dicionário de dados; F) Data de atualização.

Para a avaliação de qualidade e maturidade dos dados recuperados foi aplicado um protocolo de avaliação que tomou como base os 8 (oito) princípios de qualidade que norteiam a publicação de dados (Quadro 1). Assim como os princípios FAIR, que estabelecem que os dados devem ser Encontráveis (Findable), Acessíveis (Accessible), Interoperáveis (Interoperable) e Reutilizáveis (Reusable) (GO FAIR, 2018); e o modelo cinco estrelas de Tim Berners-Lee (BERNERS-LEE, 2009; SILVA, 2013), 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.

Quadro 1 - 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).

Durante as primeiras análises, verificou-se a necessidade de incluir no protocolo de avaliação mais três critérios a respeito da qualidade dos dados, são eles: **Credibilidade** (o quanto 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. O protocolo de avaliação está resumido no Quadro 2.

Quadro 2 - Protocolo de avaliação da pesquisa

| ITEM | CRITÉRIO | ORIGEM DO CRITÉRIO |
|------|----------------------|---|
| 01 | PRIMÁRIOS. | Princípios para DAGs. |
| 02 | ACESSÍVEL | Princípios para DAGs e Princípios <i>FAIR</i> . |
| 03 | CREDIBILIDADE | Qualidade dos dados. |
| 04 | COMPLETOS | Princípios para DAGs. |
| 05 | NÃO-DISCRIMINATÓRIOS | Princípios para DAGs. |
| 06 | CONCISÃO | Princípios <i>FAIR</i> . |

| | | |
|----|--------------------------------|---|
| 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 FAIR. |
| 13 | REUSABILIDADE | Princípios FAIR. |
| 14 | INTEROPERABILIDADE | Princípios FAIR. |
| 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.

3 RESULTADOS E DISCUSSÕES

Os setes conjuntos de dados analisados revelaram informações acerca de quais conjuntos constam com atualização periódica, quais deles dispõem de metadados adequados, capazes de fornecer informação sem dubiedade para o usuário, quais possuem a presença de dicionários de dados, quais formatos são utilizados e a atualidade dos dados disponibilizados. Nesse sentido, por meio do Quadro 3 é possível perceber que nenhum dos sete conjuntos de dados recuperados e analisados possui metadados, o que, por si só, torna difícil o processo de uso e reuso dos dados.

Quadro 3 - Situação geral dos sete conjuntos de Dados Abertos sobre compras e contratações decorrentes de Covid-19 em Pernambuco.

| | Possuem Metadados | Tem Dicionário de Dados | Estão em Formato Adequado | Apresenta Data de Criação | Apresenta Data de Atualização |
|-----------------------------------|-------------------|-------------------------|---------------------------|---------------------------|-------------------------------|
| Qtd. De Conjuntos de Dados | 0 | 1 | 3 | 2 | 1 |

Fonte: Dados da Pesquisa (2021)

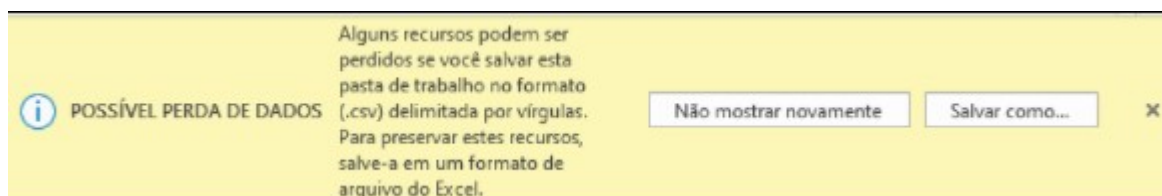
Além disso, levando em consideração a Infraestrutura Nacional de Dados (INDA) essa situação é ainda mais crítica pois, segundo ela, é orientado que, quanto mais metadados forem disponibilizados, mais eles agregam valor ao conjunto de dados, além de melhorar sua classificação e a busca sobre ele. Diante do exposto, os conjuntos de dados 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.

Em relação aos formatos disponibilizados, os dados são publicizados em três formatos (CSV/HTML/PDF). Observa-se que isso viola o critério da **consistência**, pois formatos como o PDF pode trazer dificuldade para a **leitura por máquinas**. 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. Além disso, apenas dois conjuntos de dados possuem data de criação e somente um indica sua data de atualização o que implica na violação do critério da **atualidade**, pois o usuário dos dados não terá certeza sobre se os dados estão atualizados ou não, se estão em sua versão mais recente.

Alguns *datasets* ao serem abertos emitiram alertas sobre possível perda de dados devido ao formato em que foram disponibilizados, conforme a Figura 1.

Figura 1: Alerta ao abrir o arquivo de dados.



Fonte: Os autores (2021).

Vale salientar que foi encontrado apenas um dicionário de dados para um único conjunto de dados, o que dificulta por exemplo, a compreensão dos nomes das colunas da extensa planilha em Excel que é produzida a partir dos demais conjuntos de dados publicados no formato CSV. Isso, é um empecilho, uma vez que a publicação de dados governamentais tem por objetivo aumentar a transparência das informações da administração pública e garantir que todo cidadão tenha acesso a dados que sejam acessíveis, gratuitos, completos e compreensíveis.

Um resumo da avaliação dos conjuntos de dados com base nos critérios do protocolo de avaliação apresentado nos procedimentos metodológicos pode ser visualizado no Quadro 4.

Quadro 4 – Avaliação dos conjuntos de dados com base no protocolo de avaliação da pesquisa.

| ITEM | CRITÉRIO | AValiação com relação ao critério |
|------|----------------------------|---|
| 1 | PRIMÁRIOS. | Os dados são disponibilizados com alta granularidade, sem agregação ou modificação. |
| 2 | ACESSÍVEL. | Os conjuntos de dados disponibilizados nem sempre podem ser acessados, ao abrir o arquivo é informado que há perda de dados. |
| 3 | CREDIBILIDADE | Não foi possível fazer cruzamento de dados para verificar o quanto os dados são verdadeiros. |
| 4 | COMPLETOS | Alguns conjuntos de dados são incompletos, sem metadados significativos. |
| 5 | NÃO-DISCRIMINATÓRIOS | Os dados estão disponíveis sem necessidade de solicitação ou cadastro prévio. |
| 6 | CONCISÃO | Alguns conjuntos de dados não possuem metadados para sua representação. |
| 7 | PROCESSAMENTO POR MÁQUINAS | Ao abrir alguns conjuntos de dados é informado que há perda de dados. |
| 8 | INTERPRETABILIDADE | Em alguns conjuntos de dados, faltam metadados e dicionários de dados que possam contribuir com a sua compreensão. |
| 9 | LIVRE DE LICENÇAS | Não há restrição de direito autoral, patente, propriedade intelectual ou segredo industrial |
| 10 | NÃO PROPRIETÁRIOS | Os conjuntos de dados são apresentados em formato aberto. |
| 11 | ATUALIDADE | Em alguns conjuntos de dados não é possível verificar se há atualização periódica. |
| 12 | ENCONTRABILIDADE | Os conjuntos de dados não possuem identificadores persistentes. |
| 13 | REUSABILIDADE | Apesar de estarem em formatos abertos e livres de licenças, porém, faltam metadados que possam contribuir com a sua compreensão. |
| 14 | INTEROPERABILIDADE | O fato de não haver dicionário de dados ou metadados elucidando os campos evidencia problemas de interoperabilidade. |
| 15 | MATURIDADE | Apesar de estarem em licença aberta, formato aberto e não proprietário, os conjuntos de dados carecem de identificadores persistentes para que possam ser interligados com outros conjuntos de dados. |
| 16 | CONSISTÊNCIA. | Os conjuntos de dados analisados estão limitados a |

| | | |
|----|--------------------------------|--|
| | | quatro tipos de formatos, não há metadados na grande maioria. |
| 17 | FACILIDADE DE USO POR HUMANOS. | Devido à forma de apresentação dos dados e à falta do dicionário de dados e dos metadados, não é trivial para os humanos compreenderem os arquivos de dados. |

Fontes: Os autores (2021)

Um detalhamento do Quadro 4, apontando a situação de cada critério com relação a cada um dos conjuntos de dados pode ser vista na Figura 2. Cada critério pode ser pontuado de 0 a 4, indicando o quanto o conjunto atende ou não ao critério, sendo 0 não atender ao critério e 4 atender completamente ao critério.

Figura 2 - Resultado da avaliação de cada conjuntos de dados com relação aos critérios de avaliação.

| CRITÉRIOS | CONJ 01 | CONJ 02 | CONJ 03 | CONJ 04 | CONJ 05 | CONJ 06 | CONJ 07 |
|-------------------------------|---------|---------|---------|---------|---------|---------|---------|
| Primários | 4 | 4 | 4 | 4 | 0 | 0 | 0 |
| Acessível | 3 | 3 | 3 | 4 | 3 | 3 | 3 |
| Credibilidade | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Completos | 3 | 3 | 3 | 3 | 1 | 1 | 1 |
| Não-discriminatórios | 4 | 4 | 4 | 4 | 4 | 4 | 4 |
| Concisão | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Processamento por Máquinas | 3 | 3 | 3 | 4 | 4 | 4 | 4 |
| Interpretabilidade | 3 | 3 | 3 | 3 | 0 | 0 | 0 |
| Livre de licenças | 4 | 4 | 4 | 4 | 4 | 4 | 4 |
| Não Proprietários | 4 | 4 | 4 | 4 | 4 | 4 | 4 |
| Atualidade | 1 | 0 | 0 | 1 | 0 | 0 | 0 |
| Encontrabilidade | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Reusabilidade | 3 | 1 | 3 | 3 | 1 | 1 | 1 |
| Interoperabilidade | 3 | 1 | 3 | 1 | 1 | 1 | 1 |
| Maturidade | 3 | 1 | 3 | 3 | 1 | 1 | 1 |
| Consistência | 1 | 1 | 3 | 3 | 1 | 1 | 1 |
| Facilidade de Uso por humanos | 1 | 1 | 3 | 3 | 1 | 3 | 1 |

Fonte: Os autores (2021)

É possível verificar a inconformidade dos conjuntos de dados analisados quanto aos padrões existentes para publicação de dados, visto que diversos critérios são atendidos apenas regularmente ou de forma insuficiente. Os critérios mais críticos ficam em torno daqueles que impactam o uso e reuso dos dados tais como: **Consistência, Atualidade, Facilidade de uso por humanos, Credibilidade e Encontrabilidade**. 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.

Neste sentido, consideram-se que dois dos principais problemas que mais impactaram sobre os critérios foram: o problema da descrição limitada ou ausência de descrição dos conjuntos de dados, além dos tipos de formatos em que são disponibilizados, que podem trazer dificuldades para a **leitura por homens ou máquinas**.

Observa-se, a partir das informações obtidas sobre os dados, que 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.

3 CONCLUSÃO

Por meio dos resultados obtidos, verifica-se que 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, não alcançando o principal objetivo, que é permitir ao cidadão realizar o controle social e participar ativamente das ações do estado.

Verificou-se que alguns conjuntos de dados disponíveis para visualização ainda oferecem, inclusive, barreiras técnicas para serem reutilizados pelos cidadãos para criação de novos projetos, aplicativos e serviços. Destaca-se principalmente que torna-se difícil utilizar responsabilmente 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; e sobre como coletá-los. Por isso, além dos dados, é necessária uma descrição abrangente e significativa deles, o que pode ser feito por meio de metadados (SHANKARANARAYANAN; EVEN, 2006).

Constatou-se, também, 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 é adotada uma forma padrão e consistente para publicação dos DAG pelos órgãos; que há problemas de uso correto e eficiente de metadados; não há o estabelecimento de formatos adequados de publicação; e nem o preenchimento correto das informações sobre os dados (metadados) em boa parte dos conjuntos de dados analisados. Além disso, foi observada a ausência de informações tais como quem é o mantenedor dos conjuntos de dados, quando o conjunto foi criado e qual foi a última data de atualização. Todos os problemas identificados implicam diretamente no grau de maturidade e qualidade dos dados.

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. Pois, 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 com uma tarefa complexa a cumprir para publicar dados da forma adequada (SILVA; PINHEIRO, 2019). Adicionalmente, legislações que surgem podem trazer novos desafios para a publicação de dados em formato aberto, como, por exemplo, a Lei nº 13.709 - Lei Brasileira de Proteção de Dados Pessoais (LGPD), (BRASIL, 2020), uma legislação recente que entrou em vigor em agosto de 2021, afeta a definição dos dados que podem ou não ser publicizados.

REFERÊNCIAS

ATTARD, J.; ORLANDI, F.; SCERRI, S.; AUER, S. A systematic review of open government data initiatives. **Government Information Quarterly**, Oxford; New York, v. 32, n. 4, p. 399-418, 2015. Disponível em: <https://doi.org/10.1016/j.giq.2015.07.006>. Acesso em: 10 out. 2022.

BERNERS-LEE, T. **Is your Linked Open Data 5 Star**. 2009. Disponível em: <http://www.w3.org/DesignIssues/LinkedData>. Acesso em: 02 ago. 2022.

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, 264 p. Disponível em: http://legislacao.planalto.gov.br/legisla/legislacao.nsf/Viw_Identificacao/lei%2012.5272011?OpenDocument. Acesso em: 02 ago. 2022.

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 ago. 2022.

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.

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**, [S. l.], v. 11, n. 1, p. 58-78, 2017.

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

EAVES, David. **The Three Laws of Open Government Data**. Disponível em: <http://eaves.ca/2009/09/30/three-law-of-open-government-data/>. Acesso em: 19 out. 2022.

FERNANDES, J. L. F.; CORDEIRO, D. F. Avaliação de formatos de publicação de dados abertos governamentais através de indicadores de usabilidade. **Tendências da Pesquisa Brasileira em Ciência da Informação**, Brasília, v. 9, n. 1, 2016. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/119608>. Acesso em: 02 ago. 2022.

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

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.

GO FAIR. **Global Open, Findable, Accessible, Interoperable, Reusable**. 2018. Disponível em: <https://www.go-fair.org/>. Acesso em: 03 ago. 2022.

INDA - **Infraestrutura Nacional de Dados Abertos**. 2019. Disponível em: <http://wiki.dados.gov.br/>. Acesso em 06 ago. 2022.

ISOTANI, S.; BITTENCOURT, I. I. **Dados abertos conectados**. São Paulo: Novatec, 2015. 175 p.

LIMA, C. R. M.; SÁNCHEZ-TARRAGÓ, N.; MORAES, D.; GRINGS, L.; MAIA, M. R. Emergência de saúde pública global por pandemia de covid-19. **Revista Folha de Rosto**, [S. l.], v. 6, n. 2, p. 5-21, 2020. DOI: 10.46902/2020n2p5-21. Acesso em: 30 ago. 2022.

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.

MALAMUD, Carl; O'RILLEY, *et al.* **8 Principles of Open Government Data**, December 2007. Disponível em: https://public.resource.org/8_principles.html. Acesso em: 06 ago. 2022.

MICHEL, M. H. **Metodologia e Pesquisa científica em ciências sociais**: um guia prático para acompanhamento da disciplina e elaboração de trabalhos monográficos. São Paulo: Atlas, 2009.

OPEN GOVERNMENT WORKING GROUP. **Eight principles of open government data**. [S. l.]: Open Government Working Group, 2007. Disponível em: <https://opengovdata.org/>. Acesso em: 06 ago. 2022.

OPEN KNOWLEDGE INTERNATIONAL. **What is Open Data?** Disponível em: <https://okfn.org/opendata/>. Acesso em: 04 ago. 2022.

PARYCEK, P.; SACHS, M. Open government–information flow in Web 2.0. **European Journal of ePractice**, [S. l.], v. 9, n. 1, p. 1-70, 2010.

PIRES, M. T. **Guia de dados abertos**. São Paulo: CONIP SPUK, 2015. Disponível em: <http://book-hom.nicdev.com.br/dados-abertos/introducao>. Acesso em: 22 out. 2022.

SHANKARANARAYANAN, G.; EVEN, A. The metadata enigma. **Communications of the ACM**, ACM, v. 49, n. 2, p. 88–94, 2006. Disponível em: <https://dl.acm.org/doi/10.1145/1113034.1113035>. Acesso em: 16 out. 2022.

SILVA, N. **Maturidade em Dados Abertos**: entenda as 5 estrelas. [S. l.]: Open Knowledge Brasil, 2013. Disponível em: <https://ok.org.br/noticia/maturidade-em-dados-abertos-entenda-as-5-estrelas/>. Acesso em: 10 ago. 2022.

SILVA, P. N.; PINHEIRO, M. M. K. Métrica alternativa para dados governamentais abertos na América Latina. **Transinformação**, [s. l.], v. 31, n. e190009, 2019. Disponível em: <http://dx.doi.org/10.1590/2318-0889201931e190009>. Acesso em: 10 ago. 2022.

SILVA, M. B. A apresentação de dados abertos sobre a covid-19 pelo governo brasileiro. **Revista Fontes Documentais**, [s. l.], v. 3, p. 494-503, 2020. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/151191>. Acesso em: 30 ago. 2022.

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: 04 ago. 2022.