

**UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO
CENTRO DE CIÊNCIAS JURÍDICAS E ECONÔMICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO**

DIRCEU FLAVIO MACEDO

**DADOS ABERTOS GOVERNAMENTAIS:
MODELO DE GOVERNANÇA VOLTADO A QUALIDADE DE DADOS PARA
PUBLICAÇÃO EM REDE**

**VITÓRIA (ES)
2021**

DIRCEU FLAVIO MACEDO

**DADOS ABERTOS GOVERNAMENTAIS:
MODELO DE GOVERNANÇA VOLTADO A QUALIDADE DE DADOS PARA
PUBLICAÇÃO EM REDE**

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Informação da Universidade Federal do Espírito Santo como requisito para a obtenção do título de Mestre em Ciência da Informação.

Orientadora: Prof^ª. Dra. Daniela Lucas da Silva Lemos.

Linha de pesquisa: Memória, Representação e Informação.

**VITÓRIA (ES)
2021**

Ficha catalográfica disponibilizada pelo Sistema Integrado de Bibliotecas - SIBI/UFES e elaborada pelo autor

M141d Macedo, Dirceu Flavio, 1981-
Dados abertos governamentais : modelo de governança voltado a qualidade de dados para publicação em rede / Dirceu Flavio Macedo. - 2021.
153 f. : il.

Orientadora: Daniela Lucas da Silva Lemos.
Dissertação (Mestrado em Ciência da Informação) -
Universidade Federal do Espírito Santo, Centro de Ciências Jurídicas e Econômicas.

1. Análise de conteúdo (Comunicação). 2. Governança pública. 3. Metadados. 4. Organização da informação. 5. Política de informação. 6. Representação do conhecimento. I. Lemos, Daniela Lucas da Silva. II. Universidade Federal do Espírito Santo. Centro de Ciências Jurídicas e Econômicas. III. Título.

CDU: 001

DIRCEU FLAVIO MACEDO

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Informação da Universidade Federal do Espírito Santo como requisito para obtenção do título de mestre em Ciência da Informação.

Linha de pesquisa 2: memória, representação e uso da informação.

Aprovada em 08 de outubro de 2021.

COMISSÃO EXAMINADORA



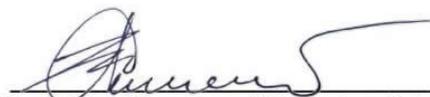
Prof^ª. Dr^ª. Daniela Lucas da Silva Lemos
Universidade Federal do Espírito Santo
Orientadora



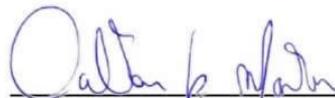
Prof^ª. Dr^ª. Lucileide Andrade de Lima do Nascimento
Universidade Federal do Espírito Santo
Examinadora Interna



Prof. Dr. Fabrício Martins Mendonça
Universidade Federal do Espírito Santo
Examinador Interno



Prof. Dr. Henrique Monteiro Cristovão
Universidade Federal do Espírito Santo
Examinador Interno



Prof. Dr. Dalton Lopes Martins
Universidade de Brasília
Examinador Externo

AGRADECIMENTOS

No local onde cresci nunca esperaram um dia que eu pudesse ser aprovado em um vestibular para uma universidade federal: “Está perdendo seu tempo, não vai conseguir, ninguém nunca conseguiu passar lá”, era o que comumente ouvia enquanto regressava ao meu lar após mais um dia de estudos.

Bom, anos se passaram e cá estou vivenciando a oportunidade de defender uma dissertação de mestrado para a banca de um prestigioso programa de pós-graduação na Ufes, a mesma instituição federal que me acolheu para a graduação há mais de 15 anos atrás, a qual sou muito grato por novamente me abrir as portas.

Abro esse texto oferecendo esse pequeno relato para agradecer em primeiro lugar aquele que nunca, jamais, me deixou esmorecer. Que me proporcionou forças de onde sequer imaginava ter. Que nunca me permitiu sentir subjugado ou inferiorizado. Que nunca me permitiu desistir, e que me impulsionou a ir em frente contra todos os prognósticos: ***Obrigado meu Deus por tudo que tem feito por mim ao longo de minha vida, e por todas as minhas conquistas!***

Agradeço a minha esposa por estar sempre ao meu lado me apoiando, incentivando e ajudando nessa empreitada tão desafiadora, muito obrigado por tudo. Agradeço ao meu pequeno Pedro, que com seu jeito especial e sorriso trouxe luz e alegria a nossa família. Meu filho, quando crescer espero que leia essas palavras e siga os ensinamentos de seu pai: corra atrás dos seus sonhos e jamais desista!

Estendo esse agradecimento a toda minha família: meus pais, minhas irmãs, meus sobrinhos queridos, que acompanharam essa jornada sempre me apoiando e me incentivando. Amo vocês do fundo do meu coração.

Também agradeço a Universidade Federal do Espírito Santo, ao Programa de Pós-Graduação em Ciência da Informação, o corpo docente e os meus colegas de turma por toda ajuda. Obrigado por me acolherem tão bem, vocês foram fantásticos!

Finalizo a seção fazendo um agradecimento especial a minha professora orientadora Daniela, que sempre me incentivou, compartilhou seu conhecimento e não me deixou abaixar a cabeça frente às adversidades. Tenho orgulho de não apenas ser seu aluno, mas poder considerá-la daqui em diante como uma amiga. Obrigado por tudo que fez por mim ao longo dessa jornada, serei eternamente grato.

Agradeço todas as dificuldades que enfrentei; não fosse por elas, eu não teria saído do lugar. As facilidades nos impedem de caminhar. Mesmo as críticas muito nos auxiliam.

Chico Xavier

RESUMO

A escalada na abertura de dados governamentais é um fenômeno originado de longos debates sobre a modernização do Estado contando, na atualidade, com diversos governos mundiais publicando um vasto conteúdo informacional em rede. Baseado nessa constatação, a pesquisa é motivada em compreender os desafios impostos à administração pública na publicação de dados de forma alinhada ao movimento aberto sob a ótica dos portais de dados. O objetivo desta pesquisa foi propor um modelo de governança voltado à qualidade de dados publicados em ambiente web, com foco na disponibilização dos recursos informacionais do governo com abrangência, consistência, eficácia e transparência. A metodologia foi o primeiro resultado da pesquisa, sendo apresentado como um instrumento científico dividido em três etapas. Na primeira etapa foi realizada uma pesquisa bibliográfica e documental evidenciando os antecedentes históricos do movimento de abertura de dados, conjuntamente a um panorama atual de iniciativas e estágios de evolução; o que levou a condução de um pré-teste nos portais de dados abertos dos EUA, Reino Unido, Brasil, e Irlanda, alcançando-se como resultado a definição dos desafios para abertura de dados governamentais com qualidade. Posteriormente, estes desafios serviram como insumo para a definição das dimensões analíticas fundamentais à proposição do modelo de governança. Na segunda etapa foi exposto o diagnóstico de maturidade para os portais de dados abertos, apresentando-se também a definição dos níveis de maturidade que foram empregados na pesquisa. O diagnóstico foi validado nos mesmos portais do pré-teste, com as notas finais sendo apresentadas e discutidas. Os resultados da validação indicaram que as iniciativas de abertura carecem de políticas de dados mais amplas e efetivas, faltando comprometimento político para colocar em prática os planos de ação formalizados, com pouca transparência na execução das ações e um insuficiente envolvimento social na sua formulação. Na terceira fase, os resultados foram tratados, gerando-se as inferências necessárias para a elaboração do modelo de governança voltado à qualidade de dados para publicação em rede, estando este fundamentado a partir dos resultados obtidos nas etapas anteriores. O modelo permitirá que os governos construam políticas de dados abrangentes e consistentes, mitigando as barreiras de publicação, e de forma consonante capacitando e encorajando os técnicos governamentais a publicarem dados com mais qualidade, beneficiando o movimento aberto e a sociedade.

Palavras-chave: Dados Abertos Governamentais; Organização da Informação; Modelo de Governança; Qualidade de Dados.

ABSTRACT

The escalation in the opening of government data is a phenomenon originated from long debates about the modernization of the State, counting, nowadays, with several world governments publishing a vast informational content on the net. Based on this finding, the research is motivated to understand the challenges imposed on the public administration in publishing data in line with the open movement from the perspective of data portals. The objective of this research was to propose a governance model aimed at the quality of data published in a web environment, with a focus on making government informational resources available with scope, consistency, efficiency, and transparency. The methodology was the first result of the research, being presented as a scientific instrument divided into three stages. In the first stage, a bibliographical and documentary research was carried out, showing the historical background of the data opening movement, together with a current panorama of initiatives and stages of evolution, which led to the conduction of a pre-test in open data portals in the US, UK, Brazil, and Ireland, achieving as a result the definition of challenges for the opening of quality government data. Subsequently, these challenges served as input for defining the fundamental analytical dimensions to the proposition of the governance model. In the second stage, the maturity diagnosis for open data portals was exposed, also presenting the definition of the maturity levels that were used in the research. The diagnosis was validated on the same portals as the pre-test, with the final grades being presented and discussed. The validation results indicated that openness initiatives lack broader and more effective data policies, lack political commitment to put into practice formalized action plans, with little transparency in the execution of actions and insufficient social involvement in their formulation. In the third phase, the results were treated, generating the necessary inferences for the elaboration of the governance model aimed at data quality for network publication, which is based on the results obtained in the previous steps. The model will allow governments to build comprehensive and consistent data policies, mitigating publication barriers, and accordingly empowering and encouraging government technicians to publish higher quality data, benefiting the open movement and society.

Keywords: Open Government Data; Information Organization; Governance Model; Data Quality.

LISTA DE FIGURAS

Figura 1 - Cronologia com os principais marcos de abertura de dados	27
Figura 2 - Evolução continuada do compartilhamento de dados na Web	48
Figura 3 - Modelo de governança com as dimensões a serem consideradas.....	52
Figura 4 - Processo metodológico delineado na pesquisa	61
Figura 5 - Indicadores do modelo de governança de dados	65
Figura 6 - Elementos relacionados a uma publicação de dados com qualidade.....	111
Figura 7 - Modelo de governança proposto.....	119
Figura 8 - Tela inicial do portal dos EUA.....	143
Figura 9 - Tela inicial do portal do Reino Unido.....	146
Figura 10 - Tela inicial do portal brasileiro.....	148
Figura 11 - Tela inicial do portal da Irlanda	150

LISTA DE QUADROS

Quadro 1 - Performance de algumas nações segundo o <i>Open Data Barometer</i>	23
Quadro 2 - Queda acentuada do Reino Unido no índice <i>OURdata Index</i>	23
Quadro 3 - Desafios para a abertura de dados governamentais	28
Quadro 4 - Quadro sinóptico sobre projetos de dados governamentais nas esferas públicas brasileiras	31
Quadro 5 - Categorias de metadados	38
Quadro 6 - Tipologia para padrões de dados	39
Quadro 7 - Exemplos de vocabulários de referência em ambientes digitais	42
Quadro 8 - Modelos conceituais que são baseados ou que utilizam ontologias	46
Quadro 9 - Descrição das etapas de levantamento bibliográfico	59
Quadro 10 - Descrição das etapas de levantamento documental	60
Quadro 11 - Técnicas e entregas de cada etapa da análise de conteúdo	62
Quadro 12 - Fundamentação diagnóstica para cada variável do indicador	68
Quadro 13 - Fundamentação dos níveis de maturidade	73
Quadro 14 - Pontuação dos indicadores e valor total do diagnóstico	74
Quadro 15 - Modelo de diagnóstico de maturidade para portais de dados abertos governamentais	75
Quadro 16 - Resultados do indicador de Infraestrutura em TI	85
Quadro 17 - Resultados do indicador de Acessibilidade e Usabilidade	88
Quadro 18 - Resultados do indicador de Curadoria e Publicação dos Dados	90
Quadro 19 - Resultados do indicador Político e Social	93
Quadro 20 - Resultados do indicador Econômico	95
Quadro 21 - Resultados do indicador de Organização e Processos Internos	97
Quadro 22 - Resultados do indicador de Aspectos Legais e de Privacidade	99
Quadro 23 - Resultados do indicador de Capacitação e Treinamento	101
Quadro 24 - <i>Ranking</i> dos portais avaliados a partir do diagnóstico	103
Quadro 25 - Conjunto de atividades relacionadas ao modelo de governança proposto	113

LISTA DE GRÁFICOS

Gráfico 1 - Comparativo do indicador de Infraestrutura em TI.....	86
Gráfico 2 - Comparativo do indicador de Acessibilidade e Usabilidade	88
Gráfico 3 - Comparativo do indicador de Curadoria e Publicação dos Dados.....	91
Gráfico 4 - Comparativo do indicador Político e Social	93
Gráfico 5 - Comparativo do indicador Econômico	96
Gráfico 6 - Comparativo do indicador de Organização e Processos Internos	98
Gráfico 7 - Comparativo do indicador de Aspectos Legais e de Privacidade	100
Gráfico 8 - Comparativo do indicador de Capacitação e Treinamento	102
Gráfico 9 - Desempenho dos portais frente aos indicadores.....	104

LISTA DE ABREVIATURAS E SIGLAS

AACR2	<i>Anglo-American Cataloguing Rule</i>
API	<i>Application Programming Interface</i>
CC	Ciência da Computação
CI	Ciência da Informação
CCG	Conselho de Cooperação do Golfo
CGU	Controladoria Geral da União
INDA	Infraestrutura Nacional de Dados Abertos
ISBD	<i>International Standard Bibliographic Description</i>
KO	<i>Knowledge Organization</i>
LAIP	Lei de Acesso à Informação Pública
LOD	<i>Linked Open Data</i>
MARC	<i>Machine-Readable Cataloging</i>
NISO	<i>National Information Standards Organization</i>
OGP	<i>Open Government Partnership</i>
ONG	Organização não Governamental
OGDA	Órgão de Governança de Dados Abertos
PDA	Plano de Dados Abertos
RDF	<i>Resource Description Framework</i>
SRI	Sistemas de Recuperação da Informação
TIC	Tecnologia da Informação e Comunicação

SUMÁRIO

1	INTRODUÇÃO	13
1.1	PROBLEMA DE PESQUISA E HIPÓTESE.....	15
1.2	OBJETIVOS	16
1.3	JUSTIFICATIVAS E CONTRIBUIÇÕES.....	17
2	ESTADO DA ARTE: DADOS ABERTOS GOVERNAMENTAIS	20
2.1	ANTECEDENTES E ASPECTOS HISTÓRICOS	20
2.2	PANORAMA ATUAL: DADOS ABERTOS GOVERNAMENTAIS EM ÂMBITO MUNDIAL.....	22
3	FUNDAMENTAÇÃO TEÓRICA	34
3.1	ORGANIZAÇÃO E REPRESENTAÇÃO DA INFORMAÇÃO E DO CONHECIMENTO EM CONTEXTOS DIGITAIS	34
3.1.1	Metadados na representação descritiva de documentos	37
3.1.2	Vocabulários controlados na normalização de conteúdo documental	40
3.1.3	Modelos conceituais para representação documental	43
3.1.4	Dados conectados (<i>Linked Data</i>): construindo a Web de Dados	47
3.2	MODELO E GOVERNANÇA DE DADOS	50
3.3	CONTRIBUIÇÃO DA FUNDAMENTAÇÃO TEÓRICA PARA A PESQUISA ...	55
4	METODOLOGIA DE PESQUISA	57
4.1	CLASSIFICAÇÃO DA PESQUISA	57
4.2	ETAPAS METODOLÓGICAS PARA A PROPOSIÇÃO DOS MODELOS DE DIAGNÓSTICO E DE GOVERNANÇA	58
4.2.1	Levantamento bibliográfico e documental	58
4.2.2	Coleta e análise dos dados	60
4.2.2.1	<i>Etapa de pré-análise: definição das categorias de análise</i>	63
4.2.2.2	<i>Etapa de exploração do material: elaboração do modelo de diagnóstico de maturidade</i>	67
4.2.2.3	<i>Etapa de tratamento dos resultados: proposta do modelo de governança</i>	83
5	RESULTADOS	84
5.1	APLICAÇÃO DO DIAGNÓSTICO DE MATURIDADE	84
5.1.1	Indicador de Infraestrutura em TI	84
5.1.2	Indicador de Acessibilidade e Usabilidade	86
5.1.3	Indicador de Curadoria e Publicação dos dados	89

5.1.4	Indicador Político e Social	91
5.1.5	Indicador Econômico	94
5.1.6	Indicador Organização e Processos Internos	96
5.1.7	Indicador de Aspectos Legais e de Privacidade	98
5.1.8	Indicador de Capacitação e Treinamento	100
5.1.9	Resultado consolidado da avaliação nos portais governamentais	103
5.2	APRESENTAÇÃO DO MODELO DE GOVERNANÇA	111
6	DISCUSSÃO DOS RESULTADOS	122
7	CONCLUSÕES E TRABALHOS FUTUROS	124
	REFERÊNCIAS	126
	APÊNDICES	138
	APÊNDICE A – REFERÊNCIAS DA AMOSTRA DE DOCUMENTOS ESCOLHIDOS PARA ANÁLISE ORGANIZADOS EM ORDEM CRONOLÓGICA DE PUBLICAÇÃO	Erro! Indicador não definido.
	APÊNDICE B – PRÉ-TESTE REALIZADO NO PORTAL DE DADOS ABERTOS DOS ESTADOS UNIDOS.....	138
	APÊNDICE C – PRÉ-TESTE REALIZADO NO PORTAL DE DADOS ABERTOS DO REINO UNIDO	146
	APÊNDICE D – PRÉ-TESTE REALIZADO NO PORTAL DE DADOS ABERTOS DO BRASIL	148
	APÊNDICE E – PRÉ-TESTE REALIZADO NO PORTAL DE DADOS ABERTOS DA IRLANDA	150

1 INTRODUÇÃO

Nos últimos anos, diversas esferas governamentais comprometidas com a modernização do Estado e com os anseios da sociedade passaram a promover a abertura de seus dados apoiando-se na moderna infraestrutura da rede mundial de computadores (DAVIES, 2010; GRAY, 2014).

O processo embrionário desse movimento de abertura teve seu epicentro no final do século XX e início do século XXI, quando indivíduos e organizações passaram a cobrar uma postura mais aberta dos governos, exigindo maior acesso às informações, vislumbrando progressos na defesa de causas sociais e cívicas na produção de novas tecnologias, no fomento de inovação, no avanço científico, nos interesses econômicos, nos estudos sobre interoperabilidade, e na criação de novos formatos para acesso a dados.

O advento das Tecnologias de Informação e Comunicação (TICs) e a popularização da Internet permitiram o surgimento de um ambiente propício para questionamento e reflexão do papel do Estado perante a sociedade. Nesse contexto, uma nova abordagem foi requerida aos governos no esforço de estreitar sua relação com a sociedade, tornando-se inevitável seu aperfeiçoamento para um modelo mais eficiente. Na sucessão de processos que culminaram em um governo cada vez mais eletrônico, aflorou-se o movimento para a abertura de dados governamentais (DAVIES, 2010; DAVIES; BAWA, 2012; GRAY, 2014).

A abertura de dados pode ser encarada como um grande movimento para compartilhamento de conhecimento e ampliação democrática, beneficiando tanto a sociedade quanto o governo no aspecto de acessibilidade a fontes de dados (BAUER; KALTENBÖCK, 2011), o que aumenta a responsabilidade, a transparência, a eficiência governamental, a inovação, o combate a corrupção, o empoderamento cívico, além de incentivar o uso e a reutilização de informações para ações de interesse (OPEN GOVERNMENT PARTNERSHIP, 2011; ATTARD *et al.*, 2015).

Projetos de abertura de dados vêm sendo objeto de análise por diversas partes interessadas, incluindo organizações, ativistas, governos e pesquisadores na busca de se compreender os preceitos estabelecidos pelo movimento aberto; se estão de fato trazendo benefícios e transformações sociais pleiteadas; ou se estão atendendo às melhores práticas de publicação de dados.

Pesquisas têm sido desenvolvidas progressivamente nos campos da Ciência da Informação (CI) e da Ciência da Computação (CC), orientadas à organização da informação em ambientes digitais com o objetivo de aumentar o escopo da interoperabilidade e da integração das bases disponíveis na Web (LEMOS; SOUZA, 2020a).

Em adição, encontra-se o uso do padrão de dados abertos ligados ou *Linked Open Data* (LOD), que prioriza a ligação entre bases de dados de fontes heterogêneas (BIZER; HEATH; BERNERS-LEE, 2009; MACHADO; SOUZA; DA GRAÇA SIMÕES, 2019). Lemos e Souza (2019) afirmam que em apoio à organização de documentos na Web recomenda-se também o uso de ontologias e vocabulários controlados para a representação de relacionamentos semânticos e conceituais, o emprego de modelos conceituais (de referência e ontológicos) para orientar a modelagem e o processo de busca e recuperação da informação na realidade digital.

No entanto, além dos desafios técnicos associados à implantação, princípios e formatos, os governos precisam lidar com outras barreiras atreladas à abertura de seus dados (KUČERA *et al.*, 2015; SANTAREM SEGUNDO, 2015; PINTO; ALMEIDA, 2020), especialmente em relação às estratégias de governança digital, exigindo-se um grande esforço interdisciplinar para a disponibilização dos dados governamentais na rede sem impor riscos associados à integridade e à confiabilidade dos dados (SHEPHERD *et al.*, 2019).

Apesar dos avanços do movimento de abertura de dados governamentais, ainda é identificável uma falta de comprometimento político das estruturas administrativas com a qualidade dos dados disponibilizados seguindo definições trabalhadas ao longo dos anos pelos ativistas de dados.

Destacam-se problemas na conscientização dos gestores públicos na tarefa técnica de publicação de dados, bem como no esclarecimento de quais princípios e disciplinas são necessários à abertura; e problemas relativos à capacitação dos técnicos públicos responsáveis pela publicização dos dados. Todas essas questões têm gerado uma queda de compromisso na qualidade da abertura dos dados, afetando negativamente as iniciativas governamentais que se propõem a abrir os dados à sociedade em rede.

Não obstante, a formalização de políticas públicas de dados por parte de algumas nações não melhorou tal cenário, mostrando-se insuficiente (LUNA-REYES; NAJAFABADI, 2019; SHEPHERD *et al.*, 2019; WANG; SHEPHERD; BUTTON, 2019).

Deve-se sanar essas lacunas e aplicá-la de forma responsável e criteriosa, com suporte de uma estrutura pública de governança comprometida com altos padrões de qualidade. Os procedimentos devem ser conduzidos de forma a garantir maior transparência, eficiência e qualidade aos dados. Isso se aplica a toda a administração pública, perpassando gestores até alcançar-se servidores públicos.

1.1 PROBLEMA DE PESQUISA E HIPÓTESE

Dados disponibilizados por esferas governamentais são responsáveis por representar grande parte da realidade de uma nação (economia, política, geografia, sociedade etc.), exigindo atenção em seu tratamento e publicação. Como afirma Davies (2010, p. 2): “Os governos coletam grandes quantidades de dados e mantêm *datasets* nacionais significativos, dessa forma dados significativos devem ser representados ou contextualizados de alguma forma - convertidos em informação”.

Entretanto, por mais que os governos lidem com inúmeras barreiras para abrir seus dados na rede (KUČERA *et al.*, 2015; SANTAREM SEGUNDO, 2015; W3C, 2017; RAUTENBERG *et al.*, 2018; PINTO; ALMEIDA, 2020), algumas deficiências da administração pública estão contribuindo para gerar alguns empecilhos à abertura, tais como: i) falta de comprometimento político dos governos com a qualidade dos dados; ii) políticas de dados não suficientemente abrangentes para garantir a qualidade na abertura dos dados; e iii) cultura de abertura de dados incipiente no próprio meio da administração pública.

Portanto, a situação problemática que se configura nessa pesquisa concentra-se no fato de que o movimento de abertura de dados governamentais possui problemas de governança de dados (SHEPHERD *et al.*, 2019), carecendo de maior atenção nas áreas de capacitação e treinamento bem como por parte dos legisladores públicos, gestores e servidores alocados administrativamente nas áreas de publicação de dados (WORLD WIDE WEB FOUNDATION, 2018).

A presente pesquisa, baseada em boas práticas de abertura de dados governamentais (BLANK, 2019; WORLD WIDE WEB FOUNDATION, 2018; OECD, 2020), reafirma a importância da orientação para gestores que tratam da formulação de políticas de dados abertos, além de capacitação para os indivíduos que lidam com a publicação de dados em quaisquer esferas governamentais.

Acredita-se que o esclarecimento sobre o uso combinado de princípios oriundos dos campos da CI (compreender o funcionamento das formas de representação da informação disponíveis), da CC (entender como os editores estão implementando os aspectos de interoperabilidade de dados entre os sistemas de informação) e da gestão em administração pública (incorporar conceitos e processos de governança), podem auxiliar no desenvolvimento de um modelo de governança que guie as autoridades envolvidas no gerenciamento de dados a uma publicação de dados em rede com maior qualidade.

De acordo com a situação problemática apresentada, a presente pesquisa busca responder a seguinte questão: *Como instituições de governo, em âmbito mundial, estão conduzindo suas práticas de publicações de dados abertos, e como podem melhorar o gerenciamento de dados abertos disponibilizados para publicação, visando qualidade em todo o ciclo de vida desses dados?* Parte-se da hipótese nessa pesquisa de que grande parte das nações estão abrindo dados governamentais pouco qualificados em rede, prejudicando o acesso e reutilização desse importante conteúdo informacional por parte dos membros integrantes da sociedade.

1.2 OBJETIVOS

O objetivo geral desta pesquisa é propor um modelo de governança voltado a qualidade de dados destinados a publicação em ambiente web, com ênfase na disponibilização qualificada e transparente de recursos informacionais no domínio do governo digital.

Pretende-se que o modelo proposto seja convertido em um instrumento orientado a gestão de dados com ênfase na governança. O modelo propõe-se a evidenciar as dimensões que precisam ser observadas na elaboração de políticas de dados incluindo os elementos a serem atendidos e as ações a serem executadas para garantir maior integridade e confiabilidade aos dados abertos publicados em âmbito mundial, resguardando o valor dos dados para que beneficiem toda a sociedade.

Intenciona-se também que o modelo de governança possa servir como documentação profícua orientada aos gestores públicos de dados, bem como a editores, professores, estudantes e entusiastas na promoção de melhorias ao campo de abertura de dados governamentais, capacitando os indivíduos que lidam no cotidiano com a publicação de dados governamentais em rede.

Como objetivos específicos almeja-se:

- identificar iniciativas relevantes de abertura de dados governamentais na literatura de abrangência nacional e internacional para fins de análise de evidências;
- selecionar portais de iniciativas governamentais que disponibilizam dados abertos para a sociedade;
- analisar e comparar os portais governamentais selecionados a partir de dimensões analíticas previamente definidas, e posteriormente discutir seus resultados.

1.3 JUSTIFICATIVAS E CONTRIBUIÇÕES

Os governos detêm bases imensas de dados, incluindo indicadores econômicos, registros de empresas e eleitorais, relatórios escolares, estatísticas de crimes, entre outras (BIZER; HEATH, 2011). Sant'ana (2009) defende que os indivíduos tenham acesso ao maior volume possível de dados sobre as ações executadas pelos agentes públicos, consolidando dessa forma uma participação mais cívica.

De forma consoante, Davies (2010) observou a partir de investigação que, a abertura de dados contribuiu para desafiar desequilíbrios de poder em vários processos burocráticos entre indivíduos e o estado (por exemplo, solicitação de financiamentos ou apelação sobre uma decisão de admissão escolar).

A integração desses dados levaria a percepção de novos cenários antes não evidentes, por exemplo, no Reino Unido, em que o serviço civil passou a fornecer informações sobre vagas em aberto de forma estruturada, facilitando aos portais de emprego externos a sua incorporação, aumentando com isso a transparência no mercado de trabalho de forma significativa (BIZER; HEATH, 2011).

Dessa forma, o movimento de abertura de dados governamentais possibilita que indivíduos e organizações tenham acesso a indicadores tributários, educacionais, e de saúde (EUROPEAN UNION, 2021), permitindo um melhor posicionamento frente a realidades do cotidiano. Para que se garanta tais benefícios é necessário que o governo reúna esforços, conjuntamente com a participação cívica, na elaboração de políticas de dados e governança mais abrangentes e realmente eficazes na disponibilização de dados em rede com qualidade.

A presente pesquisa justifica-se pela demanda social já reconhecida de abertura de dados por parte dos governos, de forma mais ordenada e responsável possível, especialmente pelo fato desse conteúdo informacional representar uma parcela significativa da realidade de toda uma nação e de seus cidadãos, sendo de extrema importância para a ampliação democrática na sociedade.

Nesse contexto, torna-se notória a relevância desta pesquisa, que visa contribuir para que os agentes que estão à frente de políticas públicas, ou que lidam com a publicação de dados, obtenham informações sobre as barreiras a serem vencidas para a implementação de iniciativas de abertura exitosas. Por consequência, fazendo-se valer de um modelo de governança que os conduza na aplicação de um conjunto de ações fundamentadas para mitigar os desafios de publicação de dados com mais qualidade em rede.

As contribuições dessa pesquisa são enumeradas a seguir:

- sistematizar um conjunto de desafios e oportunidades mapeados nas ações de abertura de dados governamentais em alguns países a fim de estreitar a relação entre a esfera pública e a sociedade;
- fomentar a representação semântica de dados governamentais, beneficiando seu acesso e consumo por diversos atores, dentre eles os sociais que buscam o desenvolvimento de aplicativos cívicos;
- conduzir a administração pública na elaboração de políticas de dados mais eficientes, contando com maior abrangência e transparência;
- capacitar os desenvolvedores e técnicos governamentais a publicarem os dados na rede com qualidade a partir do uso eficaz de um conjunto de ações de governança de dados;
- melhorar a eficiência nos organismos públicos que não dispõem de recursos ou setores específicos para publicação de dados, e encorajar servidores que atuam nessa atividade a iniciarem uma cultura de publicação de dados com qualidade dentro de seus órgãos.

Espera-se como produtos de pesquisa a concepção de um modelo de governança de dados, que facilite a compreensão e a didática pelas partes interessadas em proceder a abertura de dados governamentais. O produto resultante e toda a sua documentação (arquivos, dados etc.) serão disponibilizados no

repositório conhecido como *GitHub*¹, com o propósito de compartilhar e disseminar os resultados de pesquisa para a comunidade de prática.

Esta dissertação está organizada da seguinte maneira:

- o Capítulo 1 apresenta o contexto da pesquisa, suas questões e hipótese, os objetivos (geral e específicos), justificativas e contribuições;
- o Capítulo 2 discorre sobre o estado da arte envolvendo dados abertos governamentais situando o problema da pesquisa em contextos atuais;
- o Capítulo 3 disserta sobre a fundamentação teórica subjacente, tratando de conceitos e teorias importantes ao desenvolvimento metodológico e aplicado da pesquisa;
- o Capítulo 4 trata da classificação e dos procedimentos metodológicos adotados na condução da pesquisa;
- o Capítulo 5 apresenta os resultados associados à aplicação do modelo de diagnóstico de maturidade nos portais, e, como produto da dissertação, a proposição do modelo de governança voltado a qualidade de dados para publicação em rede;
- o Capítulo 6 compreende as discussões relacionadas a todos os resultados alcançados a partir da condução da pesquisa; e
- finalmente, o Capítulo 7 conclui a dissertação tecendo as considerações finais, e expondo as limitações, contribuições, e indicações de trabalhos futuros.

¹ *GitHub* é uma plataforma de hospedagem de código-fonte e arquivos com controle de versão.

2 ESTADO DA ARTE: DADOS ABERTOS GOVERNAMENTAIS

Nesta seção dirigida à revisão de literatura discutimos a temática dados abertos governamentais desdobradas em duas subseções tratando dos antecedentes, dos aspectos históricos e do panorama atual, em âmbito mundial.

2.1 ANTECEDENTES E ASPECTOS HISTÓRICOS

O tema “dados abertos” surgiu pela primeira vez em 1995 a partir de um documento pertencente a uma agência científica dos Estados Unidos (EUA) que promovia troca aberta de dados geofísicos e ambientais entre agências governamentais para melhor compreensão de fenômenos associados. A abertura de dados pode ser considerada uma práxis da comunidade científica que já era defendida pelo pesquisador Robert King Merton², que destacava a ampliação do conhecimento humano como um grande benefício à sociedade.

Desse modo, os pesquisadores foram os primeiros a perceber os benefícios da abertura e compartilhamento de dados, contudo foi o encontro entre teorias científicas e conceitos de *software* livre e código aberto que moldaram o panorama de dados abertos que conhecemos hoje (CHIGNARD, 2013).

Gray (2014) apresenta algumas evidências que contribuíram para o paradigma de abertura de dados nos anos de 1990 e início dos anos 2000, sobretudo na Europa e nos EUA, a saber: i) defesa de políticas e debates promovidos por ativistas, tecnólogos e *hackers* cívicos³ a respeito da infraestrutura governamental e do fornecimento de informações públicas; ii) discussões sobre políticas de informação do setor público; iii) debates sobre o potencial econômico e social das informações e dados geoespaciais, por meio dos quais viabilizaram discussões acerca de padrões, licenças e formatos.

Em paralelo, a evolução das TICs colocou a tecnologia digital e a Internet como protagonistas da digitalização de processos, impactando os governos no que diz respeito à democratização de acesso e à transparência em dados governamentais na rede (JETZEK; AVITAL; BJØRN-ANDERSEN, 2014).

² Robert King Merton é considerado um dos pais da sociologia da ciência.

³ Indivíduos que exercem um ativismo cívico empregando avançados conhecimentos computacionais.

Os governos se valeram da evolução das TICs, especialmente do potencial da Internet, na democratização do acesso à informação pela sociedade, cujo fenômeno é denominado governo eletrônico (e-Gov), e, atualmente, conceituado como governo digital. O e-Gov impulsionou a abertura de dados pelo governo, tornando-se essencial para a democratização e transparência das atividades governamentais (PINTO; ALMEIDA, 2020).

O movimento para abertura de dados governamentais deu um grande passo em 2007 quando em Sebastopol, Califórnia (EUA) ocorreu a reunião de “advogados do governo aberto” objetivando criar uma definição para dados abertos, e assim contribuir para orientar os governos em iniciativas de abertura (GRAY, 2014). A partir daí, o conceito de “dados abertos” migrou de uma investigação de ativistas e técnicos para o discurso de autoridades governamentais.

Em 2009, deu-se a publicação do “Memorando sobre Transparência e Governo Aberto” pelo então presidente dos EUA, Barack Obama (SCHROCK, 2016), colocando a cultura aberta no centro da ação pública, pleiteando mais transparência, participação e colaboração (CHIGNARD, 2013), além de estabelecer uma cooperação entre políticos, administração pública, indústria e indivíduos. Em 2011 foi lançado o *Open Government Partnership* (OGP), quando oito países fundadores (Brasil, Indonésia, México, Noruega, Filipinas, África do Sul, Reino Unido e EUA) endossaram a Declaração do Governo Aberto (BAUER; KALTENBÖCK, 2011).

O compromisso de abertura de dados foi reafirmado pelo G8⁴ em 2013 por meio da “Carta Aberta de Dados” (G8, 2013), estabelecendo que, os dados por padrão sejam abertos, respeitando-se as legislações de privacidade em vigor.

Em 2009, tem-se o primeiro marco no movimento de abertura no Brasil com o Decreto Federal nº 6.932, determinando ao Poder Executivo a disponibilização de suas informações e dados aos órgãos e entidades públicas, e provisionamento de soluções para viabilizar o compartilhamento (RIBEIRO; ALMEIDA, 2011).

No entanto, somente em 2011, com a entrada em vigor da Lei de Acesso à Informação Pública, é que se inicia a abertura de dados do governo (SILVA; PINHEIRO, 2015). Ainda em 2011, o Brasil se comprometeu com a OGP, lançando seu Portal de Dados Abertos (BRASIL, 2021b). Em 2016, o Tribunal de Contas da União (TCU) elencou cinco motivos para abertura de dados (BRASIL, 2015): i)

⁴ O G8 é uma sigla que denomina os oito países mais ricos e influentes do mundo: Estados Unidos, Japão, Alemanha, Canadá, França, Itália, Reino Unido e Rússia.

exigência de maior transparência da gestão pública pela sociedade; ii) contribuição da própria sociedade para criação de serviços inovadores; iii) qualidade dos dados; iv) novos negócios; e v) obrigatoriedade legal.

Também em 2016 foi instituída a Política de Dados Abertos do Poder Executivo Federal, criando a seguir a Infraestrutura Nacional de Dados Abertos (INDA) e o Plano de Dados Abertos (PDA) - um instrumento que operacionaliza a Política de Dados Abertos Governamentais (BRASIL, 2021a).

Logo, tem-se que o moderno significado de “abertura” para dados governamentais, bem como o contexto no qual vivenciamos atualmente é resultado de propostas engajadas de reformas, iniciativas e experiências sucessivas de abertura de dados em diversos campos de atuação ao redor do mundo.

Somam-se a isso o benefício do uso das TICs para a qualidade dos sistemas de informação do governo, o que ampliou o debate e a participação da sociedade na construção de políticas públicas (PINTO; ALMEIDA, 2020, p.5). Em decorrência desse legado, o panorama atual associado a dados abertos governamentais em contexto mundial, incluindo o Brasil, é elucidado na próxima seção.

2.2 PANORAMA ATUAL: DADOS ABERTOS GOVERNAMENTAIS EM ÂMBITO MUNDIAL

Essa seção busca descrever o cenário atual em países com papéis legados e preponderantes ao progresso de dados abertos governamentais em escala mundial, como EUA e Reino Unido, além de países que se destacam em seus respectivos continentes, como são os casos do Brasil na América Latina, da Coreia do Sul na Ásia, do Quênia na África, da Austrália na Oceania, e Áustria, Alemanha, República Tcheca, Irlanda, Espanha e França na Europa.

Uma tendência, de acordo com Kassen (2018), é integrar projetos locais em uma única plataforma, como o *European Data Portal* (EUROPEAN UNION, 2021), e o *Latin American Open Data Initiative* (INICIATIVA LATINOAMERICANA POR LOS DATOS ABIERTOS, 2021). Outra tendência são os painéis e relatórios para aferição da maturidade dos países na abertura, com destaque ao: *Open Data Maturity Report 2019* (BLANK, 2019); *OURdata Index 2019* (OECD, 2020); *Open Data Barometer* (WORLD WIDE WEB FOUNDATION, 2018); e o *Open Data Monitor* (OPEN DATA MONITOR, 2021).

O *Open Data Barometer*, por exemplo, mede as iniciativas de países que se comprometeram com a abertura de dados avaliando-se, com notas de 0 (zero) a 100, as políticas para dados abertos, o gerenciamento e a publicação, a execução das políticas, e os treinamentos (WORLD WIDE WEB FOUNDATION, 2018). O monitor segue os princípios da “Carta Internacional de Dados Abertos” (OPEN DATA CHARTER, 2015), que estabelece que os dados sejam: abertos por padrão, relevantes, publicados com celeridade, acessíveis, reutilizáveis e interoperáveis.

Os EUA e o Reino Unido, outrora líderes em atividades de dados abertos, perderam espaço nos últimos anos em participação. A edição de 2018 do *Open Data Barometer* externa preocupação, pois ambos os países assumiram compromissos de abertura e, no entanto, estão seguindo na contramão dos princípios da Carta Internacional de Dados Abertos (WORLD WIDE WEB FOUNDATION, 2018). Através do Quadro 1 observa-se a queda do Reino Unido e dos EUA, a ascensão do Canadá, Austrália e a Coreia do Sul, além da posição do Brasil.

Quadro 1 - Performance de algumas nações segundo o *Open Data Barometer*

Posição	País	Pontuação	Varição
1º	Canadá	76	+18
2º	Reino Unido	76	- 4
3º	Austrália	75	+17
4º	França	72	+17
5º	Coreia do Sul	72	+25
9º	Estados Unidos	64	-11
14º	Brasil	50	+15

Fonte: Adaptado do World Wide Web Foundation (2018).

No Quadro 2, o *OURdata Index* 2019 aponta que o Reino Unido (líder em 2017) não sustentou uma agenda de dados abertos na sua política pública, caindo para a 20ª posição. O EUA, à época, não estava disponível nessa pesquisa.

Quadro 2 - Queda acentuada do Reino Unido no índice *OURdata Index*

Posição	País	Pontuação	Varição
1º	Coreia do Sul	0.93	- 0.01
2º	França	0.90	+ 0.05
3º	Irlanda	0.77	+ 0.24
4º	Japão	0.75	- 0.05
5º	Canadá	0.73	+ 0.03
20º	Reino Unido	0.57	- 0.22

Fonte: Adaptado do OECD (2020).

Em que pese à posição intermediária do Brasil no *Open Data Barometer*, é louvável o seu crescimento nos últimos anos. Com relação ao índice *OURdata Index*, o país estava indisponível por não ser membro da OECD.

Os estudos acadêmicos corroboram os monitores. Wang e Shepherd (2020) ao analisarem 400 *datasets* do portal do Reino Unido constataram sérios problemas na oferta de dados. Apontaram que a maioria dos recursos informacionais disponíveis nos últimos 42 meses (cerca de 70%) é representada por dados fornecidos no nível um, segundo o *ranking* cinco estrelas (BIZER; HEATH; BERNERS-LEE, 2009), contra apenas 30% de dados estruturados e granulares.

Logo, conclui-se que mesmo os entusiastas do governo aberto enfrentam grandes obstáculos para publicação. Evidencia-se uma série de confusões organizacionais no Reino Unido, seja para localizar a responsabilidade da política de dados abertos do governo, seja sobre a definição de quais profissionais devem liderar a atividade, ou sobre as abordagens para publicação (SHEPHERD *et al.*, 2019). Sobre essa situação, Wang, Shepherd e Button (2019, p. 60, *tradução nossa*) afirmam que: “Se o Reino Unido, como pioneiro da iniciativa de dados abertos, está sobrecarregado com esses problemas, então os países menos avançados provavelmente experimentarão obstáculos iguais ou semelhantes”.

A realidade americana diagnosticada por Luna-Reyes e Najafabadi (2019, p.178, *tradução nossa*) atesta que, a falta de clareza da política de dados abertos governamentais “[...] levou a um período de estagnação no desenvolvimento do programa de dados abertos do governo”. Soma-se a esse fato a carência de envolvimento coordenado das diversas agências para compreender as necessidades de requisitos de dados de suas respectivas comunidades. Yi (2019) alerta que os EUA publicaram mais de 50% dos seus dados abertos em formatos que inviabilizam a leitura por máquinas, comprometendo a sua reutilização. Em contrapartida, o Canadá destaca-se com suas taxas de abertura e transparência, contando com engajamento do terceiro setor - em número superior aos dos EUA - e de Organização não Governamental (ONGs) no reuso de dados abertos governamentais (YOUNG; VERHULST, 2016).

Na África, o movimento de abertura de dados desacelerou, estando o Quênia na liderança em aspectos técnicos, com um portal de dados com recursos similares aos do Reino Unido e dos EUA (AFFUL-DADZIE E.; AFFUL-DADZIE A., 2017).

Já na Austrália, de acordo com Chatfield e Reddick (2018), o governo central é o agente inovador para a abertura de dados frente aos sete governos estaduais, realizando investimentos em infraestrutura tecnológica. Em complemento, as iniciativas estaduais pioneiras tiveram em comum a ação de gestores públicos na difusão da nova cultura de abertura entre os diferentes setores.

Partindo para o continente europeu, a Áustria se destaca na criação de um Modelo de Implementação de Governo Aberto que serve de base para projetos locais e para outras nações, sendo reaproveitado para organizar processos de publicação de dados (BLANK, 2019). Seu portal de dados abertos prima tanto pela qualidade, semântica, relevância dos dados (SÁEZ MARTÍN; ROSARIO; PÉREZ, 2016) quanto pelo seu aspecto técnico, com ótimos indicadores de segurança e eficiência na infraestrutura de TI (JUANA-ESPINOSA; LUJÁN-MORA, 2019).

Na Alemanha, a Política Nacional de Dados Abertos, em nível federal, vigora desde 2017 com a abordagem aberta por padrão para os dados publicados (BLANK, 2019; OECD, 2020). Kassen (2018), porém, informa que o governo central outorga autonomia aos demais governos locais no desenvolvimento de iniciativas próprias, o que resulta na profusão de políticas de dados abertos no país.

As iniciativas na Alemanha possuem barreiras, como a falta de padronização gráfica e dos formatos de dados entre os inúmeros portais governamentais existentes, afetando a usabilidade (VETRÒ *et al.*, 2016). Wirtz *et al.* (2016) também apontam barreiras organizacionais (legais, burocráticas, hierárquicas, sentimento de empregos ameaçados) que estão minando a atuação dos servidores públicos no esforço de cooperação para iniciativas de dados governamentais abertos no país.

Na República Tcheca, o Catálogo Nacional de Dados Abertos⁵, criado em 2015, concentra os *datasets* governamentais, sanando a grande descentralização das fontes de dados do país (PAVLÍK *et al.*, 2020). O modelo de governança para a publicação de dados, contando em seu início com forte centralização, abrange todas as esferas de governo, mas atualmente, o que se observa são os governos locais conduzindo suas próprias iniciativas de dados abertos (BLANK, 2019).

Recentemente a Irlanda, Espanha e França despontaram na Europa, possuindo desde 2018 uma liderança no continente, sendo criadores de tendências (BLANK, 2019). Ao contrário da Alemanha e República Tcheca, que restringem ao

⁵ <https://data.gov.cz/>

setor público a participação em seus modelos de governança, a Irlanda e a Espanha, de maneira inovadora, contemplam em seus modelos a participação de atores tanto do setor privado, quanto da sociedade (BLANK, 2019).

No aspecto interativo, os portais nacionais da França, Áustria e República Tcheca permitem que os usuários publiquem seus *datasets* no sítio, transformando-os em verdadeiras plataformas colaborativas (OECD, 2020). A Europa também vem inovando nas formas de aumentar o valor dos dados, implantando o recurso de *Application Programming Interface* (API)⁶ nos portais para melhorar o acesso aos dados, beneficiando dessa forma as aplicações feitas por terceiros (BLANK, 2019).

Com relação ao LOD, a Europa ainda não manifestou grande entusiasmo (BLANK, 2019). Segundo o Relatório de Maturidade de Dados Abertos do *European Union* (2021, p. 58, *tradução nossa*) “A porcentagem de dados abertos conectados é inferior a 10% em 27 países (96%). Somente na Letônia, entre 10% e 19% dos dados têm cinco estrelas”. No nível de quatro estrelas, a Alemanha se destaca na Europa, com cerca de 60% a 70% dos dados disponibilizados nessa especificação.

Apesar de grande parte da literatura sobre dados abertos governamentais indicar países ocidentais (ALTAYAR, 2018), é possível também identificar iniciativas emergentes em países do Oriente Médio e Ásia, conforme relatos a seguir.

No continente Asiático a Coreia do Sul tem sido um exemplo de alta capacidade, experiência e maturidade (ALTAYAR, 2018), contando em 2018 com cerca de 51% de *datasets* legíveis por máquina e não proprietários (YI, 2019). A China acenou com avanços a partir de 2016 com uma política de dados abertos governamentais. No entanto, o país ainda não dispõe de um portal central, contando com 19 sítios incipientes - sendo Taiwan e Hong Kong as províncias mais avançadas (WANG; CHEN; RICHARDS, 2018).

No Oriente Médio, os países do Conselho de Cooperação do Golfo (CCG) - Bahrain, Kuwait, Catar, Omã, Arábia Saudita e Emirados Árabes Unidos - estão em fase inicial, não despertando muito interesse acadêmico (SAXENA, 2017). Na Arábia Saudita, por exemplo, a oferta de dados é uma iniciativa unilateral do governo, que ignora qualquer envolvimento ou participação da sociedade (ALTAYAR, 2018).

No Brasil, o instrumento oferecido pelo governo para prover acesso e uso aos dados/informações públicas é o Portal Brasileiro de Dados Abertos (BRASIL, 2021b),

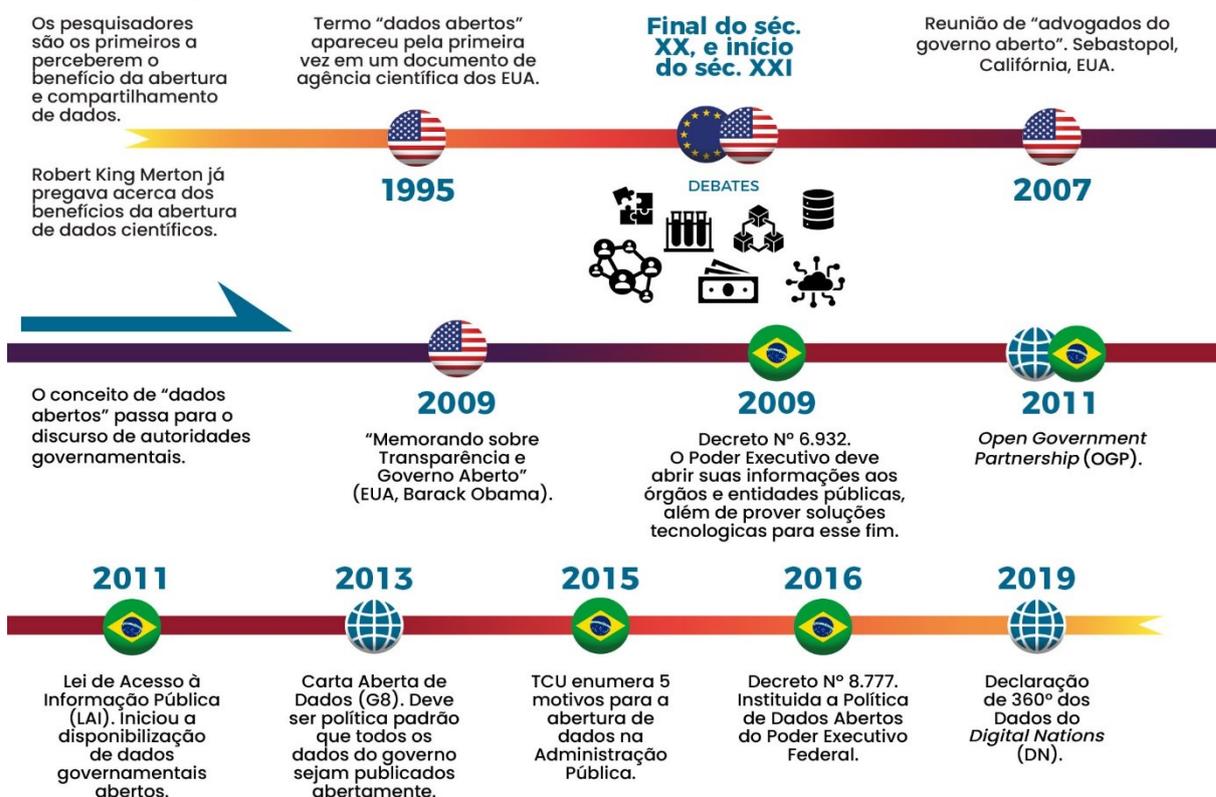
⁶ Recurso que proporciona interoperabilidade entre as aplicações desenvolvidas por terceiros e sua conexão com os *datasets* publicados.

sendo que o Poder Executivo do país aborda em seu Plano de Dados Abertos (PDA) uma determinação que obriga cada órgão/entidade a elaborar um plano de abertura com vigência de dois anos (BRASIL, 2021a).

No entanto, ao se analisar o monitor da Controladoria Geral da União (CGU) para controle das iniciativas de abertura do Poder Executivo Federal (BRASIL, 2021c), constatou-se que até agosto de 2021 registrava-se 126 órgãos que ainda não haviam criado um PDA, contra 103 órgãos com seus PDA's publicados. Partindo do painel da CGU, Silva e Pinheiro (2019) observaram a existência de *datasets* incompletos, desatualizados e sem qualquer aderência com o LOD. Contudo, atestam o valor informacional dos dados para a sociedade, como dados sobre recursos hídricos e de orçamento federal.

A partir dos resultados alcançados com a revisão de literatura, o progresso da abertura de dados governamentais foi organizado em uma linha cronológica (Figura 1) evidenciando marcos de abertura de dados tanto no Brasil quanto no exterior.

Figura 1 - Cronologia com os principais marcos de abertura de dados



Fonte: Elaborado pelo autor (2021).

Mesmo já sendo objeto de interesse e discussão no meio científico, a temática *Dados Abertos* ganha força apenas no final do século XX, inicialmente em países com

grande desenvolvimento tecnológico e engajamento de atores da sociedade. Ao final da primeira década do século XXI, os governos nacionais passam a adotar em seus discursos os benefícios da abertura de dados, incluindo o tema nas agendas governamentais.

Observa-se, no entanto, um grande salto no movimento a partir da segunda década do século XXI, com rápida expansão em alcance global entre as nações. As normatizações legais, a criação de políticas de abertura, a governança e os protocolos técnicos amadurecem a partir desse período, o que o caracteriza ainda como um fenômeno muito recente, passível, portanto, de muitos desafios.

Com relação aos desafios enfrentados pelo movimento, os resultados alcançados pela revisão de literatura também permitiram mapear e organizar no Quadro 3 as principais barreiras para a abertura de dados abertos governamentais em âmbito mundial.

Quadro 3 - Desafios para a abertura de dados governamentais

(continua)

Desafios	Descrição	Referências
Infraestrutura tecnológica dos dados	O gerenciamento dos dados governamentais se mostra deficiente e inconsistente devido a alterações frequentes nos <i>datasets</i> e manutenção nos servidores de dados.	(DAVIES, 2010; WORLD WIDE WEB FOUNDATION, 2018)
Acesso aos dados e usabilidade	A interface de acesso aos dados não possui um bom projeto de design de experiência, afetando a capacidade de localização, consumo e integração. Os portais geralmente possuem uma abordagem técnica e orientada a desenvolvedores, não contemplando os usuários finais.	(DAVIES, 2010; CHIGNARD, 2013; KITCHIN, 2013)
Publicação	Os <i>datasets</i> são geralmente incompletos, sem informações adicionais, apresentando falta de padrão e muitas vezes sem metadados significativos, evidenciando problemas de interoperabilidade e em fontes de referência comuns. As publicações de forma geral não utilizam todo o potencial da Web de Dados e os preceitos do LOD, essenciais para conectar <i>datasets</i> heterogêneos e descobrir recursos relacionados.	(DAVIES; FRANK, 2013; CHIGNARD, 2013; W3C, 2017)
Políticos e sociais	Os dados de maior impacto social ainda permanecem amplamente fora do escopo de abertura. A abertura de dados capacita os empoderados, fortalecendo ainda mais aqueles que já têm poder e privando os demais.	(CHIGNARD, 2013; KITCHIN, 2013; WORLD WIDE WEB FOUNDATION, 2018)
Econômicos	Os dados abertos carecem de um modelo financeiro sustentável, além de sofrerem críticas por fomentarem a redução do papel do estado através da comercialização dos serviços públicos.	(KITCHIN, 2013)

Quadro 3 - Desafios para a abertura de dados governamentais

(conclusão)

Desafios	Descrição	Referências
Organização e processos internos	Se as autoridades públicas não perceberem o valor para abrir dados, não alocarão os recursos necessários para garantir a sua realização. Soma-se a isso, a necessidade de proteger seus interesses pessoais como uma forte razão para que mantenham alguns dados secretos ou inacessíveis.	(DAVIES; BAWA, 2012)
Legais e privacidade	A natureza dos dados abertos impõe novas demandas ao governo para prover privacidade, segurança e impedir o uso indevido das informações. Os governos devem determinar a política sobre quais dados devem ser compartilhados e suas circunstâncias adequadas	(DAVIES; BAWA, 2012; CHIGNARD, 2013; W3C, 2017).
Capacitação e treinamento	Grande parte dos usuários não possui as habilidades necessárias para avaliar a qualidade e a adequação ao uso de dados, além de não possuir o alcance do que podem fazer com informações associadas. Soma-se a isso, a necessidade de estudantes recém-formados e profissionais que lidam com tratamento e análise de dados serem devidamente treinados e capacitados em várias competências interdisciplinares.	(WORLD WIDE WEB FOUNDATION, 2018; BLANK, 2019).

Fonte: Elaborado pelo autor (2021).

O Quadro 3 ajuda a compreender de forma sucinta quais as principais áreas em que os desafios se fazem presentes, apontando barreiras de natureza informacional, organizacional, política, social, legal e tecnológica. Percebe-se que apesar do desenvolvimento da área de tratamento e publicação dos dados, os governos ainda demandam por políticas públicas de dados mais abrangentes, com conjuntos de atividades e ações que envolvam uma maior participação dos membros que compõem a sociedade, além de ser substancial uma maior transparência na execução das ações que os governos formalizam nos documentos sistematizadores de políticas.

O Estado da Arte explicitou a forma precária com que grande parte das nações vêm abrindo seus dados, tanto em aspectos tecnológicos (variados tipos de soluções) e licenças de uso (não aberto) para acesso e reúso de dados, quanto nas práticas de organização da informação estabelecidas. Observou-se a utilização de vários tipos de soluções tecnológicas para publicação de dados, o que acarreta a produção de conteúdo em diversos formatos e linguagens, gerando desafios às instituições de governo com interesse na publicização de dados em rede.

As quedas de representatividade de nações pioneiras e ricas como EUA e Reino Unido também mostram que ser detentora de grande tecnologia não é sinônimo

de qualidade para abertura. Falta maior comprometimento político de estruturas administrativas com a qualidade dos dados disponibilizados seguindo os princípios do movimento aberto. O objetivo é assegurar o direito de acesso aos dados de acordo com os preceitos e definições consolidados ao longo dos anos graças ao envolvimento de grande espectro da sociedade.

Em contrapartida, é perceptível que algumas nações como Irlanda, França, Canadá e Coréia do Sul tomaram o protagonismo para o avanço do movimento. Identificou-se nesses países o claro respaldo governamental na construção das políticas públicas, com respeito à transparência, à disponibilização dos dados e boas práticas de governança.

Em relação ao tratamento ou curadoria dos dados presentes nas bases de dados dos governos pesquisados, percebe-se lacunas associadas a aspectos semânticos da informação, o que vai dificultar questões ligadas à interoperabilidade entre as comunidades de governo para prover condições necessárias à integração plena de informações heterogêneas entre sistemas e bases de dados disponíveis na Web (BIZER; HEATH; BERNERS-LEE, 2009).

Percebe-se a ausência de modelos ou padrões de organização e representação da informação usados para a constituição de bases de dados. Poder-se-ia explorar o uso de regras de catalogação, padrões de metadados, vocabulários e ontologias (LEMOS; MENDONÇA; SOUZA, 2020), vistos como instrumentos de organização seminais ao campo da Ciência da Informação que ajudam nos processos de classificação e indexação de dados na Web.

No Brasil, a evolução continua visto que, alguns dispositivos legais contribuem para normatizar a abertura de dados nos portais eletrônicos de transparência governamental do Brasil, podendo ser observada a sua manifestação tanto na União, quanto nos Estados, Municípios e Distrito Federal. O Quadro 4 apresenta algumas iniciativas de projetos voltados à publicização de dados abertos governamentais nas esferas públicas brasileiras.

Quadro 4 - Quadro sinóptico sobre projetos de dados governamentais nas esferas públicas brasileiras

(continua)

Projetos	Escopo	Tecnologias utilizadas	Modelagem de dados	Formato de dados	Meios de acesso
Sistema Fiscalize	Trata informações orçamentárias da câmara dos deputados.	Extraí informação dos dados orçamentários gerando relatórios via <i>Oracle Reports</i> .	Modelagem relacional de bancos de dados.	PDF.	Disponível na Web. Acesso direto aos dados.
Sistema Siga Brasil	Trata informações orçamentárias do senado federal.	Gera relatórios dinâmicos através de técnicas de modelagem dimensional de dados e da tecnologia <i>Business Objects</i> .	Modelagem relacional de bancos de dados.	PDF, XLS e CSV.	Disponível na Web. Acesso indireto aos dados via aplicação.
Portais de dados abertos estaduais	Trata de dados abertos de sete Estados: SP, RS, PE, MG, GO, AL e ES.	Os <i>datasets</i> permitem gerar relatórios, com informações sobre os dados procurados.	Modelagem relacional de bancos de dados.	PDF, DOC, HTML, CSV, JSON, XML, ODS, XML, RTF, TXT, KML, GEOJSON e PNG.	Disponível na Web. Acesso direto aos dados.
Portal de dados abertos da câmara dos deputados	Trata de dados abertos legislativos. Conjunto de dados tratado: "Deputados".	O <i>dataset</i> disponibilizado permite gerar relatórios, incluindo informações (metadados) sobre os parlamentares.	Modelagem entidade-relacionamento.	XML, JSON, CSV, XLSL e ODS.	Disponível na Web. Acesso direto aos dados. Cada deputado possui um URI.
Portais governamentais	Trata dos sítios oficiais do governo eletrônico da esfera federal.	Heterogeneidade nas tecnologias utilizadas. Algumas permitem buscas avançadas e cruzamento de dados com bases externas, outras disponibilizam dados "fechados" atrapalhando a reutilização.	Modelagem relacional de bancos de dados.	Muitos dados em formato fechado. Alguns documentos impressos digitalizados em formato de imagem.	Disponível na Web. Acesso ativo: direto aos <i>datasets</i> disponíveis. Acesso passivo: acesso a <i>datasets</i> ainda indisponíveis via solicitação.
Governo eletrônico no Rio Grande do Norte	Trata de dados abertos em 46 sítios eletrônicos de prefeituras.	Os <i>datasets</i> disponibilizados permitem gerar relatórios, incluindo informações sobre os dados.	Modelagem relacional de bancos de dados.	Diversos formatos eletrônicos, incluindo abertos.	Disponível na Web. 78,3% dos sítios possibilitam acesso direto.

Quadro 4 - Quadro sinóptico sobre projetos de dados governamentais nas esferas públicas brasileiras

(conclusão)

Projetos	Escopo	Tecnologias utilizadas	Modelagem de dados	Formato de dados	Meios de acesso
Portal Brasileiro de Dados Abertos	Trata de 25 <i>datasets</i> disponibilizados (Agricultura, Saúde, Infraestrutura, Educação).	Plataforma CKAN. Ao selecionar um recurso em um <i>dataset</i> é exibida uma página com outros elementos de metadados descritivos relacionados ao arquivo contendo os dados.	Modelagem relacional de bancos de dados.	CSV, JSON, XML, HTML, PDF, ZIP, DATA.	Disponível na Web. Acesso direto aos dados.
Portais do governo federal	Trata dos sítios do governo eletrônico federal e do portal de dados abertos.	Prevê a prospecção de ontologias e vocabulários existentes a serem reutilizados para publicação de dados governamentais abertos.	Modelagem relacional de bancos de dados.	Maior parte dos dados em formatos proprietários.	Disponível na Web. Acesso direto (com outros indiretos) aos dados.
Guia de aplicativos do governo federal	Trata de aplicativos para dispositivos móveis criados por órgãos do executivo.	Tecnologia <i>Android</i> em mais de 90% dos aplicativos. Fonte de dados está relacionada aos conteúdos produzidos pelo órgão que construiu o aplicativo.	Modelagem relacional de bancos de dados.	Formatos sem padronização, alguns em PDF.	Dos 129 aplicativos, 103 possuem acesso direto na plataforma <i>Android</i> .
Publicação de dados abertos e semânticos no Brasil	Trata do Portal Brasileiro de Dados Abertos (PBDA) e outras iniciativas no país.	Plataforma CKAN no PBDA e Alagoas. Outras iniciativas utilizam soluções próprias que permitem gerar relatórios.	Modelagem relacional de bancos de dados.	Principais: CSV e HTML. Sem um formato padrão.	Disponível na Web. Acesso direto aos dados.
Aplicativos cívicos utilizando dados abertos do governo	Trata de 18 aplicativos sobre financiamento de campanha eleitoral, atuação política etc.	Plataforma CKAN; aplicativos: iOS, <i>Android</i> , <i>Windows</i> . <i>Linked Data</i> nos padrões: XML, URI, SPARQL, RDF. Ontologias no padrão OWL.	Modelagem relacional de bancos de dados. Prevê modelo ontológico.	Utilizam os formatos existentes nas diversas bases.	Disponível na Web. Acesso direto aos dados.
Portais de transparência da administração pública	Trata dos portais de dados abertos do Brasil e dos Estados de SP, MG, RJ, BA, RS, e suas capitais.	Plataforma CKAN no Brasil. Em outras iniciativas são plataformas próprias para se gerar relatórios, incluindo informações sobre os dados procurados.	Modelagem relacional de bancos de dados.	PDF, CSV, ODS, HTML, XLS, TXT, XML, JSON.	Disponível na Web. Acesso direto aos dados. Maior parte gratuita.

Fonte: Adaptado de Macedo e Lemos (2019).

Os resultados expostos por Macedo e Lemos (2019) sobre projetos voltados à publicização de dados abertos governamentais nas esferas públicas brasileiras sugerem que a preocupação no Brasil é dispor os dados na Web em formatos não proprietários e legíveis por máquina. Porém, em relação aos preceitos do LOD, existe uma baixa adesão por parte das instâncias governamentais brasileiras em oferecer os dados abertos em um formato compatível. A situação se agrava nas esferas estaduais e municipais, onde a padronização é praticamente nula.

O estudo ratifica que apenas a formalização de uma política pública de dados não tem se mostrado suficiente para abrir dados com a qualificação desejável pelas partes interessadas na reutilização do conteúdo governamental. Nesse sentido, as ações mais exitosas de abertura de dados nas esferas internacionais estão cada vez mais alinhadas aos interesses da abertura informacional, o que vem impactando positivamente todo o ciclo de vida dos dados. Essa visão é ratificada pela “Declaração de 360° dos Dados” elaborada em 2019 pelas Nações Digitais (DIGITAL NATIONS, 2019), cúpula de países que enfatizam a necessidade de uma maior abrangência no gerenciamento de dados abertos governamentais.

Países como Irlanda e Espanha já demonstraram ser possível incluir no desenvolvimento de ações a participação de iniciativas privadas de utilidade pública com origem na sociedade. Acrescenta-se a importância da promoção de eventos como os *hackathons* (CRUSOE; AHLIN, 2019), por agregar entusiastas a fim de explorar dados abertos, e assim propor projetos de *software* ou *hardware* fazendo uso desses dados.

A partir da conclusão do estudo tem-se o seguinte recorte com sugestões de aprimoramento: i) maior colaboração entre as organizações interessadas, focando no impacto das políticas para se ter resultados mais concretos; ii) criação de uma cultura de dados incorporada como padrão dentro das rotinas de trabalho; iii) mudança na mentalidade de abertura por “obrigação” para uma cultura em que a divulgação de dados seja incentivada, tornando claro além do aspecto legal, a sua importância social; iv) harmonização dos interesses políticos por trás da abertura de dados, frente à complexidade e os riscos que os servidores públicos enfrentam ao abrir os dados, como exposição de dados e privacidade.

3 FUNDAMENTAÇÃO TEÓRICA

O presente capítulo constitui-se de uma abordagem teórica conceitual fundamental para embasar os procedimentos metodológicos e o desenvolvimento dos produtos da pesquisa. Assim, visando a um encadeamento salutar no percurso teórico metodológico adotado, o Capítulo foi organizado em três seções, conforme apresentado a seguir. A Seção 3.1 trata da organização e representação da informação e do conhecimento em contextos digitais, apontando assuntos seminais aos campos da Ciência da Informação e da Ciência da Computação sobre desafios inerentes à organização e tratamento da informação em ambiente web. A Seção 3.2 explana sobre modelo e processo de governança no gerenciamento dos dados como ativo estratégico. E a Seção 3.3 enfatiza as contribuições e implicações do Capítulo ao desenvolvimento da presente pesquisa.

3.1 ORGANIZAÇÃO E REPRESENTAÇÃO DA INFORMAÇÃO E DO CONHECIMENTO EM CONTEXTOS DIGITAIS

A necessidade de organizar conteúdos de qualquer natureza é inerente à atividade humana, observada por toda sua rica trajetória histórica e cultural. Sobre a abordagem humana a respeito da aquisição e organização do conhecimento Sayão (2001, p. 82) comenta, “Constroem-se, assim, esquemas abstratos da realidade, nos quais as coisas são reduzidas a seus perfis mais convenientes”. Não apenas limitando-se a organizar todo o conhecimento acumulado durante séculos, o homem desenvolveu formas de conseguir representá-lo.

Essa característica tornou possível um melhor raciocínio sobre as coisas do mundo a partir de processos de formalização sobre objetos e suas relações (CAMPOS, 2004). Esse conjunto de ações permitiu a humanidade melhorar a compreensão acerca dos fenômenos, contribuindo para uma transmissão eficiente da informação advinda de todo o conhecimento acumulado e organizado.

O recente fenômeno da “explosão da informação” (LE COADIC, 1996; SARACEVIC, 2009) repercutiu de forma dramática no mundo contemporâneo, afetando a adequada organização e representação dos volumosos estoques gerados e sua posterior recuperação. Estamos vivendo uma era em que a informação se caracteriza como um objeto de consumo cobiçado e indispensável para a realização

de diversas atividades do cotidiano, demandadas por corporações, governos e indivíduos.

O grande volume de informação resultante dessa produção e consumo vem sendo objeto de estudo por parte da Ciência da Informação (CI), sendo parte integrante de sua gênese. Nesse contexto, faz-se necessário desenvolver métodos eficientes para a sua organização, representação, armazenamento e posterior recuperação, sejam na produção científica, documental ou quaisquer outros registros de memória em suportes físicos ou digitais.

De acordo com Saracevic (1996), a CI busca tornar acessível um acervo crescente de conhecimento, provendo soluções para tais desafios. Já a representação do conhecimento deixou de ser uma preocupação exclusiva do mundo da documentação, tornando-se relevante para outros domínios, como na estrutura de banco de dados, estrutura sintática e semântica da linguagem natural e incluindo também problemas comuns na representação em sistemas de informação, bem como em sistemas especialistas da Inteligência Artificial (VICKERY, 1986).

Saracevic (2009) esclarece que um dos pilares da CI é justamente prover soluções aos desafios oriundos da explosão universal da informação, potencializada em nossos tempos pelo pujante ambiente digital e seus intermináveis formatos gerados. Portanto, com relação às novas formas de produção e difusão de conteúdos digitais, a complexa realidade em que se apresenta o ambiente eletrônico vem exigindo representações mais sofisticadas para uma organização e recuperação da informação eficientes.

Os desafios inerentes à organização e representação são comuns tanto a CC quanto a CI. De acordo com Campos (2004), estes desafios na CC servem para auxiliar a implantação de estruturas computáveis, enquanto na CI possibilitam a elaboração de linguagens documentárias verbais e notacionais, visando à recuperação de informação e à organização dos conteúdos informacionais de documentos.

A CI debruça-se na elaboração de variados tipos de conhecimentos no esforço hercúleo de organizar domínios distintos, e assim atender a uma de suas principais premissas que é a recuperação da informação (LANCASTER, 1986).

A partir desse contexto, Hjørland (2002, p. 422) apresenta onze abordagens para a análise de domínio, quais sejam: i) produção de guias de literatura e portais de assuntos; ii) produção de classificações especiais e tesouros; iii) pesquisa sobre

indexação e recuperação de especialidades; iv) estudos empíricos de usuários; v) estudos bibliométricos; vi) estudos históricos; vii) estudos de documentos e gêneros; viii) estudos epistemológicos e críticos; xix) estudos terminológicos, linguagens para finalidades especiais, estudos do discurso; x) estudos de estruturas e instituições de comunicação científica; e xi) análise de domínio em cognição profissional e inteligência artificial.

A presente pesquisa se apoia em duas dessas abordagens: i) as classificações especiais e tesouros: organizam as estruturas lógicas de categorias e conceitos em um domínio, bem como as relações semânticas entre os conceitos; e ii) as especialidades de indexação e recuperação: organizam documentos ou coleções únicas para otimizar a capacidade de recuperação e visibilidade de seus “potenciais epistemológicos” específicos.

Na visão de Hjørland (2003), a Organização do Conhecimento (*Knowledge Organization* - KO) em CI significa especialmente a organização de informações em registros bibliográficos (índices de citações, registros completos de texto e a Internet). Como afirma Hjørland (2007, p. 369), “O campo de KO em sistemas de informação está, portanto, preocupado com a construção, uso e avaliação de ferramentas semânticas para recuperação da informação”.

A combinação das competências oriundas da CC e CI está resultando em soluções cada vez mais eficientes para a organização e representação da informação e do conhecimento, refletindo em sofisticados sistemas de recuperação da informação (SRI). Pesquisas recentes (LEMOS; SOUZA, 2019; LEMOS; SOUZA, 2020b) apontam para diversas iniciativas nos campos da CI e CC voltadas à exploração semântica da informação com o propósito de melhorar a eficácia dos SRI's. Podem-se destacar:

- web Semântica e sua proposta emergente de dados abertos conectados;
- instrumentos voltados à representação de relacionamentos semânticos e conceituais, incluindo ontologias e vocabulários controlados. Propõe-se resolver problemas de interoperabilidade em sistemas e bases de dados, além das dificuldades intrínsecas da linguagem natural;
- modelos conceituais, de referência e ontológicos que orientam a modelagem da realidade documental e o processo de busca e recuperação da informação em contextos digitais.

Segundo Lemos e Souza (2018), os processos, produtos e teorias subjacentes à organização e representação da informação e do conhecimento em sistemas de

informação podem ser arranjados nas seguintes disciplinas, sendo objetos de estudo tanto da CI quanto da CC:

- **representação descritiva:** destacando o conceito “metadados” e seus padrões de descrição;
- **representação temática:** destacando o papel dos vocabulários controlados nos SRIs;
- **modelos para a organização da realidade documental:** destacando o uso de modelos conceituais baseados em ontologias.

O uso combinado de cada um desses ordenamentos tem como objetivo aprimorar a representação da informação, permitindo uma interoperabilidade mais eficiente entre diferentes sistemas de informação na rede, e, finalmente, compondo importante requisito na proposição de um modelo de governança voltado a qualidade de dados para publicação em rede.

Assim sendo, a Seção 3.1.1 busca elucidar sobre metadados e seu papel em descrever recursos de informação; a Seção 3.1.2 trata da utilização de vocabulários controlados como uma estrutura “semântica” que visa descrever mais precisamente os conteúdos envolvidos nos recursos de informação; a Seção 3.1.3 aborda o emprego de modelos na organização e representação da realidade documental, incluindo vocabulários, padrões de metadados e ontologias no suporte à modelagem conceitual de domínios; e a Seção 3.1.4 disserta sobre os conceitos do *Linked Data*, apresentando o expoente máximo da Web de Dados: o *Linked Open Data*.

3.1.1 Metadados na representação descritiva de documentos

Metadados em sua origem etimológica são dados sobre dados (IFLA, 2005; GILLILAND, 2016; RAUTENBERG *et al.*, 2018). Segundo Siqueira e Silva (2011, p. 11) metadados são “[...] informações estruturadas que descrevem, identificam, localizam ou tornam mais fácil a recuperação, o uso ou o gerenciamento de fontes de informação digital”, atuando, portanto, de forma decisiva na representação descritiva de documentos.

Os metadados são um produto do desenvolvimento histórico de regras de catalogação (SIQUEIRA; SILVA, 2011), mantendo conexões evolutivas com códigos como: i) *International Standard Bibliographic Description* (ISBD); ii) *Anglo-American Cataloguing Rule* (AACR2); e iii) *Machine-Readable Cataloging* (MARC).

A Web remodelou os meios de descrever e gerenciar recursos de informação, nesse sentido os metadados despontam como um instrumento primário neste esforço e um importante elo na cadeia de valor das economias do conhecimento (DUVAL *et al.*, 2002). Desde então, o metadado encontra-se inserido na raiz da Internet como suporte de acesso a recursos digitais na Web, porém, com a explosão informacional da Web e as novas demandas, os processos anteriores de geração de metadados mostram-se insuficientes e incompatíveis (SIQUEIRA; SILVA, 2011).

Com novos conteúdos sendo produzidos nos mais diversos formatos multimídia, fica cada vez mais importante estabelecer padrões de metadados para que sistemas distintos possam trocar dados sem prejuízo de perda de conteúdo, sendo seu uso fundamental para prover interoperabilidade.

Sobre a importância dos metadados, Riley (2017, p. 6) é claro ao afirmar que “[...] interoperabilidade, a troca eficaz de conteúdo entre sistemas, conta com metadados que descrevem esse conteúdo para que os sistemas envolvidos possam efetivamente perfilar o material recebido e combiná-lo com suas estruturas internas”.

Gilliand (2016) propõe delimitar os metadados em categorias distintas, sendo organizados para aplicações e finalidades distintas conforme disposto no Quadro 5.

Quadro 5 - Categorias de metadados

Categoria	Definição dos metadados
Administrativo	Metadados usados no gerenciamento e administração de coleções e recursos de informação.
Descritivo	Metadados usados para identificar, autenticar e descrever coleções e recursos de informações confiáveis relacionados.
Preservação	Metadados relacionados ao gerenciamento de preservação de coleções e recursos de informação.
Técnico	Metadados relacionados a como um sistema funciona ou se comporta.
Uso	Metadados relacionados ao nível e tipo de uso de coleções e recursos de informação.

Fonte: Adaptado de Gilliland (2016).

A proposta de categorias de metadados fornecido por Gilliland (2016) procura refletir os aspectos principais de sua funcionalidade, auxiliando em seu reuso e compreensão. Ainda segundo a autora, todos os objetos de informação em geral, independentemente da forma física ou intelectual, possuem três recursos (conteúdo, contexto e estrutura) que podem e devem ser refletidos por metadados.

Em recente projeto aplicado sobre ativos de patrimônio cultural aberto, Lemos, Sá e Souza (2019) recomendam o uso de padrões de metadados e ontologias multimídia bem fundamentadas para documentação de acervos culturais em 3D, pois demonstram ser uma solução para descrição de conteúdo multimídia processável por máquina e baseada em semântica.

Alguns dos padrões de metadados destacados (LEMOS; SÁ; SOUZA, 2019) são: i) MPEG-7, como um padrão que busca prover um vocabulário rico e comum para multimídia, incluindo descritores primitivos (extraídos da própria mídia) e de alto nível (destinados à descrição semântica do conteúdo da mídia); ii) *Dublin Core*, como padrão para descrever recursos na Web, sendo que na atualidade já é uma realidade esforços para sua integração com a linguagem *Resource Description Framework* (RDF); iii) *VRA Core*, como um padrão de metadados reconhecido internacionalmente, usado para objetos de patrimônio cultural, incluindo: pinturas, desenhos, esculturas, arquitetura, fotografias. Outros padrões apontados por Gilliland (2016) são: MARC, BIBFRAME, EAD, LIDO, *Australian Recordkeeping Metadata Schema*.

A autora também propõe uma tipologia com objetivo de organizar padrões para tratamento de dados em categorias, conforme sistematizadas no Quadro 6.

Quadro 6 - Tipologia para padrões de dados

(continua)

Padrões	Exemplos
Padrões de estrutura de dados (conjuntos de elementos de metadados, esquemas). Essas são “categorias” ou “contêineres” de dados que compõem um registro ou outro objeto de informação.	Formato MARC (Catalogação legível por máquina), Descrição de arquivamento codificado (EAD), BIBFRAME (Estrutura bibliográfica), <i>Dublin Core Metadata Element Set</i> , Categorias para a descrição de obras de arte, <i>VRA Core</i> .
Padrões de valor dos dados (vocabulários controlados, dicionário de sinônimos, listas controladas). Estes são os termos, nomes e outros valores que são usados para preencher padrões de estrutura de dados ou conjuntos de elementos de metadados.	Cabeçalhos de assunto da Biblioteca do Congresso, Arquivo de autoridade de nomes e <i>Thesaurus</i> para materiais gráficos; <i>Getty Art & Architecture Thesaurus</i> , União da lista de nomes de artistas (ULAN) e <i>Thesaurus</i> de nomes geográficos; <i>ICONCLASS</i> ; Cabeçalhos de Assunto Médico.

Quadro 6 - Tipologia para padrões de dados

(conclusão)

Padrões	Exemplos
Padrões de conteúdo de dados (regras e códigos de catalogação). Estas são diretrizes para o formato e sintaxe dos valores de dados usados para preencher os elementos de metadados.	Regras de catalogação anglo-americanas, descrição e acesso de recursos, descrição bibliográfica do padrão internacional, catalogação de objetos culturais e descrição de arquivos: um padrão de conteúdo.
Formato de dados / padrões de intercâmbio técnico (padrões de metadados expressos em formato legível por máquina). Esse tipo de padrão geralmente é uma manifestação de um padrão específico da estrutura de dados, codificado ou marcado para o processamento da máquina.	Recurso Descrição Estrutura, MARC21, MARCXML, EAD XML DTD, METS, BIBFRAME, LIDO XML, XML simples de Dublin Core, XML qualificado de Dublin Core, XML qualificado de Dublin, XML do VRA Core 4.0.

Fonte: Adaptado de Gilliland (2016).

A organização sistematizada no Quadro 6 opera de modo a fornecer maneiras padronizadas de estruturar e expressar esses padrões para processamento, publicação e implantação computacional. Mesmo com padrões de metadados mais flexíveis e modulares como o *Dublin Core*, o arcabouço da Web demandará inúmeros esquemas de metadados para as diversas necessidades existentes, incluindo dados (SIQUEIRA; SILVA, 2011).

Portanto, torna-se relevante dentro dos estudos da organização e representação da informação e do conhecimento o aperfeiçoamento desses recursos para a melhoria contínua da recuperação de conjuntos de dados em sistemas de informação.

3.1.2 Vocabulários controlados na normalização de conteúdo documental

Além da criação e padronização, os produtores desejam normalizar seus metadados através do controle dos valores reais usados; o uso de vocabulários controlados é uma forma de se fazer isso (RILEY, 2017).

Vocabulário controlado é uma lista predeterminada de termos que pode incluir uma forma de estrutura semântica (LANCASTER, 2004; RILEY, 2017). Riley (2017, p. 17) destaca que “Essas listas geralmente identificam uma palavra ou frase preferida

para um determinado conceito e, às vezes, fornecem mapeamentos de outros termos para o conceito preferido”. O uso de instrumentos como vocabulários controlados atenua os indesejáveis efeitos da heterogeneidade semântica na recuperação da informação – polissemia, sinonímia, fatores hierárquicos – frutos da natureza ambígua da linguagem natural (LEMOS; SOUZA, 2018).

Para Riley (2017), os vocabulários controlados possuem em seu escopo apenas um único idioma, apesar de grandes comunidades como a ISO trabalharem para vinculá-los a outros ou definir novos vocabulários para organizar termos de idiomas distintos. Muitas categorias de vocabulários, quase sinônimos, têm sido cunhadas, como: i) ontologias; ii) vocabulários controlados; iii) e outros como: tesouros, taxonomias, listas de código, redes semânticas. (RAUTENBERG *et al.*, 2018; W3C, 2017).

Para a *National Information Standards Organization* (NISO), os vocabulários controlados “[...] são utilizados para a representação de objetos de conteúdo em sistemas de organização do conhecimento e incluem listas, anéis de sinônimo, taxonomias e tesouros” (ANSI, 2005, p. 1).

Segundo a ANSI (2005, p. 18), “Taxonomia é um vocabulário controlado que consiste em termos preferenciais, todos conectados em uma hierarquia ou poli hierarquia”, enquanto o Tesouro, de acordo com a ANSI (2005, p. 18), seria um “[...] vocabulário controlado, organizado em uma ordem conhecida e estruturado para que as várias relações entre os termos sejam exibidas claramente e identificadas por indicadores padronizados de relacionamento”.

No escopo da Web Semântica, para que seu cenário fique organizado e bem representado é preciso que se estabeleça um conjunto de vocabulários de referência (incluindo vocabulários controlados, padrões de metadados e ontologias), como maneira de facilitar a interoperabilidade entre metadados.

No Quadro 7 apresenta-se alguns vocabulários de referência usados em ambientes digitais.

Quadro 7 - Exemplos de vocabulários de referência em ambientes digitais

Vocabulário	Descrição
Dublin Core	Iniciativa pioneira na definição de um vocabulário para a descrição de metadados. Indicado para a descrição de recursos (documentos).
Friend of a Friend (FOAF)	Vocabulário adequado para a definição de metadados sobre pessoas, seus interesses, relacionamentos e atividades. O vocabulário tem um conjunto central de classes e propriedades.
Simple Knowledge Organization System (SKOS)	Fornecer uma maneira de representar esquemas de classificação, tais como vocabulários controlados, taxonomias e tesouros. Muitos desses sistemas compartilham uma estrutura semelhante e são usados em aplicações similares. SKOS captura grande parte dessas semelhanças e as torna explícitas, permitindo o compartilhamento de dados entre essas diversas aplicações. Os diversos conceitos podem ser relacionados semanticamente entre si, em hierarquias e redes de associação informal, e agrupada em esquemas conceituais.
Schema.org	Fornecer uma coleção de vocabulários que podem ser utilizados para embutir metadados em páginas Web, e são entendidas pelas principais máquinas de busca. Os metadados podem ser embutidos utilizando microdados, RDFa ou JSON-LD.
PROV	Vocabulário que se refere à proveniência dos dados: quem gerou; como foram gerados; quais foram às fontes.
DCAT	Permite a criação de catálogos com descrições de conjuntos de dados. O uso de uma forma padrão para a descrição dos catálogos aumenta a capacidade de descoberta e permite que as aplicações possam encontrar metadados distribuídos em diferentes catálogos.

Fonte: Adaptado de Laufer (2015).

Para Lemos e Souza (2018), o foco na organização de conteúdos de documentos redundante em linguagens documentárias que podem auxiliar:

- nos processos de análise e descrição de documentos, permitindo a criação padronizada de metadados ao nomear, de forma consistente, os pontos de acesso aos documentos e a informação neles contida;
- no processo de busca em um SRI através da padronização e expansão do vocabulário das consultas.

Vocabulários controlados capturam um consenso da comunidade sobre um domínio específico, sendo que seu reuso aumenta a interoperabilidade e evita as redundâncias (RAUTENBERG *et al.*, 2018). Desta feita, o uso de vocabulários controlados torna-se uma peça importante dentro da organização e representação da informação e do conhecimento em sistemas de informação, evitando ruídos desnecessários na experiência entre usuários e os SRI.

3.1.3 Modelos conceituais para representação documental

Segundo Sayão (2001), modelos buscam a formalização do universo através de meios de expressões controláveis; derivam da necessidade humana de entender a complexa realidade ao seu redor, tornando-se representações simplificadas e inteligíveis do mundo. Já em sistemas de informação, modelos são:

[...] uma representação de um recorte da realidade, que, de acordo com a sua função utilitária e por meio do seu modo de expressão, sua estrutura e suas igualdades e desigualdades em relação ao seu original, tenta comunicar algo sobre o real (SAYÃO, 2001, p. 86).

Mylopoulos (1992, p. 3) faz a seguinte contribuição: “Modelagem conceitual é a atividade de descrever formalmente alguns aspectos do mundo físico e social ao nosso redor para fins de entendimento e comunicação”.

Para Sayão (2001), são três as características básicas dos modelos: i) Mapeamento: sempre modelam alguma coisa, ou seja, são representações de “originais” (ou “protótipos”), naturais ou artificiais, que, por sua vez, também podem ser modelados; ii) redução: modelos geralmente não mapeiam todos os atributos do original que eles representam, mas unicamente aqueles que são relevantes para quem modela; iii) pragmatismo: modelos em si não são pertencentes à mesma classe que seus originais, cumprindo suas funções de substituição orientadas para objetivos dependentes de operações mentais ou factuais, em um tempo específico.

Os modelos conceituais despontam em relação aos processos de representação descritiva e temática justamente por permitirem uma melhor descrição do domínio a ser representado, e, por conseguinte, preservam sua integridade semântica. Quanto mais soubermos sobre um domínio específico e quanto mais preciso for sua representação, maiores serão as chances de construir sistemas e serviços computacionais consistentes com a realidade desse domínio (GUIZZARDI, 2005). “A modelagem conceitual tem uma vantagem sobre a linguagem natural ou notações diagramáticas, na medida em que se baseia em uma notação formal que permite capturar a semântica da aplicação” (MYLOPOULOS, 1992, p. 3).

Nesse sentido, o uso de ontologias está sendo cada vez mais explorado na representação de conteúdos em suporte à modelagem conceitual de domínios. Tal característica tem feito as ontologias ganharem grande atenção como técnica para organização da informação e do conhecimento (ALMEIDA, 2014). Mendonça e

Almeida (2012) defendem o emprego de modelos baseados em ontologias para sanarem os problemas de interoperabilidade entre sistemas, que muito sofreram com a inconsistência da modelagem conceitual em seus primórdios.

Na atualidade, as ontologias fornecem um meio eficaz para representar informações de alto nível de abstração, sendo consideradas essenciais para a interoperabilidade semântica e a troca de informações entre computadores e humanos em ambientes complexos (VASSILAKIS; LEPOURAS, 2006). A importância do emprego de ontologias é ressaltada por Lemos e Souza (2019, p. 106) “Ontologias são de interesse como instrumento de organização de conhecimento que versa especialmente na análise de conceitos e relacionamentos de um domínio”.

Diferente do seu sentido filosófico, referente a um sistema particular de categorias, a ontologia (com “o” minúsculo) em sistemas de informação seria de acordo com Guarino (1998, p. 4) “[...] um artefato de engenharia, constituído por um vocabulário específico usado para descrever certa realidade, além de um conjunto de suposições explícitas sobre o significado pretendido das palavras do vocabulário”. Guarino (1998) juntamente com Guizzardi (2005) versam sobre a oportunidade de se desenvolver ou reutilizar níveis diferentes de ontologias de acordo com o seu propósito, classificando-as nas seguintes tipologias:

- **ontologias de fundamentação:** baseada nas teorias da ontologia formal, que descrevem o conhecimento sobre a realidade de uma maneira que é independente de linguagem, de estados de coisas particulares e de estados epistêmicos de agentes conhecedores. A representação dessas teorias em um artefato concreto é conhecida como uma ontologia de fundamentação;
- **ontologias de domínio:** descrevem o vocabulário relacionado em um domínio específico (como medicina ou automóveis). Tem uma posição central na visão da Web Semântica;
- **ontologias de tarefa:** descrevem o vocabulário relacionado em uma tarefa genérica (como diagnóstico ou venda), especializando os termos introduzidos na ontologia de nível superior;
- **ontologias de aplicação:** descrevem conceitos dependendo de um domínio e tarefa em particular, que geralmente são especializações de ambas as ontologias relacionadas. Em geral esses conceitos correspondem às funções desempenhadas pelas entidades do domínio durante a execução de uma determinada atividade.

Guizzardi *et al.* (2008) enfatizam a diferença do termo “ontologia” quando aplicado nas comunidades de: Computação no contexto de modelagem conceitual: usado de acordo com sua definição em Filosofia; Web Semântica: responsáveis por representar um domínio particular (biologia molecular, direito, governo) expressado em alguma linguagem de representação de conhecimento (RDF, OWL, *F-Logic*) ou de modelagem conceitual (UML, EER). Para os pesquisadores em Web Semântica, uma ontologia é um documento ou arquivo que define formalmente as relações entre termos e conceitos (SOUZA; ALVARENGA, 2004).

O emprego de ontologias, e suas modelagens semânticas oriundas de uma abordagem conceitual, é um passo natural em direção à organização e descrição dos complexos estoques informacionais existentes (especialmente em ambientes digitais como a Web) preservando a semântica, melhorando a recuperação e a interoperabilidade dos conteúdos.

Berners-Lee, Hendler e Lassila (2001) já adiantavam que a Web Semântica teria como proposta anotar os conteúdos de documentos usando ontologias de domínio, diferente do método comumente empregado para gerar conhecimento sobre documentos: os tradicionais metadados.

Como apontam Lemos e Souza (2019), as ontologias possuem os mesmos princípios dos vocabulários controlados (operam com linguagem natural e fazem a delimitação de termos e de relações); todavia, diferem entre si em relação à semântica envolvida na terminologia das linguagens documentárias nos seguintes aspectos: i) as ontologias incluem axiomas formais (através de declarações lógicas) que restringem a utilização do vocabulário; ii) outra distinção está no fato de as especificações de relações no contexto das ontologias serem em número superior as dos tesouros, por exemplo.

Assim, ontologias sendo vistas como sistemas de organização de conhecimento podem ser endereçadas, por exemplo, para organização e controle da terminologia usada em metadados destinados a mídias diversas; além de auxiliar usuários na produção de descrições mais sistemáticas e consistentes por meio de conhecimento explícito acerca de um domínio (LEMOS; SOUZA, 2019, p. 106).

O Quadro 8 apresenta alguns modelos conceituais baseados ou que possuem ontologias em seu escopo, visando uma maior integração semântica e interoperabilidade de dados em ambientes digitais.

Quadro 8 - Modelos conceituais que são baseados ou que utilizam ontologias

Modelos conceituais baseados em ontologias	
Conceptual reference model (CIDOC CRM)	Caracteriza-se como uma ontologia de alto nível para habilitar a integração de informação para dados de memória cultural e sua correlação com a informação bibliotecária e arquivística. Explica, ainda, que o modelo analisa os conceitos comuns por trás das estruturas de dados e metadados para suportar a transformação, mediação e integração de dados (DOERR, 2003).
Descriptive Ontology for Linguistic and Cognitive Engineering (DOLCE)	Pertence à biblioteca <i>WonderWeb</i> de ontologias fundamentais. Atua como ponto de partida para comparar e elucidar os relacionamentos com outras ontologias e para esclarecer as suposições ocultas subjacentes às ontologias ou recursos linguísticos existentes. Aplicado com sucesso em domínios como direito, biomedicina e agricultura (OBERLE <i>et al.</i> , 2007).
Suggested Upper Merged Ontology (SUMO)	Tentativa de vincular categorias e relações provenientes de diferentes ontologias de nível superior, a fim de melhorar a interoperabilidade, comunicação e a pesquisa na área da Web Semântica (OBERLE <i>et al.</i> , 2007).
Unified Foundational Ontology (UFO)	Ontologia de fundamentação baseada em um número de teorias das áreas de Ontologias Formais, Lógica Filosófica, Filosofia da Linguagem, Linguística e Psicologia Cognitiva. Aplicada para avaliar, reprojeter e integrar os modelos de linguagens de modelagem conceitual, e para prover semântica de mundo real para seus elementos de modelagem (GUIZZARDI <i>et al.</i> , 2008).
Modelo conceitual de dados (EDM)	Utilizado no projeto Europeia, que envolve uma biblioteca digital que oferece vasto acervo cultural <i>online</i> . Busca promover alinhamento semântico entre variados padrões de metadados (<i>Lightweight Information Describing Objects</i> – LIDO; <i>Encoded Archival Description</i> – EAD; MARC e <i>Dublin Core</i>). Ele pode ser considerado um modelo conceitual baseado em ontologias a partir do momento em que promove o enriquecimento de dados culturais por meio da estrutura semântica subjacente aos padrões W3C (LEMOS; SOUZA, 2019).
Common Data Model (CDM)	O CDM é um modelo comum de metadados dos recursos publicados pela União Europeia (EU), sendo uma ontologia baseada no modelo FRBR (<i>Functional Requirements for Bibliographic Records</i>) e descrita por meio de tecnologias RDF-S e OWL. É capaz de representar as relações entre os tipos de recursos administrados pela UE e suas visões de acordo com o modelo FRBR (UNIÃO EUROPÉIA, 2021).
ECCF Core Ontology	A <i>European Commission Conceptual Framework</i> (ECCF) é um modelo de conhecimento para descrições de instituições da UE e qualquer outra organização de serviço público. Pretende ser uma ontologia central a ser estendida com classes e propriedades específicas para cada domínio de política e serviço (UNIÃO EUROPÉIA, 2021).

Fonte: Elaborado pelo autor (2021).

Em ambientes como o LOD, o campo da linguagem é um dos principais entraves à interoperabilidade, pois cada domínio possui suas próprias nomenclaturas,

relações e definições para termos (uma mesma palavra pode assumir significados diferentes entre campos distintos). As ontologias teriam o efeito de restringir o significado de um dado conjunto de termos, no entanto as relações responsáveis pela interligação também precisam concordar entre os domínios, o que nem sempre é possível (ROCHA; CAMPOS; COSTA, 2016).

3.1.4 Dados conectados (*Linked Data*): construindo a Web de Dados

Dados Conectados (*Linked Data*) refere-se a um conjunto de práticas recomendadas para publicar e interligar dados estruturados na Web (BIZER; HEATH, 2011; AUER, 2014). A Web de Dados seria o resultado das ligações desses conjuntos de dados, podendo ser descrita com mais precisão como uma rede de coisas no mundo, descrita por dados na Web (BIZER; HEATH; BERNERS-LEE, 2009). Conectar dados passou a ser adotado por inúmeros provedores de dados, levando à criação de um espaço de dados global que contém muitos bilhões de asserções - a Web de Dados Conectados (BIZER; HEATH; BERNERS-LEE, 2009; BIZER; HEATH, 2011; AUER, 2014).

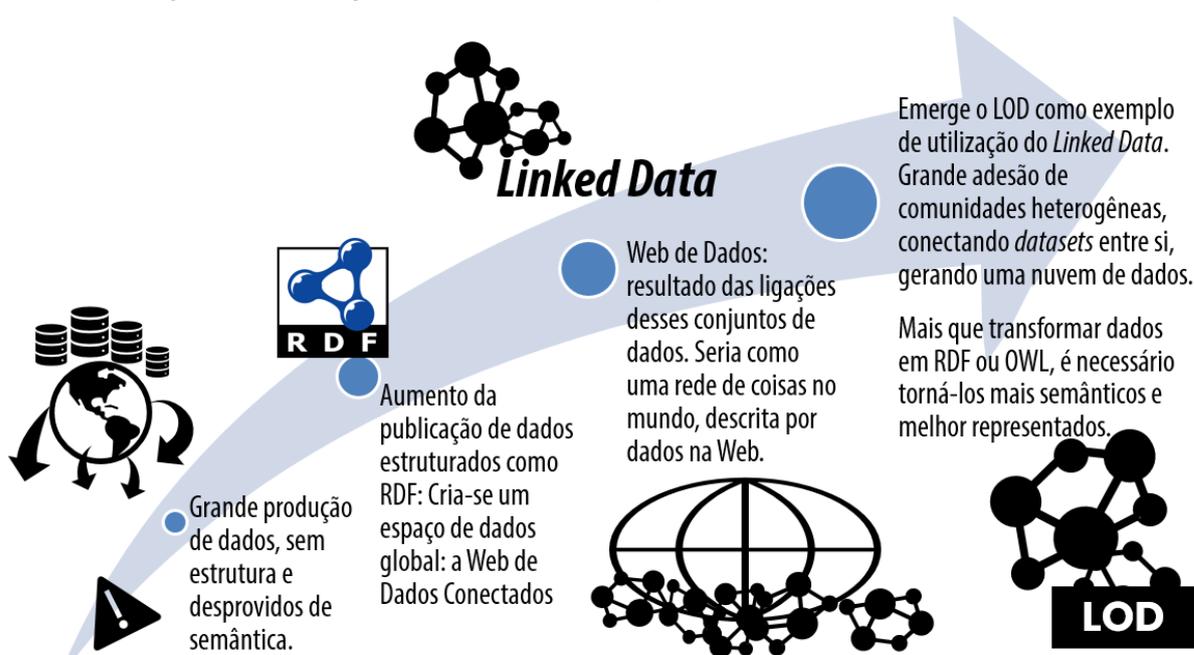
A ideia de conectar dados na Web desponta como uma promissora solução para enfrentar um dos maiores desafios na área de gerenciamento inteligente de informações: a exploração da Web como plataforma para integração de dados e informações (BERNERS-LEE, 2006; AUER, 2014). De acordo com Bizer, Heath e Berners-Lee (2009) a proposta do *Linked Data* ampliou o uso da Web, oferecendo uma alternativa sustentável e tecnológica para o grande volume de dados gerados, e já contando com muitas comunidades produtoras de conhecimento.

Os governos digitais, impulsionados pelo movimento aberto, foram uma das comunidades que mais passaram a publicar dados, tendo como uma das consequências imediatas um despejo massivo de dados na Web. Rapidamente constatou-se que esses dados foram publicados sem maiores critérios ou desprovidos de padronização, o que comprometeu grande parte de sua estrutura e semântica (BIZER; HEATH; BERNERS-LEE, 2009).

O movimento de Dados Abertos Conectados (LOD) surge nesse contexto, como uma medida que visa apoiar a publicação dos dados abertos na Web de forma mais estruturada, garantindo maior interoperabilidade, semântica e uma recuperação mais precisa desses estoques. Os governos pioneiros em conectar dados abertos

foram os EUA e Reino Unido (DING; PERISTERAS; HAUSENBLAS, 2012), sendo que o último se comprometeu em ofertar qualquer conjunto de dados brutos em dados abertos conectados (SHERIDAN; TENNISON, 2010). Toda essa escalada evolutiva é esclarecida a partir da Figura 2.

Figura 2 - Evolução continuada do compartilhamento de dados na Web



Fonte: Elaborado pelo autor (2021).

O LOD estabelece uma nova maneira de publicar e consumir conhecimento, possuindo características adequadas às necessidades governamentais, tais como: propõe o uso de padrões abertos; fornece informações úteis sobre determinadas comunidades; carrega em si o poder da colaboração; consumidores também podem ser publicadores; mais e mais dados sendo gerados e potencialmente "ligáveis" (CORDEIRO *et al.*, 2011).

Acrescentando aos preceitos do *Linked Data* já enunciados anteriormente, Bizer, Heath e Berners-Lee (2009) apontam o núcleo para conectar dados na Web e plenamente adotado pelo movimento LOD, são eles: i) refere-se aos dados publicados na Web de forma que sejam legíveis por máquina; ii) seu significado é definido explicitamente; iii) está vinculado a outros conjuntos de dados externos; iv) pode ser vinculado a partir de *datasets* externos.

Conectar dados abertos na Web irá compor a ideia por trás do *Linked Open Data*. Para o presente estudo, um conteúdo é considerado "Aberto" desde que seja acessível a qualquer indivíduo, podendo este usar, alterar e compartilhar livremente o

material informacional respeitando-se os princípios que asseguram a proveniência do conteúdo (OPEN KNOWLEDGE FOUNDATION, 2021). Já no escopo governamental, os dados do governo serão considerados abertos se forem divulgados publicamente de acordo com os oito princípios consagrados pelo *Open Government Working Group* (2007) em Sebastopol (EUA):

- i) completo: todos os dados públicos são disponibilizados. Dados públicos são dados que não estão sujeitos a limitações válidas de privacidade, segurança ou privilégio;
- ii) primário: os dados são apresentados como coletados na fonte, com o nível mais alto possível de granularidade, não em formas agregadas ou modificadas;
- iii) oportuno: os dados são disponibilizados o mais rápido possível para preservar o seu valor;
- iv) acessível: os dados estão disponíveis para a maior variedade de usuários para o maior número de finalidades;
- v) processável por máquina: os dados são razoavelmente estruturados para permitir o processamento automatizado;
- vi) não discriminatório: os dados estão disponíveis para qualquer pessoa, sem necessidade de registro;
- vii) não proprietário: os dados estão disponíveis em um formato sobre o qual nenhuma entidade tenha o controle exclusivo;
- viii) sem licença: os dados não estão sujeitos a nenhum direito autoral, patente, marca comercial ou regulamento de segredo comercial. Restrições razoáveis de privacidade, segurança e privilégios podem ser permitidos.

Embora esses princípios não sejam imutáveis nem juridicamente vinculativos, eles são amplamente considerados pela comunidade global como diretrizes gerais sobre abertura de dados por parte do governo (BAUER; KALTENBÖCK, 2011).

Em 2010, Berners-Lee propôs um sistema de classificação por estrelas para incentivar os governos a publicarem bons dados abertos na Web. As estrelas variam desde um estágio mais simples - como a mera publicação de dados em qualquer formato com licença aberta - até um estágio mais avançado - que seria conectar os dados de acordo com as melhores práticas do LOD (OPEN DATA, 2010):

- ★ Disponível na Web (qualquer que seja o formato), mas com uma licença aberta para ser Dado Aberto.
- ★★ Disponível como dados estruturados legíveis por máquina. Por exemplo: Excel em vez de digitalizar imagens de uma tabela.
- ★★★ Opções anteriores acrescido de formato não proprietário. Por exemplo: CSV em vez de Excel.
- ★★★★ Além disso, usar os padrões abertos do W3C (RDF e SPARQL) para identificar as coisas, para que as pessoas possam apontar para elas.
- ★★★★★ Todas as opções anteriores, acrescido de: conectar dados de fontes heterogêneas objetivado prover contexto (LOD).

Ligar informações de fontes heterogêneas contribuem para atingir os objetivos do governo aberto de forma plena: se os dados puderem ser colocados em um novo contexto, aplicativos com maior valor poderão ser gerados (W3C, 2013). A capacidade de confrontar dados de diferentes contextos é uma qualidade que permite, por exemplo, traçar um recorte ou mesmo uma abordagem única sobre um cenário de que se tem pouco conhecimento. Portanto, cabe aos editores do governo estarem cada vez mais atentos às capacidades do LOD.

3.2 MODELO E GOVERNANÇA DE DADOS

Ao aumentarem sua presença on-line, os governos viram-se pressionados a oferecerem serviços com mais qualidade e inovação. Inserido no contexto do governo eletrônico, a abertura de dados governamentais vem exigindo atenção, especialmente por enfrentar desafios heterogêneos relacionados às questões tecnológicas, econômicas, sociais, políticas, além de incômodas barreiras na cultura organizacional, dentre outras. Essas questões vêm afetando a forma e a qualidade da oferta de dados ao grande público.

Tal situação está impondo aos governos uma necessidade de incorporação de estratégias gerenciais, demandando da administração pública um investimento em modelos de governança para melhor articulação dos processos internos. O objetivo é atingir uma maior qualidade na abertura de dados governamentais em auxílio as comunidades que necessitam dos dados, o que acaba por beneficiar a qualidade e

eficiência dos portais de dados abertos governamentais, sendo estes os principais expoentes das iniciativas de abertura.

Para Horvath (2017), a governança ergue-se como uma coleção de estruturas, regulamentos, normas, padrões, processos e práticas, sejam formais ou informais, que objetivam regulamentar, coordenar e orientar a gestão para que um objetivo específico seja atingido.

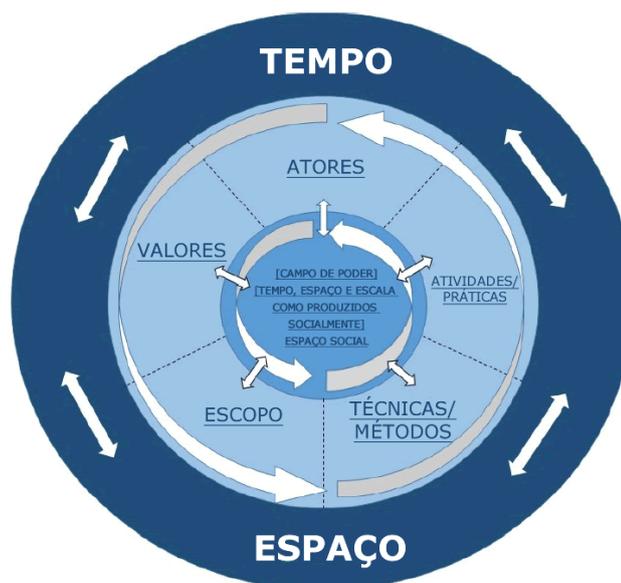
Após realizar uma extensa pesquisa sobre governança, Horvath (2017) analisou os resultados e sugeriu cinco domínios que devem ser considerados para elaboração de um modelo de governança que se proponha a lograr êxito.

- **atores:** constitui-se de um enorme tecido social, tais como: o estado, governos, burocratas, empresas, mercados, sindicatos, indústria, mídia, organizações globais, sociedade, especialistas, ONGs, comunidades de prática (científica, profissional, empresarial), dentre outros.
- **atividades/práticas:** as práticas ou atividades envolvidas incluem a formulação de políticas, orientação, planejamento, financiamento, provisão, controle, descentralização, representação, cooperação, coordenação, participação, integração etc.
- **técnicas/métodos.** vertente ampla que abrange: estruturas legais a diretrizes e regulamentações; de códigos de conduta de governança a normas e valores culturais; de políticas de tecnologia a auditorias, indicadores, avaliações e classificações; e de governança compartilhada, em rede, delegada e cooperativa.
- **escopo:** orientado para a prática abrangendo a capacidade de se atingir objetivos específicos, alguns exemplos: definir direção; exercer poder e autoridade; fortalecer os sistemas democráticos; instituir sistemas equitativos e sólidos; alcançar e manter a sustentabilidade econômica e a viabilidade financeira de longo prazo; salvaguardar os ativos; alcançar mudanças organizacionais e institucionais; promover justiça, transparência e responsabilidade; melhorar a eficiência e a produtividade; capacitação.
- **valores:** valores e escopo estão relacionados, com o escopo objetivando determinar os meios pelos quais os atores podem atingir os valores determinados. Exemplos de valores: busca pela promoção de

responsabilidade, justiça, transparência e prestação de contas; fortalecimento do sistema democrático de governo com participação cidadã; compartilhamento de compromissos e valores éticos etc.

Objetivando melhor compreensão, Horvath (2017) criou um modelo representado pela Figura 3, em que propõe organizar as cinco dimensões.

Figura 3 - Modelo de governança com as dimensões a serem consideradas.



Fonte: Adaptado de Horvath (2017).

Para Horvath (2017), os domínios não devem ser vistos de forma isolada, estando intrinsecamente relacionados. Já "governança" não deve ser tratada no centro do diagrama, devendo ser um fenômeno que se eleva da confluência de uma ampla gama de processos não expressos, tornando-se dessa forma o objeto a ser analisado e não necessariamente o instrumento de análise. O centro se comporta como um campo relacional onde as variáveis que são produzidas nas interações sociais (espaço, tempo, escala e poder) podem ser localizadas. Horvath (2017, p. 39) defende que o *continuum* tempo-espaço deve ficar em um círculo exterior, "[...] refletindo o fato de que o tempo e o espaço não são apenas produzidos socialmente, mas também podem ser independentes e externos às interações humano e social".

A instauração de um modelo de governança permite que os recursos disponíveis sejam empregados de forma mais eficiente, e em conformidade com as prioridades de uma estratégia estabelecida previamente (ANNE, *et al.* 2017). Ainda segundo Anne *et al.* (2017, p.23) "A governança fornece estrutura para a tomada de

decisão organizacional, esclarecendo quem tem autoridade para tomar diferentes tipos de decisões e quem é consultado no processar”.

Portanto, o conceito de governança emerge com intuito de atingir determinado objetivo a partir de valores previamente compartilhados pelos seus atores. Com relação aos instrumentos, tem-se a liberdade de usar as mais variadas atividades e métodos que estiverem à disposição. Para as organizações públicas que lidam especificamente com a ativação de iniciativas em abertura de dados, a governança de dados agrega um conjunto de esforços de gerenciamento dos processos de toda a cadeia relacionada a esse tipo de empreendimento.

A governança de dados é o processo de definir e aplicar prioridades para gerenciar e usar dados como um ativo estratégico na infraestrutura de dados, sendo para isso necessário o estabelecimento de um órgão de governança para instituir as autoridades, estruturas organizacionais, políticas e procedimentos objetivando supervisionar e coordenar a política de gerenciamento de dados (ESTADOS UNIDOS, 2020a). A governança de dados “[...] ajuda a priorizar as alocações de recursos de dados para responder às principais perguntas da agência e atender às necessidades das partes interessadas” (ESTADOS UNIDOS, 2020a, p. 5).

Não apenas alçando os dados como ativo estratégico fundamental a tomada de decisões realísticas, um órgão de governança de dados também precisa cuidar de toda a infraestrutura relacionada ao gerenciamento de dados. Sua atenção está, portanto, repousada nas seguintes atividades (ESTADOS UNIDOS, 2020a):

- **identificação de dados:** identificar ativos de dados e desenvolver um inventário com metadados apropriados;
- **política de gerenciamento de dados:** desenvolver breves declarações de intenção de gerenciamento, e regras fundamentais para conduzir a criação, aquisição, privacidade, integridade, segurança, qualidade e uso de dados e informações;
- **gerenciamento de problemas de dados:** criar processos para identificar e remover os obstáculos que afetam o uso eficaz dos dados.
- **avaliação de dados:** desenvolver processos para medir a qualidade, utilidade e impacto dos dados;
- **supervisão de dados:** monitorar os ativos de dados da organização e quaisquer ações tomadas para melhorá-los;

- **comunicações de dados:** aproximar as informações entre funcionários e gerentes. Uma comunicação alinhada e transparente é fundamental para aprimorar os processos de gerenciamento de dados.

O órgão de governança de dados precisa definir a visão da utilização dos dados alinhada às necessidades da agência, além disso, cabe ao órgão deliberar sobre os valores da agência que servirão de guia na tomada de decisão e priorização de recursos no gerenciamento de dados (ESTADOS UNIDOS, 2020a).

Nesse sentido, Horvath (2017) apregoa que a governança deve se valer de instrumentos como indicadores e avaliações para se atingir os objetivos da gestão. Anne *et al.* (2017) sugerem como método de avaliação o uso de diagnóstico contendo graus de capacidade para avaliar a maturidade do elemento a ser investigado. No âmbito da governança de dados, tais avaliações de maturidade devem ser conduzidas de forma periódica para medir o progresso e priorizar os próximos passos (ESTADOS UNIDOS, 2020a).

Uma avaliação de maturidade analisa todos os aspectos das políticas, procedimentos e operações da agência relacionados a dados e infraestrutura de dados, incluindo governança de dados, gerenciamento de dados, cultura de dados, sistemas e ferramentas de dados, análise de dados, habilidades e capacidade da equipe, capacidade de recursos e conformidade com lei e política (ESTADOS UNIDOS, 2020a, p. 9).

Portanto, deve-se oferecer atenção a dimensão de maturidade de dados, pois seu parecer permite que o órgão público faça uma autoavaliação em relação às melhores práticas, determine lacunas, informe as decisões de investimento e identifique áreas a serem priorizadas (ESTADOS UNIDOS, 2020a).

Um dos desafios a implantação de órgãos de governança de dados na administração pública está relacionado as nuances institucionais. A governança e o gerenciamento de dados eficazes geralmente exigem uma mudança na cultura organizacional, o que é difícil, pois cada agência tem uma maturidade e níveis de evolução distintos. Sugere-se, portanto, que as agências incorporem modelos de governança que façam mais sentido a sua realidade (HORVATH, 2017).

De todo modo, o sucesso em se atingir os objetivos estabelecidos com a abertura de dados governamentais passa por uma abordagem em governança de dados na administração pública conduzida de forma eficiente e madura. Aspectos como planejamento, monitoramento, coordenação, definição de prioridades, devem

ser levados em consideração pelos representantes da gestão pública nos esforços de mitigar as barreiras que obstruem a abertura de dados com melhor qualidade.

No rol de métodos que conduzem uma melhor governança de dados destacam-se os modelos de maturidade, utilizados como importantes instrumentos de diagnóstico, especialmente para os gestores que lidam diretamente com a competência de dados. Tais modelos contribuem não apenas com a análise e o diagnóstico de determinado escopo, mas também servem para auxiliar o processo de governança de dados como um todo, apontando lacunas, indicando prioridades de investimento e alertando sobre eventuais carências que porventura sejam de difícil identificação por parte das autoridades envolvidas na gestão.

3.3 CONTRIBUIÇÃO DA FUNDAMENTAÇÃO TEÓRICA PARA A PESQUISA

O arcabouço conceitual, teórico-metodológico apresentado no Capítulo 3 torna-se importante para o desenvolvimento da pesquisa, especialmente por ser responsável pela elucidação de conceitos seminais que tangem a organização e representação dos conteúdos informacionais digitais e seus padrões semânticos, o estudo de modelos conceituais para organização do conhecimento, além dos conceitos gerais de governança e de governança de dados.

Como um dos objetivos da pesquisa é a investigação de iniciativas relevantes de abertura de dados governamentais, faz-se necessário ter a compreensão e domínio dos fundamentos teóricos que irão nortear a condução da exploração desses objetos empíricos.

A importância no domínio das competências apresentadas na fundamentação inclui a habilidade em saber exatamente o que pesquisar e identificar nos ambientes a serem investigados para que sejam usados posteriormente como inferência na análise dos dados e elaboração de resultados mais conclusivos.

Os portais de dados abertos governamentais publicam um conteúdo informacional extremamente variado, demandando um domínio sobre os padrões de representação de dados. Portanto, um dos interesses do presente estudo, e da pesquisa em CI, é compreender o funcionamento das formas de representação disponíveis, seja na descritiva, temática ou nos modelos conceituais.

O estudo dos modelos conceituais ontológicos orientará a análise dos domínios governamentais investigados, especialmente na forma como os editores

governamentais estão resolvendo as questões de interoperabilidade de dados entre os diversos sistemas de informação que utilizam o portal para compartilhar seus dados e metadados. Já o conhecimento oriundo do *Linked Data* e do LOD será importante para oferecer a melhor curadoria na elaboração das ações de publicação mais alinhadas com os preceitos abertos.

Finalmente, a abordagem em governança de dados auxiliará nas questões enfrentadas pelas agências públicas responsáveis pela disponibilização dos dados, especialmente os desafios políticos, econômicos, sociais, de capacitação, e da organização interna da administração pública.

O domínio dos assuntos delineados no presente Capítulo, especialmente os aspectos técnicos relacionados dos padrões semânticos, modelagem conceitual ontológica e governança de dados, irá auxiliar na perseguição das respostas frente às questões da pesquisa, buscando investigar a abertura de dados na rede pelos governos e concomitantemente propor melhorias na gestão pública de dados.

Os assuntos tratados no capítulo visam não apenas melhorar a qualidade dos dados publicados, como também servirá para nortear a construção do modelo de governança de dados a partir dos desafios enfrentados, e a proposição de atividades alinhadas com a governança de dados para o melhor gerenciamento da publicação e do monitoramento dos portais.

Espera-se, dessa forma, garantir melhores condições para que os dados publicados tenham maior qualidade, e que favoreçam a vida das pessoas em termos de acesso à informação para uso em sociedade.

4 METODOLOGIA DE PESQUISA

Para cumprir os objetivos propostos, a presente pesquisa fez uso de métodos científicos oriundos das Ciências Sociais Aplicadas, que, de acordo com Valentim (2005), compreendem um conjunto de técnicas e instrumentos utilizados para o desenvolvimento de um determinado estudo, visando subsidiar e apoiar o pesquisador nas atividades inerentes à realização da pesquisa.

Gil (2019) destaca que pesquisas se referem aos mais diversos objetos e perseguem objetivos distintos, sendo natural o esforço por classificá-las e torná-las uma atividade importante na organização dos fatos e interpretações salutaras. Assim, a Seção 4.1 trata dos fundamentos da pesquisa; e a Seção 4.2 discorre sobre os procedimentos de coleta e análise dos dados.

4.1 CLASSIFICAÇÃO DA PESQUISA

Em relação a sua finalidade, é considerada uma pesquisa aplicada, pois é voltada à aquisição de conhecimentos com vistas à aplicação de instrumento diagnóstico em situações reais e específicas envolvendo portais de dados abertos governamentais.

Com base na abordagem do problema, esta pesquisa pode ser classificada como sendo de natureza quali-quantitativa em vista à necessidade de analisar, avaliar dados a partir de critérios numéricos, interpretar e descrever um cenário em sua realidade corrente.

Com relação aos seus objetivos, a pesquisa pode ser classificada em duas tipologias, a saber:

- exploratória/descritiva: tem como objetivo proporcionar maior familiaridade com o problema, com vistas a torná-lo mais explícito, e descrever as características do fenômeno investigado. Portanto, será averiguado como os editores estão disponibilizando o conteúdo informacional relativo aos dados abertos nos portais governamentais. Além disso, as tecnologias, os padrões de anotação (vocabulários, modelos, ontologias, metadados), a navegação, o acesso aos dados, as políticas de dados, a governança, a documentação técnica, e a interação dos portais serão analisados (comparativamente) e descritos para que

suas características e usos sejam compreendidos, analisados e interpretados.

- explicativa: tem como propósito identificar e explicar fatores que determinam ou contribuem para a ocorrência do fenômeno estudado. A explicação de certos eventos relacionados aos portais de dados abertos governamentais torna-se essencial para entendimento e clareza dos dados coletados; além de ser importante para qualificar o modelo de governança de dados proposto.

Finalmente, em relação aos procedimentos técnicos para a coleta e a análise dos dados, a pesquisa foi assim constituída:

- bibliográfica: elaborada a partir de material já publicado, constituído principalmente de livros, artigos de periódicos, eventos científicos, teses, dissertações e material disponibilizado na Internet.
- documental: elaborada a partir de material interno às organizações envolvidas nos projetos analisados, incluindo bases de dados, *datasets*, manuais, relatórios de pesquisas, legislação, dentre outros.

4.2 ETAPAS METODOLÓGICAS PARA A PROPOSIÇÃO DOS MODELOS DE DIAGNÓSTICO E DE GOVERNANÇA

Para fins de proposição do modelo de diagnóstico de maturidade dos portais governamentais que, por sua vez, se desdobrará em insumos para a elaboração do modelo de governança voltado a qualidade de dados para publicação em rede, foram empregados uma combinação de estratégias metodológicas que são descritas detalhadamente nas seções que se seguem.

4.2.1 Levantamento bibliográfico e documental

Com o propósito de organizar e melhorar a compreensão a respeito de todo o levantamento de conteúdo pertinente à pesquisa, foram elaborados os Quadro 9 e 10 que descrevem, de forma categorizada e sistemática, todo o conjunto de práticas e estratégias empregadas para durante a realização dos levantamentos bibliográfico-e documental instrumentalizado durante a pesquisa.

Quadro 9 - Descrição das etapas de levantamento bibliográfico

Pesquisa bibliográfica		
Fontes de informação	Bases de dados	Foram consultadas bases de dados específicas que retratam tanto material em CI quanto em CC no Portal de Periódico da Capes, a saber: SciELO – <i>Scientific Electronic Library Online</i> ; <i>Google Scholar</i> ; <i>Library and Information Science Abstracts (LISA)</i> ; Base de Dados em Ciência da Informação (BRAPCI); <i>Web of Science</i> ; <i>Computer and Information Systems</i> ; <i>Computers and Applied Sciences Complete</i> ; <i>Information Science & Technology Abstracts (ISTA)</i> ; <i>Library, Information Science & Technology Abstracts with Full Text</i> ; <i>Scopus (Elsevier B.V)</i> ; e Biblioteca Brasileira de Teses e Dissertações (BDTD).
	Periódicos nacionais	Investigação em periódicos nacionais em CI com Qualis “A” a exemplo da <i>Perspectivas em Ciência da Informação e Informação & Sociedade</i> .
	Periódicos estrangeiros	Investigação em Revistas internacionais bem qualificadas como <i>Annual Review of Information Science and Technology</i> ; <i>IEEE Intelligent Systems</i> ; e <i>International Journal on Semantic Web and Information Systems</i> .
	Evento científico	Foram também consultados os Anais do Encontro Nacional de Pesquisa e Pós-graduação em Ciência da Informação (Enancib), um dos principais eventos científicos nesse campo no país, com um recorte dos últimos dez anos.
Palavras-chave usadas na busca	Para a recuperação dos documentos nas bases de dados científicas supracitadas, foi empregada a técnica de busca por palavras-chave que refletem o universo do assunto. Foram adotados termos nos idiomas português e inglês, incluindo a combinação de dois ou mais destes termos, de forma a tornar os resultados mais refinados.	
	Português	“barreiras”, “dados abertos governamentais”, “dados abertos conectados”, “desafios”, “estado da arte”, “modelos semânticos”, “ontologia”.
	Inglês	“ <i>barriers</i> ”, “ <i>challenges</i> ”, “ <i>linked open data</i> ”, “ <i>LOD</i> ”, “ <i>enrichment semantic</i> ”, “ <i>open government data</i> ”, “ <i>OGD</i> ”, “ <i>ontology</i> ”, “ <i>semantic expressiveness</i> ”, “ <i>state of the art</i> ”.

Fonte: Elaborado pelo autor (2021).

Quadro 10 - Descrição das etapas de levantamento documental

Pesquisa documental			
Fontes de informação	Documentos institucionais	A partir de conhecimento prévio do pesquisador e analisando manuais de dados abertos do governo brasileiro foi possível se chegar a repositórios digitais contendo documentos, relatórios e estatísticas de diversos organismos relacionados diretamente com a pesquisa ou subjacente, como por exemplo: <i>World Wide Web Consortium (W3C)</i> , organização responsável pela criação dos padrões para a Internet; documentos de organizações governamentais (<i>European Union, Organisation for Economic Co-operation and Development – OECD</i>); documentos de organizações não governamentais (<i>International Aid Transparency Initiative, Open Data Charter, Open Government Data Group, Open Government Partnership; Open Knowledge Foundation, World Wide Web Foundation</i>).	
	Documentos governamentais	EUA	<i>Memorandum on Transparency and Open Government; H.R.1770 - OPEN Government Data Act; Federal Data Strategy; 2020 Action Plan.</i>
		Reino Unido	<i>U.K National Data Strategy 2020.</i>
		Brasil	<i>Lei Federal Nº 12.527/2011 (Lei de Acesso à Informação Pública - LAI); Decreto Federal Nº 8.777; Plano de ação 2021-2022; Plano de Dados Abertos, vigência junho/2021 a junho/2023.</i>
		Irlanda	<i>Open Data Strategy 2017-2022; Public Service Data Strategy 2019-2023.</i>

Fonte: Elaborado pelo autor (2021).

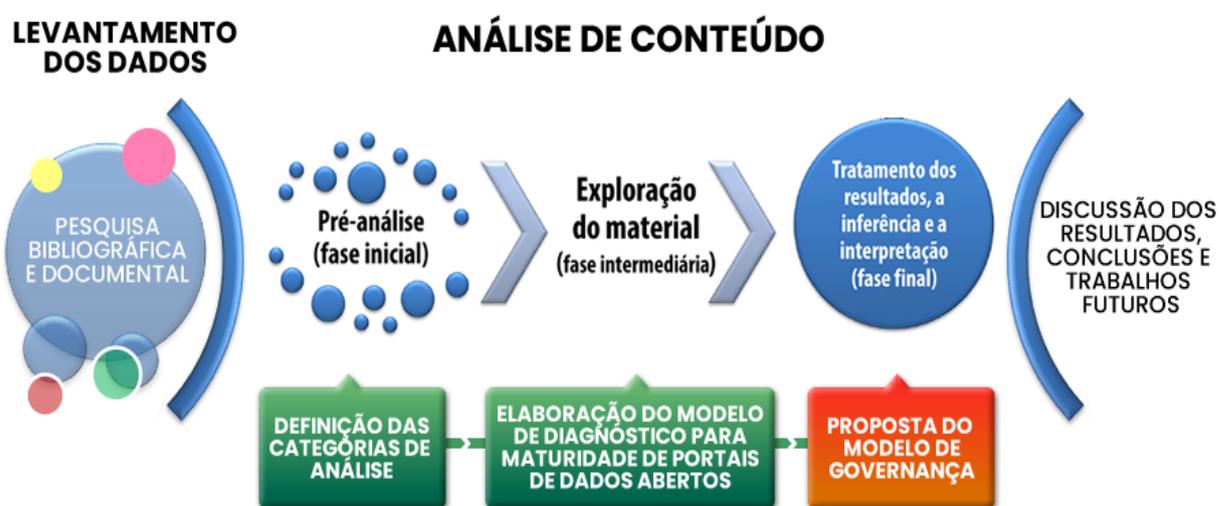
4.2.2 Coleta e análise dos dados

O material coletado nas etapas de levantamento bibliográfico e documental foi examinado a partir de técnicas oriundas do método de análise de conteúdo proposto por Bardin (2016). Para Valentim (2005), tal técnica possui uma relação estreita com diversas ciências, fato que contribui para a riqueza do método. Bardin (2016) aborda que a sutileza dos métodos de análise de conteúdo tem como objetivo superar a incerteza e promover o enriquecimento da leitura.

A análise de conteúdo adota como estratégia o emprego de um conjunto de métodos de análise das comunicações para fins de extração, análise e interpretação dos dados, e assim criar categorias que permitam a inferência de conhecimentos presentes nas mensagens (BARDIN, 2016).

O método combina técnicas com abordagens quantitativa e qualitativa visando obter maior conhecimento sobre o fenômeno a ser pesquisado, permitindo inferências mais seguras no momento da análise dos dados e com resultados mais concisos (VALENTIM, 2005). O processo metodológico delineado na pesquisa, qual seja, desde o levantamento bibliográfico e documental, passando pela coleta e análise dos dados, definição das categorias de análise, desenvolvimento do diagnóstico de maturidade, e chegando a produção do modelo de governança, está condensado no esquema gráfico representado pela Figura 4 juntamente às etapas previstas no método de Bardin (2016).

Figura 4 - Processo metodológico delineado na pesquisa



Fonte: Elaborado pelo autor (2021).

Com relação ao método de análise de conteúdo, o Quadro 11 detalha as técnicas empregadas em cada uma de suas etapas e as respectivas entregas esperadas para a pesquisa.

Quadro 11 - Técnicas e entregas de cada etapa da análise de conteúdo

Etapas	Técnicas para a coleta e análise dos dados		Entregas principais
	Técnica	Descrição	
Pré-análise (fase inicial)	<i>Identificação</i>	Aplicação da técnica de leitura flutuante (primeiro contato com a literatura).	Definição das categorias de análise
	<i>Seleção dos documentos</i>	Constituição do <i>corpus</i> com base nas regras de exaustividade, representatividade, homogeneidade, pertinência e exclusividade. Trata-se do material coletado tanto na pesquisa bibliográfica e documental, quanto nas fontes documentais dos portais elencados para o pré-teste .	
	<i>Coleta e análise dos dados</i>	Condução da coleta e análise dos dados e de um pré-teste para auxiliar na definição das categorias de análise.	
Exploração do material (fase intermediária)	<i>Análise</i>	Aplicação do diagnóstico de maturidade nos portais de dados abertos avaliados no pré-teste, de modo a obter um exame detalhado sobre os dados coletados e registrados.	Modelo de diagnóstico de maturidade endereçado aos portais de dados abertos governamentais
	<i>Interpretação</i>	Extração de indicadores e variáveis de requisitos a partir dos resultados aferidos que, por sua vez, permitirão inferências de conhecimentos presentes nos dados.	
	<i>Síntese de conteúdo</i>	Resultado do esforço de síntese da comunicação, destacando seus aspectos mais importantes a serem trabalhados na fase seguinte.	
Tratamento dos resultados, a inferência e a interpretação (fase final)	A partir do arcabouço teórico-metodológico alcançado com o diagnóstico, o modelo de governança é proposto.		Modelo de governança voltado a qualidade de dados para publicação em rede

Fonte: Elaborado pelo autor (2021).

A seguir, as etapas detalhadas no Quadro 11 serão descritas para fins de entendimento dos procedimentos metodológicos empregados na pesquisa: i) pré-análise (fase inicial); ii) exploração do material (fase intermediária); e iii) tratamento dos resultados, a inferência e a interpretação (fase final).

4.2.2.1 Etapa de pré-análise: definição das categorias de análise

A revisão bibliográfica e documental trouxe à tona publicações distribuídas, incluindo artigos científicos, periódicos, legislação, documentos de organizações governamentais e não-governamentais, dentre outras.

Dada a natureza interdisciplinar da pesquisa, delimitou-se o recorte temático nos campos da CI e da CC por possuírem um aporte teórico metodológico consolidado e relevante às práticas de abertura de dados, especialmente na área de organização e representação da informação e do conhecimento em ambientes digitais. O campo da administração pública também foi investigado como complemento aos da informação e da tecnologia no aspecto de governança e política de dados.

Cerca de 100 publicações estavam em conformidade com os critérios definidos na etapa de levantamento bibliográfico e documental. Desse total, mais de 30 publicações foram selecionadas em função de análise prévia do resumo e palavras-chave, podendo ser consultadas a partir do Apêndice A. As publicações descartadas encontravam-se ou repetidas ou com temáticas não relacionadas de forma explícita com a proposta de levantar as iniciativas e desafios do movimento de dados abertos governamentais.

Para melhor compreensão da realidade da governança de dados em iniciativas de abertura de dados governamentais foi realizado um pré-teste em portais de dados abertos governamentais pré-selecionados. Esse primeiro contato visou consolidar o amadurecimento e posterior padronização das categorias de análise a serem trabalhadas nas etapas subsequentes.

O pré-teste foi conduzido a partir do resultado da análise da etapa de revisão de literatura e dos desafios enfrentados pelos governos na abertura de dados. Dessa forma, a escolha dos portais a serem investigados foram assim justificados:

- EUA (ESTADOS UNIDOS, 2021a) e Reino Unido (REINO UNIDO, 2021a), por serem os pioneiros no movimento aberto, sendo precursores na criação dos portais de dados abertos e os primeiros a aderirem ao LOD. Finalmente, ambas as iniciativas despertam interesse da investigação por estarem em queda tanto nos *rankings* de maturidade, quanto nos resultados elucidados nas pesquisas científicas;

- Brasil (BRASIL, 2021b), por se tratar de um dos países fundadores da OGP, além de despertar o interesse inicial da pesquisa que ansiava por entender o nível de maturidade da iniciativa brasileira;
- Irlanda (IRLANDA, 2021a), por se tratar de um país que está em grande ascensão no movimento de abertura de dados governamentais apresentando grandes inovações através de seu portal.

Os critérios secundários para a seleção dos portais foram assim definidos: i) possuírem um portal federal centralizado; ii) estarem como membros ativos do OGP⁷, sendo esta organização alinhada aos princípios consagrados pela reunião de Sebastopol. Para a realização do pré-teste foi definido como recorte temporal para a investigação dos portais todo o período de maio de 2021.

Após a execução do pré-teste, cujos resultados são exibidos em detalhe nos Apêndices B, C, D, E, constatou-se que as iniciativas ainda estão aquém de um panorama razoável de governança de dados para uma decisão de abertura que seja favorável em termos de valor informacional e transparência à sociedade, alinhando ou confirmando, portanto, com os desafios evidenciados e discutidos no capítulo sobre o estado da arte. Nota-se uma carência de visibilidade em relação às políticas públicas de dados em vigor, sua abrangência, envolvimento social, ações a serem desempenhadas e a transparência na execução das atividades.

A partir da exploração da documentação técnica dos portais feita no pré-teste, observou-se que as iniciativas se esforçam em divulgar os padrões de metadados que adotam para a representação semântica dos dados, mas sem indicação da localização dos modelos de anotação usados e, tampouco, se a publicação de dados segue, por exemplo, o padrão LOD, se restringindo apenas a menção da importância do *ranking* 5 estrelas.

No aspecto tecnológico, todos os portais analisados utilizam o CKAN como plataforma de modelagem de metadados, o que indica a grande predileção das iniciativas internacionais por esse *software* gratuito. No entanto, a forma como cada portal explorou os recursos oferecidos pelos CKAN variou significativamente de uma iniciativa para outra.

Outras questões relacionadas a treinamento, a capacitação, a legislação e a própria disseminação da cultura de dados precisam de maior atenção e

⁷ <https://www.opengovpartnership.org/our-members/>

amadurecimento. Muitos eventos para a promoção da temática de dados abertos ou a divulgação de cursos e eventos estão desatualizados ou são inexistentes.

Com os desafios identificados e organizados a partir da coleta e análise dos dados foi possível lançar um olhar crítico sobre determinados portais de dados abertos, consolidando, portanto, a etapa de pré-teste.

Em se tratando de um instrumento de análise, o pré-teste serviu para consolidação dos desafios como dimensões analíticas, logo, categorias de análise à luz do método de Bardin (2016). Tais dimensões foram consideradas apropriadas como indicadores seguros a serem utilizados na elaboração de um modelo de diagnóstico de maturidade para portais de dados abertos governamentais. Nesse sentido, os desafios foram compilados e organizados com alguns ajustes de nomenclatura e apresentados como indicadores conforme Figura 5.

Figura 5 - Indicadores do modelo de governança de dados



Fonte: Elaborado pelo autor (2021).

Os 8 indicadores foram dispostos em um arranjo cíclico por possuírem características complementares, ou seja, não estão exercendo forças hierárquicas umas sobre as outras. A seguir será apresentado o escopo de cada um dos indicadores:

- i) **infraestrutura em TI:** adequar o *hardware*, *software* e TIC's da organização para que sustentem os sistemas de informação responsáveis por prover capacidade aos portais de dados abertos;
- ii) **acessibilidade e usabilidade:** garantir que qualquer indivíduo possa acessar e usar os dados, além de oferecer recursos que permitam a navegação pelo portal sem obstrução, com o emprego de uma linguagem acessível em favor de uma experiência plena de uso;
- iii) **curadoria e publicação dos dados:** definir na organização a autoridade que irá zelar e cuidar do repositório de dados, além de garantir que os *datasets* sejam disponibilizados em conformidade com as melhores práticas de qualidade e publicação;
- iv) **político e social:** identificar e disponibilizar datasets de maior impacto social, utilizando o portal de dados abertos como instrumento de captação dos anseios da sociedade, vitrine para divulgação dos impactos dos dados e monitoramento das ações de implementação dos dados abertos pelas agências governamentais;
- v) **econômico:** valorizar as iniciativas que permitam a viabilidade econômica e sustentável do portal de dados abertos, incluindo a promoção da participação de atores da iniciativa pública e privada e, concomitante, apontando os benefícios econômicos da reutilização dos dados para a melhoria do serviço público em geral;
- vi) **organização e processos internos:** criar uma estratégia de gerenciamento, manutenção e uso dos dados dentro do governo, visando torná-los ativos estratégicos para a melhor tomada de decisão, e fomentando uma cultura de dados na administração;
- vii) **aspectos legais e de privacidade:** determinar claramente quais os termos relacionados a privacidade e a segurança estão atribuídos aos dados e metadados, tornando o tipo de licença de uso claramente identificável nos *datasets* disponibilizados no portal;
- viii) **capacitação e treinamento:** fomentar a cultura de dados entre a sociedade, além de aprimorar as habilidades e competências ligadas a dados em alunos da rede de ensino e servidores públicos que trabalham diretamente com processos ligados a curadoria de dados.

4.2.2.2 Etapa de exploração do material: elaboração do modelo de diagnóstico de maturidade

Esta etapa foi conduzida visando a elaboração e aplicação do modelo de diagnóstico de maturidade em portais de dados abertos. O instrumento foi proposto a partir da definição das categorias de análise, determinadas na fase anterior, e foi aplicado em portais de dados abertos governamentais com o intuito de avaliar a sua maturidade frente aos indicadores recomendados.

O Quadro 12, a seguir, apresenta os indicadores e suas variáveis, extraídas a partir da síntese de conteúdo consolidada na fase de pré-análise. As variáveis estão distribuídas em cada um dos indicadores, e acompanhadas de fundamentação diagnóstica.

Quadro 12 - Fundamentação diagnóstica para cada variável do indicador

(continua)

Indicador	Variável	Fundamentação diagnóstica
1. Infraestrutura em TI	1.1. Uso de sistema de gerenciamento de dados	Um sistema de gerenciamento de dados (em inglês, <i>Data Management System</i> - DMS) torna mais eficiente a descrição, a publicação, o compartilhamento, e a gestão do conteúdo nos portais. O uso de plataformas gratuitas que dispensam o uso de <i>softwares</i> proprietários ⁸ propiciam economicidade aos governos, garantia de atualizações futuras e novos recursos através de uma comunidade de prática atuante.
	1.2. Informações nos <i>datasets</i>	As informações sobre as datas de publicação e atualização garantem um melhor controle do versionamento dos dados. <i>Datasets</i> desatualizados podem indicar problemas de ordem técnica ou organizacional. A informação sobre o editor de dados permite maior transparência e atribuição de responsabilidade, melhorando a qualidade de procedência dos <i>datasets</i> .
	1.3. Informação sobre modificação nos <i>datasets</i>	A partir da divulgação do tempo de atraso entre a publicação do <i>dataset</i> em sua fonte original, até seu espelhamento no portal (não devendo ser confundido com frequência de atualização), os usuários poderão acompanhar a celeridade na oferta dos dados, constituindo-se em uma métrica de qualidade para a gestão da infraestrutura em TI. Espera-se que a latência seja nula, com as versões atualizadas dos dados sendo ofertadas no menor intervalo de tempo possível.
	1.4. Disponibilidade dos servidores	Informações sobre disponibilidade dos servidores são importantes para aferir se a infraestrutura está conseguindo prover um serviço ativo e confiável em tempo integral, de forma a permitir o usufruto do serviço tanto por indivíduos quanto por aplicações em qualquer momento.
2. Acessibilidade e Usabilidade	2.1. Navegação do portal	A elaboração de um projeto de <i>design</i> ou de arquitetura da informação conciso e padronizado permite que os usuários sejam conduzidos aos elementos principais das páginas sem obstáculos. O emprego de recursos variados de navegação visa garantir o acesso aos <i>datasets</i> , documentos e demais conteúdos de forma plena e sem barreiras.
	2.2. Busca por <i>datasets</i>	A grande quantidade de <i>datasets</i> disponibilizados nos portais pode se tornar um problema para sua encontrabilidade, portanto é necessário oferecer o maior número de recursos para localização dos <i>datasets</i> que estão catalogados nos sítios governamentais.

⁸ Ao utilizar algum tipo de solução proprietária, o órgão público acaba ficando dependente do fornecedor do *software*.

Quadro 12 - Fundamentação diagnóstica para cada variável do indicador

(continua)

Indicador	Variável	Fundamentação diagnóstica
2. Acessibilidade e Usabilidade	2.3. Apresentação dos dados	A iniciativa deve prover o maior número de recursos possíveis na apresentação dos dados, pois muitos dos indivíduos que acessam os portais não possuem os recursos tecnológicos ou capacidades cognitivas para processar e analisar o conteúdo informacional dos <i>datasets</i> .
	2.4. <i>Download</i> dos <i>datasets</i>	Com o intuito de atender aos diferentes perfis que acessam o portal para reutilizar os dados, recomenda-se que as iniciativas governamentais de abertura garantam recursos variados de <i>download</i> e consumo dos <i>datasets</i> .
	2.5. Reportar erro nos <i>datasets</i>	Em virtude da amplitude dos <i>datasets</i> , faz-se necessária a disponibilização de canais específicos para reportar possíveis erros. Dessa forma, encoraja-se o maior envolvimento cidadão com o serviço, além de criar a necessidade de executar ações de manutenção e monitoramento de forma periódica, pelo mantenedor governamental, visando a qualidade do serviço.
	2.6. Interação com os <i>datasets</i>	Ao incluir recursos de interação entre o público e os <i>datasets</i> , o portal contribuirá para a maior transparência acerca da efetividade de um determinado conjunto de dados, indicando, de forma preliminar, se ele está atendendo realmente aos anseios da sociedade.
3. Curadoria e publicação dos dados	3.1. Abrangência legal da publicação de dados	A regularização do compartilhamento dos mesmos princípios de abertura de dados em todas as esferas governamentais permitirá que sejam elaboradas políticas de dados mais abrangentes e eficientes. A partir disso, os portais de dados abertos serão beneficiados, com aumento da padronização entre as iniciativas, ganho de interoperabilidade técnica entre os <i>datasets</i> , além do provimento de uma capacitação uniforme para os funcionários de todas as esferas.
	3.2. Estratégia para qualidade de publicação	A governança de dados instituída precisa ter a predileção pelas melhores práticas de publicação de dados disponíveis. Agindo dessa forma, garante-se dados com maior representatividade semântica, beneficiando o acesso, a recuperação e o reúso dos dados.
	3.3. Documentação técnica sobre metadados	A documentação técnica se faz necessária para auxiliar os editores de dados, desenvolvedores ou qualquer outro interessado nas especificações, no tratamento e na disponibilização dos dados junto ao portal, permitindo uma melhor exploração do potencial dos dados e sua posterior interoperabilidade com outras bases.

Quadro 12 - Fundamentação diagnóstica para cada variável do indicador

(continua)

Indicador	Variável	Fundamentação diagnóstica
3. Curadoria e publicação dos dados	3.4. Perfis autorizados a publicar	Apesar dos dados serem oriundos das agências e órgãos governamentais, espera-se que qualquer organização independente ou indivíduo possa colaborar publicando dados que sejam de interesse público, tornando o processo plural e inclusivo.
4. Político e Social	4.1. Consulta pública para diagnosticar prioridade de abertura.	Desde o início dos debates sobre abertura de dados do governo, os ativistas pleiteiam consultas públicas para a definição dos dados mais relevantes a serem abertos. A consulta é um mecanismo que visa oferecer transparência sobre as ações desencadeadas pelos governos, objetivando reduzir o poder decisório de uma minoria sobre a maioria frente a um conteúdo de grande valor que são os dados.
	4.2. Reúso de dados com impacto político, social e ambiental	Uma melhor percepção da sociedade sobre o real valor dos dados governamentais é conquistada a partir do compartilhamento de exemplos reais dos impactos políticos, sociais e ambientais advindos de sua abertura. Esse tipo de transparência contribui para que a sociedade se posicione frente ao uso sustentável dos dados como ativo transformador, valorizando a cultura de dados.
	4.3. Quantitativo de órgãos que já abriram dados	Os dados não devem ser controlados, represados, ou ocultos por uma pequena parcela de pessoas em posição privilegiada na administração pública. Portanto, o governo deve monitorar quais órgãos estão publicando regularmente seus dados, e quais ainda estão pendentes, permitindo que qualquer indivíduo fiscalize de forma transparente, e cobre esclarecimentos sobre eventuais retenções de dados injustificadas ou não previstas na legislação.
	4.4. Reúso de dados em prol de comunidades marginalizadas	O uso dos dados é pleiteado por diversos atores sociais, especialmente por aqueles que lutam em defesa das comunidades mais vulneráveis e marginalizadas. Para que não ocorram discriminações, é imprescindível que os governos promovam periodicamente iniciativas de reúso de dados em apoio a essas comunidades.
5. Econômico	5.1. Estratégia de ativação de negócios	A partir da ativação econômica dos dados, pode-se tornar o portal mais sustentável e viável economicamente, de forma a garantir reinvestimentos na infraestrutura em TI e capacitação.

Quadro 12 - Fundamentação diagnóstica para cada variável do indicador

(continua)

Indicador	Variável	Fundamentação diagnóstica
5. Econômico	5.2. Reúso de dados em prol do serviço público	Ao promover editais públicos, o governo fomentará a transformação econômica através da reutilização dos dados. Serviços antes nunca previstos ou planejados pela administração podem ser criados a partir do reúso dos dados, garantindo que o governo inove sem a necessidade de despender vultosas quantias, o que permitiria economicidade aos cofres públicos.
	5.3. Reúso de dados com impactos econômicos	A transparência dos impactos econômicos a partir do reúso de dados abertos contribui para a consolidação dos dados como ativos estratégicos criando uma cultura de dados e justificando mais reinvestimentos para a infraestrutura de dados do portal.
6. Organização e processos internos	6.1. Cultura de dados na administração pública	As barreiras administrativas só começarão a ser desbloqueadas a partir da criação de uma cultura de dados na administração pública. A governança deve estar preparada para estimular a liderança, a gestão, a supervisão e as políticas que contribuirão para a criação desse novo ambiente.
	6.2. Reúso dos dados como ativo estratégico	A moderna administração pública está colocando os dados baseados em evidências no centro das decisões. A governança deve garantir uma estrutura organizacional que contribua para que os dados sejam utilizados como ativos estratégicos dentro das instituições evitando-se seu ocultamento e mal uso por uma elite pública.
	6.3. Reúso de dados pela administração pública	Publicar os benefícios da reutilização dos dados abertos, dentro da administração pública, melhora a percepção do governo junto a sociedade e reforça os princípios de eficiência e transparência. Além disso, a cada nova conquista reconhecida pelo público, tal prática teria o potencial de angariar para a abertura de dados a adesão de novos elementos da gestão pública.
7. Aspectos legais e de privacidade	7.1. Legislação referente a regularização de abertura	Disponibilizar a legislação que regula a abertura de dados reafirma o compromisso do governo com a transparência e uso legalizado dos dados, evitando-se futuros transtornos administrativos e institucionais.

Quadro 12 - Fundamentação diagnóstica para cada variável do indicador

(conclusão)

Indicador	Variável	Fundamentação diagnóstica
7. Aspectos legais e de privacidade	7.2. Licença de uso dos dados	Como os dados abertos governamentais tratam de informações sensíveis à sociedade, toda a licença disponibilizada nos <i>datasets</i> deve estar acompanhada do detalhamento acerca da sua abrangência de uso.
	7.3. Recurso de denunciar <i>datasets</i>	Problemas podem ocorrer na disponibilização dos dados no portal, dentre eles está a publicação de dados que não deveriam ser abertos por ferirem princípios de legalidade ou privacidade previstos em lei. Para essas situações, deve-se disponibilizar um canal especializado para tratar do tema.
8. Capacitação e treinamento	8.1. Promoção da temática de dados abertos	A promoção de eventos de dados abertos serve para melhorar a capacitação dos servidores em competência de dados, além de promover a cultura de dados na administração pública e na sociedade.
	8.2. Habilidades e competências de dados nos currículos educacionais	Os currículos escolares devem ser modernizados adotando competências de dados com intuito de preparar os indivíduos, de forma mais adequada, para a nova cultura de dados. As ações de governança de dados devem prever em sua política esse tipo de abordagem educacional.
	8.3. Capacitação dos funcionários do governo	A capacitação dos servidores deve ser constante, tendo o portal de dados um papel central nessa questão. A administração deve disponibilizar de forma aberta conteúdos que auxiliem tanto os seus funcionários, quanto servidores de outras esferas, além de entusiastas.

Fonte: Elaborado pelo autor (2021).

No percurso para a elaboração de um modelo de governança, faz-se necessário dentro de um instrumento de diagnóstico a formalização dos graus de capacidade explorados, tornando dessa forma mais precisa a avaliação de maturidade da organização a ser investigada (ANNE *et al.*, 2017).

Nesse sentido, o Quadro 13 fundamenta os níveis de graduação que compõem a avaliação, que variam entre 0 (zero) e 3 (três), sendo 0, determinante da escassez do elemento avaliado, e 3, determinante do grau máximo de maturidade.

Quadro 13 - Fundamentação dos níveis de maturidade

Nível de maturidade	Fundamentação	Pontuação
Nível 0	É o estágio mais primitivo da avaliação. Existe uma inércia relacionada a aspectos técnicos, de atendimento a boas práticas, sustentabilidade, governança e principalmente transparência nas ações desenvolvidas através do portal.	0
Nível 1	É o estágio inicial de conformidade com o elemento avaliado. É identificável uma preocupação emergente com tecnologias, acesso e uso dos dados, boas práticas de publicação, mas ainda carecendo de uma melhor abordagem frente a governança de dados, sustentabilidade, transparência e envolvimento com os setores sociais.	1
Nível 2	É o estágio estabelecido da avaliação. Percebe-se uma preocupação formal com as boas práticas de infraestrutura e publicação, uma maior preocupação com transparência no portal e sustentabilidade, o início de um maior envolvimento com outros atores sociais, faltando amadurecer e transparecer as diretrizes e sistematizações demandadas por uma governança de dados ideal.	2
Nível 3	É o estágio de alta capacidade da avaliação. Nesse nível, o governo consegue prover ao portal de dados abertos as melhores tecnologias, as práticas de publicação mais modernas, uma governança de dados moderna e transparente que atende à gestão interna e capacita seus funcionários, além de um consolidado envolvimento com diversos setores que possuam grande representatividade na sociedade.	3

Fonte: Elaborado pelo autor (2021).

O modelo de diagnóstico de maturidade proposto é composto por 8 indicadores, perfazendo 30 variáveis de requisitos ao todo. O Quadro 14 apresenta a composição e o total das variáveis, a pontuação máxima de cada indicador e a pontuação final estabelecida para o diagnóstico.

Quadro 14 - Pontuação dos indicadores e valor total do diagnóstico

Indicador	Número de variáveis	Pontuação máxima do indicador
Infraestrutura em TI	4	12
Acessibilidade e Usabilidade	6	18
Curadoria e Publicação de Dados	4	12
Político e Social	4	12
Econômico	3	9
Organização e Processos Internos	3	9
Aspectos Legais e de Privacidade	3	9
Capacitação e Treinamento	3	9
Diagnóstico de maturidade para portais de dados abertos	Total de variáveis	Pontuação total do diagnóstico
	30	90

Fonte: Elaborado pelo autor (2021).

Portanto, a pontuação máxima que uma iniciativa de dados abertos governamentais avaliada através do modelo de diagnóstico de maturidade pode alcançar é de 90 pontos.

Com os indicadores e as variáveis de requisitos bem fundamentados, condizentes com a exploração semântica do conteúdo coletado, e com os níveis de maturidade estabelecidos, foi possível propor o modelo de diagnóstico de maturidade para portais de dados abertos governamentais, apresentado no Quadro 15.

O instrumento foi responsável por nortear os parâmetros adotados durante a coleta de dados a partir da avaliação dos portais de dados abertos, e atribuir as pontuações determinadas.

Quadro 15 - Modelo de diagnóstico de maturidade para portais de dados abertos governamentais

(continua)

Indicador	Variável	Níveis de maturidade			
		Nível 0	Nível 1	Nível 2	Nível 3
1. Infraestrutura em TI	1.1. O portal de dados abertos utiliza algum DMS para viabilizar a descrição, publicação, compartilhamento, reutilização dos dados e gerenciamento de conteúdo?	O portal utiliza páginas HTML estáticas, sem qualquer uso de um DMS.	O portal utiliza um DMS do tipo proprietário, gerando custo à instituição.	O portal utiliza um DMS desenvolvido pelo próprio governo, que apesar dos custos inferiores pode ocasionar problemas de manutenção e interoperabilidade.	O portal utiliza um DMS de código aberto, moderno e gratuito. Consegue-se a redução de custos, tratamento de interoperabilidade e acesso a inovações desenvolvidas por comunidades de práticas ativas.
	1.2. O portal de dados abertos informa a data original de publicação, a data da última atualização, e o editor que publicou os dados?	Nenhuma informação é divulgada.	É disponibilizado ao menos uma informação.	É disponibilizado ao menos duas informações.	É informada a data original de publicação, a data da última atualização, além da identificação do editor responsável pela publicação do <i>dataset</i> .
	1.3. O portal de dados abertos informa o tempo de atraso entre a modificação do <i>dataset</i> na fonte original até a atualização no portal?	Nenhuma informação é divulgada.	Nenhuma informação é divulgada, mas é oferecido um link para a fonte original, podendo-se confrontar os dados.	A informação é divulgada, levando-se alguns dias para refletir as modificações no portal.	A informação é divulgada, levando-se poucas horas ou minutos para refletir as modificações no portal.
	1.4. O portal de dados abertos informa o tempo que o servidor ficou disponível no mês?	Nenhuma informação é divulgada.	A informação é divulgada, com o servidor disponível em média abaixo de 95% do tempo no período de um ano.	A informação é divulgada, com o servidor disponível em média entre 95% e 98% do tempo no período de um ano.	A informação é divulgada, com o servidor ficando em sua totalidade no ar, ou disponível em média por 99% do tempo no período de um ano.

Quadro 15 - Modelo de diagnóstico de maturidade para portais de dados abertos governamentais

(continua)

Indicador	Variável	Níveis de maturidade			
		Nível 0	Nível 1	Nível 2	Nível 3
2. Acessibilidade e Usabilidade	2.1. Navegação e apresentação dos dados no portal.	O portal possui uma navegação sem destaque ou estrutura.	O portal possui uma navegação destacada, e uma navegação auxiliar nas páginas internas.	O portal possui uma navegação destacada, e uma navegação auxiliar nas páginas internas, mas com páginas sem padrão gráfico e consistência visual.	O portal possui uma navegação destacada, uma navegação auxiliar nas páginas internas, e conta com páginas seguindo um padrão de arquitetura da informação que reforça a consistência visual na navegação.
	2.2 Busca por <i>datasets</i> .	O portal não oferece recursos para localização dos <i>datasets</i> .	O portal oferece um campo de busca para localização dos <i>datasets</i> .	O portal oferece um campo de busca para localização dos <i>datasets</i> , além de filtros de refinamento.	O portal oferece uma busca para localização dos <i>datasets</i> , filtros de refinamento, além de uma organização temática por <i>datasets</i> .
	2.3. Apresentação dos dados	Os dados são publicados com nenhum elemento descritivo.	Os dados são publicados com poucos elementos descritivos ou usando linguagem complexa.	Os dados são publicados com elementos descritivos, em linguagem acessível, e podendo conter uma pré-visualização.	Os dados são publicados com elementos descritivos em, linguagem acessível, e podendo conter uma pré-visualização acompanhada de gráficos ou análise relevante.
	2.4. Recursos para <i>download</i> dos <i>datasets</i> .	Não há opção de <i>download</i> .	É disponibilizado apenas um <i>link</i> para o <i>download</i> do <i>dataset</i> .	É disponibilizado o <i>link</i> para o <i>download</i> do <i>dataset</i> , e opções de exportação de dados em formatos diversos (HTML, PDF, CSV, JSON etc.).	É disponibilizado o <i>link</i> para o <i>download</i> do <i>dataset</i> , opções adicionais de exportação de dados em formatos diversos, além de acesso aos dados através de uma <i>Application Programming Interface</i> (API).

Quadro 15 - Modelo de diagnóstico de maturidade para portais de dados abertos governamentais

(continua)

Indicador	Variável	Níveis de maturidade			
		Nível 0	Nível 1	Nível 2	Nível 3
2. Acessibilidade e Usabilidade	2.5. O portal de dados abertos possui recursos que possibilitam os usuários reportarem erros nos <i>datasets</i> ?	Nenhum recurso é oferecido.	É oferecido o <i>e-mail</i> do editor de dados, ou um formulário genérico de contato.	É oferecido um <i>e-mail</i> , ou um formulário que trate especificamente de situações de erro.	É oferecido um recurso para que se abra um chamado e reporte o erro, possibilitando acompanhar a solicitação até a sua resolução.
	2.6. O portal de dados abertos possui recursos de interação que permitem classificar, comentar e avaliar por meio de pesquisa os <i>datasets</i> ?	Nenhum recurso de interação é oferecido.	É oferecido ao menos um recurso de interação.	São oferecidos ao menos dois recursos interativos.	O conteúdo disponibilizado no portal pode ser classificado, comentado e avaliado por meio de pesquisa, sendo esse <i>feedback</i> considerado em melhorias futuras para o portal.
3. Curadoria e Publicação dos Dados	3.1. Qual a abrangência da publicação de dados definidos pela legislação ou pela política de dados abertos?	Nenhuma informação é mencionada.	A abrangência é restrita a nível nacional.	A abrangência é restrita a nível nacional e regional.	A abrangência é em nível nacional, regional e local.
	3.2. A política de dados abertos prevê uma estratégia para a qualidade da publicação de dados na rede?	Nenhuma estratégia foi identificada.	Considera abrir os dados em formato aberto não proprietário e legíveis por máquina.	Considera abrir os dados em formato aberto, legíveis por máquina e com identificadores comuns (URIs).	Considera abrir os dados no nível 5 estrelas do LOD, prevendo que sejam ligados a outros dados de fontes externas com semântica bem definida para prover contexto (ex.: uso e reuso de ontologias de domínios específicos para organização e interligação semântica de conteúdo dos <i>datasets</i>).

Quadro 15 - Modelo de diagnóstico de maturidade para portais de dados abertos governamentais

(continua)

Indicador	Variável	Níveis de maturidade			
		Nível 0	Nível 1	Nível 2	Nível 3
3. Curadoria e Publicação dos Dados	3.3. O governo publica em seu portal de dados abertos alguma documentação técnica sobre os modelos de metadados usados?	Nenhuma documentação foi disponibilizada.	A documentação é divulgada, mas não dispõe de um dicionário de dados.	A documentação é divulgada contendo o dicionário de dados.	A documentação é divulgada contendo o dicionário de dados, além de informações sobre as fontes externas com as quais se estabelecem as relações com o esquema de metadados.
	3.4. Quais atores estão autorizados a publicarem dados no portal?	Apenas editores do governo.	Editores do governo e organizações não governamentais.	Editores do governo, organizações não governamentais, e setor privado.	Editores do governo, organizações não governamentais, setor privado, e quaisquer indivíduo.
4. Político e Social	4.1. O portal fornece uma consulta pública visando diagnosticar quais dados governamentais poderiam ter prioridade de abertura?	Nenhuma consulta pública aberta identificada.	É realizada uma consulta pública aberta, mas com restrições à participação.	É realizada uma consulta pública aberta e sem limitação de participação.	É realizada uma consulta pública aberta, sem limitação de participação, e com os resultados sendo comprovadamente implantados no portal obedecendo um cronograma.
	4.2. O governo publica em seu portal de dados abertos exemplos de serviços de reúso de dados que estão trazendo impactos ou benefícios políticos, sociais e ambientais para o país?	Nenhum exemplo é divulgado.	Os exemplos são divulgados, mas acompanhados de informações básicas.	Os exemplos são divulgados, contendo relatórios detalhados para consulta.	Os exemplos são divulgados através de aplicações que reutilizam dados em prol da luta contra a desigualdade, preservação da natureza e combate à corrupção.

Quadro 15 - Modelo de diagnóstico de maturidade para portais de dados abertos governamentais

(continua)

Indicador	Variável	Níveis de maturidade			
		Nível 0	Nível 1	Nível 2	Nível 3
4. Político e Social	4.3. O governo detalha em seu portal de dados abertos o quantitativo de órgãos e agências que já abriram seus dados?	Nenhum detalhamento foi identificado.	Entre 1% e 40% dos órgãos e agências governamentais já abriram seus dados.	Entre 40% e 80% dos órgãos e agências governamentais já abriram seus dados.	Entre 80% e 100% dos órgãos e agências governamentais já abriram seus dados.
	4.4. O governo promove em seu portal de dados abertos iniciativas de reutilização de dados com o intuito de criar aplicações para comunidades marginalizadas?	Nenhuma iniciativa é divulgada.	Concursos públicos e <i>hackathons</i> são promovidos, mas não estão atualizados.	Concursos públicos e <i>hackathons</i> recentes são promovidos, mas com envolvimento apenas do setor público.	Concursos públicos e <i>hackathons</i> recentes são promovidos com envolvimento do setor público e de diversos membros da sociedade.
5. Econômico	5.1. A política de dados abertos prevê alguma estratégia de ativação de negócios através do uso de dados abertos?	Nenhuma estratégia é mencionada na política de dados abertos.	A estratégia é apenas mencionada na política de dados abertos, sem maiores detalhes.	A estratégia é mencionada na política de dados abertos, contendo sua sistematização.	A estratégia é mencionada na política de dados abertos, contendo sua sistematização e diretrizes para sua condução (ações, tempo de execução e responsabilidades).
	5.2. O governo promove em seu portal de dados abertos iniciativas de reutilização de dados com o intuito de criar aplicações inovadoras para o serviço público?	Nenhuma iniciativa é divulgada.	Concursos públicos e <i>hackathons</i> são promovidos, mas não estão atualizados.	Concursos públicos e <i>hackathons</i> recentes são promovidos, mas com envolvimento apenas do setor público.	Concursos públicos e <i>hackathons</i> recentes são promovidos com envolvimento do setor público e de diversos membros da sociedade.

Quadro 15 - Modelo de diagnóstico de maturidade para portais de dados abertos governamentais

(continua)

Indicador	Variável	Níveis de maturidade			
		Nível 0	Nível 1	Nível 2	Nível 3
5. Econômico	5.3. O governo publica em seu portal de dados abertos exemplos de reuso de dados que estão trazendo impactos ou benefícios econômicos para o país?	Nenhum exemplo é divulgado.	Os exemplos são divulgados, mas acompanhados de informações básicas.	Os exemplos são divulgados, contendo relatórios detalhados para consulta.	Os exemplos são divulgados através de aplicações que reutilizam dados em prol do desenvolvimento econômico, geração de emprego e renda e promoção de inovação.
	6.1. A política de dados abertos contempla iniciativas que iniciem uma cultura de abertura interna no âmbito dos órgãos e agências governamentais?	Nenhuma iniciativa é mencionada na política de dados abertos.	A iniciativa é apenas mencionada na política de dados abertos, sem maiores detalhes.	A iniciativa é abordada na política de dados abertos, contendo sua sistematização.	A iniciativa é abordada na política de dados abertos, contendo sua sistematização e diretrizes que visam estimular liderança, gestão, supervisão, e políticas de comunicação interna.
6. Organização e Processos Internos	6.2. A política de dados abertos contempla um plano de gerenciamento, manutenção e uso de dados como ativos estratégicos dentro da administração pública?	Nenhum plano é mencionado na política de dados abertos.	O plano é mencionado na política de dados abertos, sem maiores detalhes.	O plano é mencionado na política de dados abertos, contendo sua sistematização.	O plano é mencionado na política de dados abertos, contendo sua sistematização e diretrizes visando garantir a existência de autoridades, funções, estruturas organizacionais, e recursos para apoiar o gerenciamento, manutenção e uso de dados como ativos estratégicos.

Quadro 15 - Modelo de diagnóstico de maturidade para portais de dados abertos governamentais

(continua)

Indicador	Variável	Níveis de maturidade			
		Nível 0	Nível 1	Nível 2	Nível 3
6. Organização e Processos Internos	6.3. O governo compartilha em seu portal de dados abertos exemplos de reutilização de dados dentro da administração pública que estão beneficiando a ativação de melhores políticas e processos de tomada de decisão?	Nenhum exemplo é divulgado.	Exemplos são divulgados, mas acompanhados de informações básicas.	Exemplos são divulgados, contendo relatórios detalhados para consulta.	Exemplos são divulgados contendo relatórios detalhados para consulta, e acompanhado de infográficos para análise e melhor compreensão dos impactos gerados dentro da administração pública e os benefícios na tomada de decisão.
7. Aspectos Legais e de Privacidade	7.1. A legislação referente a regularização da abertura de dados está presente no portal?	A legislação não é disponibilizada.	Apenas alguns trechos da legislação são mencionados.	A legislação pode ser localizada no portal, mas de difícil acesso ou oculta.	A legislação está disponível no portal, de forma transparente, visível e de fácil acesso.
	7.2. Os dados abertos são publicados com sua licença de uso?	Os <i>dados</i> são publicados sem qualquer tipo de menção acerca de sua licença.	A licença é exibida, mas sem detalhamento da extensão de seu uso.	A licença é exibida, contendo o detalhamento da extensão de seu uso, mas de difícil localização na página.	A licença é exibida, contendo o detalhamento da extensão de seu uso, e localizada de forma clara e transparente para o usuário.
	7.3. O governo oferece em seu portal de dados abertos algum recurso que permita denunciar um <i>dataset</i> que esteja em desacordo com os aspectos legais e de privacidade?	Nenhum recurso é oferecido.	É oferecido o <i>e-mail</i> do editor de dados, ou um formulário genérico de contato.	É oferecido um <i>e-mail</i> , ou um formulário que trate especificamente de denúncia relacionada a aspectos legais e de privacidade.	É oferecido um recurso para que se abra um chamado e reporte a denúncia relacionada a aspectos legais e de privacidade, possibilitando acompanhar a solicitação até a sua resolução.

Quadro 15 - Modelo de diagnóstico de maturidade para portais de dados abertos governamentais

(conclusão)

Indicador	Variável	Níveis de maturidade			
		Nível 0	Nível 1	Nível 2	Nível 3
8. Capacitação e Treinamento	8.1. O governo disponibiliza em seu portal de dados abertos eventos para promoção da temática de abertura de dados?	Nenhum evento é disponibilizado.	Eventos são disponibilizados, mas não estão atualizados.	Eventos recentes são disponibilizados, mas com envolvimento apenas do setor público.	Eventos recentes são disponibilizados com envolvimento do setor público e de diversos membros da sociedade.
	8.2. O governo divulga em seu portal de dados abertos ou na política de dados as ações para desenvolver as habilidades e competências no entendimento do ecossistema de dados nos currículos educacionais?	Não há divulgação de ações.	É apenas divulgado uma manifestação de ação, sem ainda estar em prática.	São divulgadas as ações que já estão sendo desenvolvidas, mas apenas com o envolvimento do setor público.	São divulgadas as ações que já estão sendo desenvolvidas, com o envolvimento do setor público e de organizações que representam a sociedade.
	8.3. O governo divulga em seu portal de dados abertos conteúdo relativo a iniciativas de capacitação em competências de dados abertos a seus funcionários?	Nenhuma iniciativa de capacitação em competência de dados abertos para funcionários do governo é divulgada.	São disponibilizados no portal conteúdos (guias e documentos) para apoiar a iniciativa de capacitação dos funcionários.	São disponibilizados no portal conteúdos (guias e documentos), além de programas de treinamento para apoiar a capacitação dos funcionários.	São disponibilizados no portal conteúdos (guias e documentos), programas de treinamento, além de ferramentas ou <i>softwares</i> de dados para apoiar a iniciativa de capacitação dos funcionários.

Fonte: Elaborado pelo autor (2021).

O modelo de diagnóstico de maturidade em portais de dados abertos foi aplicado entre os dias 11 de julho e 17 de agosto de 2021 nos mesmos portais de dados abertos governamentais submetidos ao pré-teste, quais sejam, Estados Unidos, Reino Unido, Brasil e Irlanda. Conjuntamente com o resultado do diagnóstico está prevista a apresentação, por meio de cálculo de média, do nível de maturidade alcançado por cada iniciativa avaliada. O resultado de toda essa investigação é apresentado na seção 5.1.

4.2.2.3 Etapa de tratamento dos resultados: proposta do modelo de governança

Conforme determina a análise de conteúdo de Bardin (2016), a etapa final do método se ocupa de estabelecer a sistematização dos passos relativos ao apropriado tratamento dos resultados, da inferência e da interpretação.

Portanto, o primeiro passo estabelecido para a elaboração do modelo foi dar tratamento aos resultados alcançados pela etapa de exploração do material, que teve como marco a elaboração e a validação do diagnóstico de maturidade.

O passo seguinte foi estabelecer as inferências obtidas a partir dos resultados alcançados pelo tratamento, chegando-se à elucidação de 3 (três) elementos relacionados a uma publicação de dados com qualidade em rede. Esses elementos irão compor a ideia subjacente ao modelo de governança, constituindo-se de verdadeiros insumos a sua proposição.

De posse das inferências, deu-se a interpretação final. Os 3 elementos se estabeleceram de forma harmônica com as 8 categorias de análise levantadas a partir dos desafios de abertura de dados, formando o núcleo do modelo. Para complementar, todas as ações e princípios estabelecidos no modelo de governança foram fundamentadas a partir de todas as inferências semânticas advindas do arcabouço teórico do estado da arte, da organização, da representação da informação, do conhecimento, e da governança de dados.

Finalmente, o modelo de governança foi dividido em dois artefatos eixos para melhor eficácia e interpretação das proposições, ambos indissociáveis e devendo operar de forma complementar, sendo eles: i) As atividades relacionadas ao modelo de governança proposto; ii) O modelo de governança voltado a qualidade de dados para publicação em rede. Os resultados dessa etapa são apresentados na seção 5.2.

5 RESULTADOS

O presente Capítulo se encarrega de cumprir as últimas etapas metodológicas da pesquisa, apresentando os resultados previstos da dissertação, quais sejam: a aplicação do modelo de diagnóstico de maturidade nos portais de dados abertos dos EUA, Reino Unido, Brasil e Irlanda, e, finalmente, a proposição do modelo de governança voltado a qualidade de dados para publicação em rede.

Portando, a Seção 5.1 trata da aplicação do diagnóstico de maturidade e a coleta dos resultados a partir da validação nos portais; enquanto a Seção 5.2 apresenta o modelo de governança, discorrendo sobre sua operação.

5.1 APLICAÇÃO DO DIAGNÓSTICO DE MATURIDADE

A presente seção detalha a aplicação do modelo de diagnóstico nos portais de dados abertos dos EUA, Reino Unido, Brasil e Irlanda. As avaliações foram divididas entre os 8 indicadores, com a exposição dos dados coletados em cada variável de requisito, concluindo com a apresentação da pontuação obtida pelos portais no respectivo indicador, acrescida de gráficos comparativos.

Ao final, serão apresentados os dados compilados e sintetizados frente aos 8 indicadores, com os portais analisados recebendo notas e avaliações finais respectivamente. A Seção é encerrada com uma discussão direcionada aos 4 portais avaliados, consolidando os resultados a serem trabalhados na proposição do modelo de governança.

A ordem dos indicadores segue a mesma do documento, quais sejam: i) Infraestrutura em TI; ii) Acessibilidade e Usabilidade; iii) Curadoria e Publicação de Dados; iv) Político e Social; v) Econômico; vi) Organização e Processos Internos; vii) Aspectos Legais e de Privacidade; e viii) Capacitação e Treinamento.

5.1.1 Indicador de Infraestrutura em TI

Com relação à Infraestrutura em TI, a iniciativa americana adota como DMS a plataforma CKAN versão 2.8.7, atendendo o nível 3 para a variável 1.1. Os *datasets* do portal informam as datas de publicação, a atualização e a descrição do editor, atendendo o nível 3 para a variável 1.2. A variável 1.3 está em conformidade com o

nível 1, sendo identificados *datasets* que oferecem o *link* para a sua fonte original, enquanto a variável 1.4 não possui informação disponível, ficando no nível 0.

A iniciativa do Reino Unido adota como DMS a plataforma CKAN em sua versão 2.8.3, atendendo o nível 3 da variável 1.1. Os *datasets* do portal informam apenas a data de atualização e a descrição do editor, atendendo ao nível 2 da variável 1.2. As variáveis 1.3 e 1.4 não foram identificadas na investigação do portal, ficando, ambas, no nível 0.

O portal brasileiro adota como DMS a plataforma CKAN em sua versão 2.6.8, atendendo o nível 3 para a variável 1.1. Os *datasets* do portal informam as datas de publicação, a atualização e a descrição do editor, atendendo o nível 3 para a variável 1.2. A variável 1.3 está em conformidade com o nível 1, sendo identificados *datasets* que oferecem o *link* para sua fonte original, enquanto a variável 1.4 não possui informação disponível, ficando no nível 0.

O portal irlandês utiliza como DMS o CKAN na versão 2.8.7, ficando a variável 1.1 no nível 3. Os *datasets* informam as datas de publicação e a atualização com a identificação do editor responsável, configurando o nível 3 para a variável 1.2. A variável 1.3 atingiu o nível 1, sendo disponibilizado um *link* para a fonte original. Na variável 1.4, o portal informou em seu monitor (IRLANDA, 2021b) a disponibilidade média de 99,6% durante o ano de 2021, atendendo ao nível 3. O Quadro 16 apresenta os resultados alcançados por cada portal na avaliação do indicador.

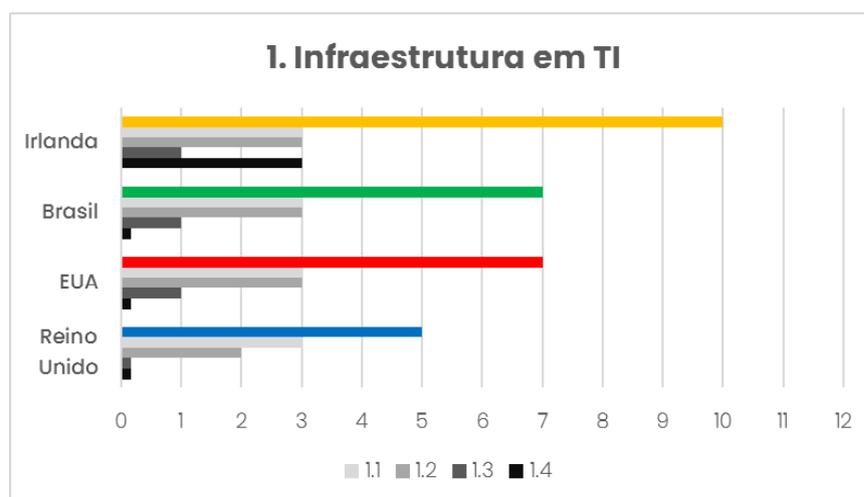
Quadro 16 - Resultados do indicador de Infraestrutura em TI

Infraestrutura em TI							
Posição	País	Pontuação nas variáveis				Total	%
		1.1	1.2	1.3	1.4		
1º	Irlanda	3	3	1	3	10	83,3%
2º	Brasil	3	3	1	0	7	58,3%
2º	EUA	3	3	1	0	7	58,3%
3º	Reino Unido	3	2	0	0	5	41,7%

Fonte: Elaborado pelo autor (2021).

A Irlanda liderou o *ranking* do indicador de Infraestrutura em TI, logo a seguir vieram empatados o Brasil e os EUA, e em último lugar o Reino Unido. Todas as iniciativas avaliadas aderiram a uma plataforma de código aberto gratuita para gerenciamento dos dados, especificamente o CKAN. O Gráfico 1 apresenta o desempenho comparativo entre os portais.

Gráfico 1 - Comparativo do indicador de Infraestrutura em TI



Fonte: Elaborado pelo autor (2021).

Em geral os *datasets* dos portais avaliados detalham bem seus conteúdos, ficando a desejar nas variáveis destinadas a informar o tempo de atraso de atualização e a disponibilidade dos dados, sendo apenas este último atendido plenamente pela Irlanda, o que explica sua folga na primeira posição.

5.1.2 Indicador de Acessibilidade e Usabilidade

Avaliando o portal norte-americano sobre a ótica do indicador de Acessibilidade e Usabilidade, a variável 2.1 fica em conformidade com o nível 3, possuindo muitos recursos de navegação, além de um projeto de designer gráfico consistente. A variável 2.2 está em conformidade com os recursos explanados no nível 3, com a organização temática de *datasets* sendo acessível pelo *link* de Tópicos. A variável 2.3 fica no nível 2, com os conjuntos de dados sendo bem descritos, detalhados, e, no caso dos geoespaciais contendo uma pré-visualização. Na variável 2.4, o portal oferece muitos formatos de *download*, suportando APIs, e oferta de coleção de dados, ficando, logo, no nível 3. A variável 2.5 atingiu o nível 2, sendo oferecido um canal de Contato para reportar erros nos *datasets*, mas não permitindo o seu acompanhamento (ESTADOS UNIDOS, 2021a). A variável 2.6 fica no nível 1, sendo identificado apenas o recurso de “submissão de história de dados”, levando ao Contato do portal que dispõem de um *link* para o fórum *Open Data Stack Exchange* (ESTADOS UNIDOS, 2021d), onde pode-se interagir e tirar dúvidas com outros usuários.

No Reino Unido, a variável 2.1 fica em conformidade com o nível 3, dispondo o portal de muitos recursos de navegação e com grande consistência visual, estando devidamente padronizado com todo o domínio governamental. A variável 2.2 também atinge o patamar mais elevado, conquistando o nível 3. Já na variável 2.3, o portal possui uma boa descrição dos *datasets*, com algumas bases oferecendo uma pré-visualização, ficando a variável 2.3 no nível 2. Na variável 2.4, o portal logra o nível 3, oferecendo *download* em muitos formatos, incluindo a existência de APIs como recurso para exportação de dados. A variável 2.5 também atingiu o nível 3, permitindo a qualquer indivíduo abrir uma reclamação na página de Suporte (REINO UNIDO, 2021a) e fazer o acompanhamento. Por último, o portal oferece apenas uma pesquisa de avaliação dos *datasets* (REINO UNIDO, 2021d) como forma de interação, ficando a variável 2.6 estagnada no nível 1 de maturidade.

No Brasil, a variável 2.1 fica em conformidade com o nível 3, possuindo uma navegação destacada e recursos auxiliares, além de um padrão gráfico coeso e consistente. A variável 2.2 está em conformidade com os recursos explanados no nível 2. Na variável 2.3, foram identificados muitos *datasets* com uma descrição enxuta, e não contando com outros recursos de apresentação, ficando no nível 1. Na variável 2.4, o portal oferece muitos formatos de *download*, com suporte a APIs, ficando em nível 3. A variável 2.5 atingiu o nível 3, sendo oferecido um canal de ouvidoria denominado Fala.br (BRASIL, 2021e) para abertura de denúncias. Na variável 2.6, o portal permite uma classificação do *dataset*, além de uma pesquisa de avaliação do conjunto de dados, portanto, a variável fica incluída no nível 2 por oferecer dois recursos de interação.

Já na Irlanda, o portal atingiu o nível 3 nas variáveis 2.1, 2.2, 2.3 e 2.4. Na variável 2.3, o acompanhamento gráfico foi identificado em alguns *datasets* de economia, não sendo observado em outros *datasets*. Na variável 2.4, o portal informa que dentre os 12.774 *datasets* disponibilizados, cerca de 7.627 suportam APIs. Na variável 2.5, o portal é aderente ao nível 1, oferecendo apenas o *e-mail* do editor para contato. Na variável 2.6, o portal atinge o nível 3, com recursos para classificar, comentar, e avaliar os *datasets* por um questionário de satisfação em Recursos e Publicações (IRLANDA, 2021c).

O Quadro 17 apresenta os resultados alcançados por cada portal, com a posição e respectiva pontuação de cada país.

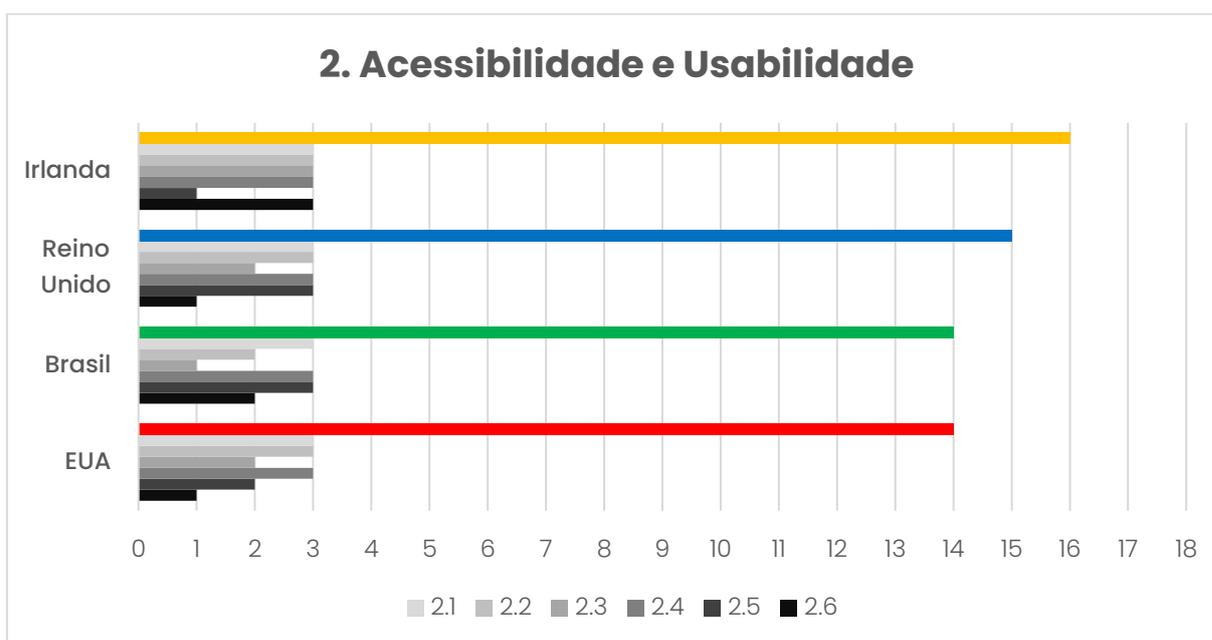
Quadro 17 - Resultados do indicador de Acessibilidade e Usabilidade

Acessibilidade e Usabilidade									
Posição	País	Pontuação nas variáveis						Total	%
		2.1	2.2	2.3	2.4	2.5	2.6		
1º	Irlanda	3	3	3	3	1	3	16	88,9%
2º	Reino Unido	3	3	2	3	3	1	15	83,3%
3º	Brasil	3	2	1	3	3	2	14	77,8%
3º	EUA	3	3	2	3	2	1	14	77,8%

Fonte: Elaborado pelo autor (2021).

A Irlanda ficou à frente no indicador de Acessibilidade e Usabilidade, com o Reino Unido em segundo lugar. Empatados em terceiro lugar vieram o Brasil e os EUA. As iniciativas possuem um projeto de designer gráfico padronizado, contendo vários recursos que facilitam a navegação. O Gráfico 2 exibe o desempenho comparativo entre os portais.

Gráfico 2 - Comparativo do indicador de Acessibilidade e Usabilidade



Fonte: Elaborado pelo autor (2021).

Esse foi um dos indicadores mais equilibrados, com as iniciativas apresentando uma boa capacidade nos níveis avaliados. O portal brasileiro perdeu performance em comparação às outras iniciativas no quesito de busca por *datasets* e apresentação dos dados, mas recuperou-se na interação que proporciona e no canal de ouvidoria para tratamento de erros. O portal irlandês tem como grande ponto fraco a inexistência de um canal especializado para reporte de problemas, afetando uma participação cidadã

mais efetiva. Já os portais dos EUA e do Reino Unido não exploraram todo o potencial interativo em seus conjuntos de dados, o que afeta a avaliação e a percepção da utilidade dos *datasets* frente ao escrutínio público.

5.1.3 Indicador de Curadoria e Publicação dos dados

Avaliando o portal dos EUA a partir do indicador de Curadoria e Publicação de Dados, a variável 3.1 é elucidada através do documento *OPEN Government Data Act*. Mesmo o portal incorporando fontes de dados de governos estaduais e locais, o documento aponta como requisito legal a abertura das bases federais especificamente (ESTADOS UNIDOS, 2017), limitando assim, a variável ao nível 1. Com relação a variável 3.2, tanto o Plano de Ação da Estratégia Federal de Dados (ESTADOS UNIDOS, 2020b), quanto a área de Recursos (ESTADOS UNIDOS, 2021a), estabelecem para qualidade de publicação de dados as diretrizes estabelecidas no nível 2, não mencionando aderência ao LOD. Na documentação técnica, a área de Recursos do portal fornece conteúdo variado relacionado aos padrões de metadados, fontes externas e dicionários (ESTADOS UNIDOS, 2021b), atendendo ao nível 3. Na variável 3.4, o portal estipula que apenas editores do governo podem publicar no sítio (ESTADOS UNIDOS, 2021a), configurando então aderência ao nível 0.

No Reino Unido, as variáveis 3.1 e 3.2 foram aferidas a partir da investigação do documento de Estratégia Nacional de Dados 2020 (REINO UNIDO, 2020). A partir da análise do documento é possível identificar que a abrangência da política de dados contempla o governo central e as autoridades regionais e locais de todo o Reino Unido, além disso a estratégia também inclui aderência ao padrão LOD (REINO UNIDO, 2020), logrando as variáveis 3.1 e 3.2 o nível 3. Com relação à documentação técnica, o portal ostenta uma área específica de Documentação (REINO UNIDO, 2021b), com informações acerca do uso de metadados orientados a publicação, ficando a variável 3.3 no nível 3. O portal estipula que apenas o governo central, autoridades locais e órgãos públicos podem publicar no sítio, ficando a variável 3.4 no nível 0.

No Brasil, as variáveis 3.1 e 3.2 foram elucidadas através da análise do documento de Plano de Dados Abertos 2021-2023 (BRASIL, 2021a). Apesar do portal estar aberto à publicação de dados de órgãos públicos de outras esferas, a abrangência do Plano é restrita apenas ao Poder Executivo Federal, ficando a variável

3.1 no nível 1. Com relação a variável 3.2, o Plano de Dados Abertos estabelece como referência para qualidade de dados o modelo LOD, atendendo ao nível 3. Na documentação técnica, o portal possui a área da Infraestrutura Nacional de Dados Abertos (INDA) contendo a cartilha de publicação e manual de catalogação (BRASIL, 2021d), atendendo ao nível 3. Na variável 3.4, o portal estipula que apenas órgãos públicos podem publicar no sítio, configurando então aderência ao nível 0.

Na Irlanda, as informações necessárias para determinar o nível das variáveis 3.1, 3.2 e 3.3 foram extraídas do portal através do *link* de Recursos e Publicações (IRLANDA, 2021a). Nesse diretório está contido a Estratégia de Dados Abertos 2017-2022 (IRLANDA, 2017), que atesta a abrangência da política de dados abertos para todos as esferas públicas do país, conferindo o nível 3 para a variável 3.1. As variáveis 3.2 e 3.3 também atendem ao nível 3, com as informações divulgadas respectivamente na página de Recursos e Publicações (*link do framework*) (IRLANDA, 2021a), e no documento de Diretrizes de Publicação (IRLANDA, 2021d). Com relação a variável 3.4, o link Editores (IRLANDA, 2021a) atesta que a tarefa de publicar dados no portal é restrita a editores do governo, ficando no nível 0.

O Quadro 18 apresenta os resultados alcançados por cada portal, com a posição e respectiva pontuação de cada país.

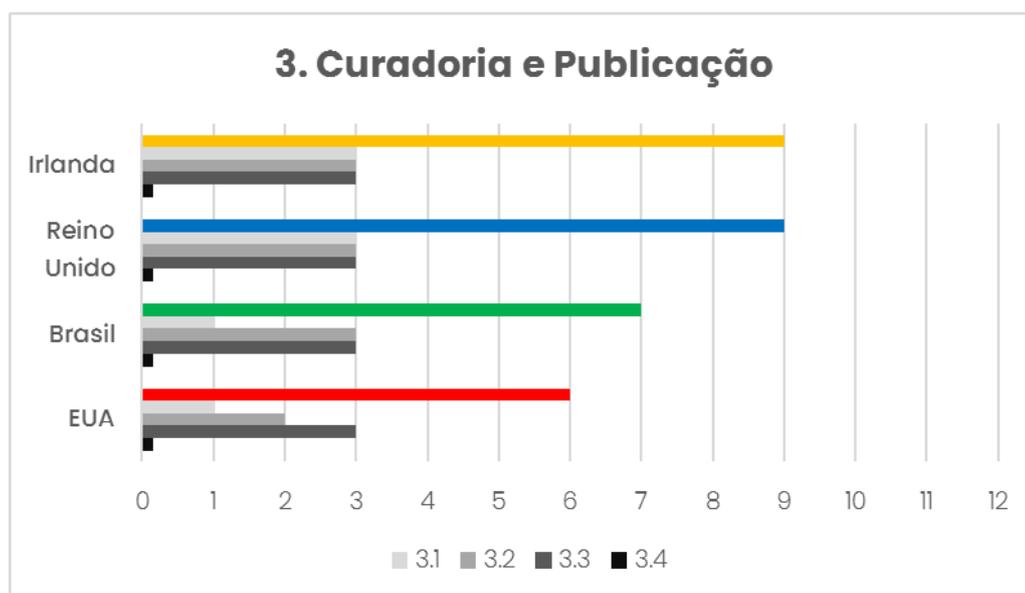
Quadro 18 - Resultados do indicador de Curadoria e Publicação dos Dados

Curadoria e Publicação dos Dados							
Posição	País	Pontuação nas variáveis				Total	%
		3.1	3.2	3.3	3.4		
1º	Irlanda	3	3	3	0	9	75,0%
1º	Reino Unido	3	3	3	0	9	75,0%
2º	Brasil	1	3	3	0	7	58,3%
3º	EUA	1	2	3	0	6	50,0%

Fonte: Elaborado pelo autor (2021).

Irlanda e Reino Unido ficaram empatados em primeiro lugar no indicador de Curadoria e Publicação dos Dados, com o Brasil em segundo lugar e os EUA em terceiro. Um ponto observável é que países continentais e federativos como o Brasil e EUA não estabeleceram a abrangência de sua política para todas as esferas públicas, isso os fez perder terreno na avaliação. O Gráfico 3 exhibe o desempenho comparativo entre os portais.

Gráfico 3 - Comparativo do indicador de Curadoria e Publicação dos Dados



Fonte: Elaborado pelo autor (2021).

No aspecto de qualidade de publicação, todas as iniciativas formalizam a aderência ao LOD em seus planos, excetuando-se os EUA. Com relação à disponibilização de material técnico, todas as iniciativas atenderam a alta capacidade, possuindo conteúdo satisfatório e devidamente *linkado* com recursos externos. O ponto negativo da avaliação do indicador foi a permanência restrita dos portais aos editores governamentais, ficando todos zerados em suas respectivas pontuações. Este é um terreno crítico, que ainda precisa ser muito debatido e trabalhado pelas iniciativas. O amadurecimento desse ponto permitirá um maior engajamento e participação cívica ao movimento de abertura, além de beneficiar a criação de uma cultura de dados na sociedade.

5.1.4 Indicador Político e Social

Iniciando o indicador na avaliação do portal norte-americano, a variável 4.1 atingiu o nível 2, podendo ser solicitado na página de Contato um novo conjunto de dados, porém, não sendo informado sobre os *datasets* abertos a partir dessa solicitação. A variável 4.2 pode ser analisada a partir da página de Destaques (ESTADOS UNIDOS, 2021a), sendo observado alguns exemplos com relatórios mais aprofundados, atingindo o nível 2. A variável 4.3 ficou em nível 2, com o monitor (ESTADOS UNIDOS, 2021e) do portal informando que dentre os 42 órgãos listados,

33 já estão abrindo suas bases, totalizando 78% de aderência. Na variável 4.4, não foi possível localizar concursos e *hackathons* recentes divulgados na página de Eventos (ESTADOS UNIDOS, 2021a), ficando a variável no nível 1.

No Reino Unido, a página de Suporte possui um formulário para a solicitação de novos *datasets* (REINO UNIDO, 2021a), porém não é informado quais conjuntos foram publicados através dessa prática, ficando a variável 4.1 no nível 2. Nenhum exemplo de reutilização de dados que estão trazendo impactos ou benefícios políticos, sociais e ambientais para o país foi divulgado, apenas a menção dos benefícios do uso de dados abertos no documento de Estratégia Nacional de Dados. Nenhum detalhamento do quantitativo de órgãos e agências que já abriram seus dados foi identificado, e finalmente nenhuma promoção a iniciativas de reutilização de dados com o intuito de criar aplicações para comunidades marginalizadas foi observado no portal, ficando, portanto, as variáveis 4.2, 4.3 e 4.4 no nível 0.

No Brasil, a página de Perguntas e Respostas esclarece que um *dataset* pode ser sugerido através do canal de ouvidoria Fala.br (BRASIL, 2021b). Como não existe uma transparência no portal que apresente os *datasets* implantados a partir dessa solicitação pública, a variável 4.1 fica confinada ao nível 2. A variável 4.2 atinge o nível 3, com a página de Aplicativos exibindo exemplos de reúso de dados que estão trazendo impactos (BRASIL, 2021b). Com relação a variável 4.3, o portal apresenta o monitor da CGU (BRASIL, 2021c), que atestou 74% dos entes federados abrindo suas bases, ficando a variável no nível 2. Na variável 4.4, através da página de Concursos (BRASIL, 2021b) foi possível identificar um *hackathon* de cunho social, mas datado de 2014, ficando a variável no nível 1.

Na Irlanda, a variável 4.1 é atendida ao nível 2, podendo-se sugerir novos *datasets* a partir de *link* Sugerir Dados na navegação principal (IRLANDA, 2021a), mas não ficando claro quais conjuntos de dados foram aceitos para publicação. A variável 4.2 pôde ser explorada pela Estratégia de Dados Abertos (IRLANDA, 2017), e pelos *links* de Impactos e Vitrines do portal (IRLANDA, 2021a), sendo esse último com exemplos de reutilização de dados para monitorar aspectos ambientais, de moradia e como o dinheiro público é gasto, garantindo a essa variável o nível 3.

Com relação à variável 4.3, o portal monitora os órgãos que publicam no sítio pelo *link* de Estatísticas (IRLANDA, 2021a), mas sem informar a totalidade e quais órgãos ainda não publicam, ficando no nível 0. Na variável 4.4, a área de Últimas

Notícias do portal (IRLANDA, 2021a) atestou que o último *hackathon* (*datathon*) promovido com viés social ocorreu em abril de 2020, ficando a variável no nível 1.

O Quadro 19 apresenta os resultados alcançados por cada portal, com a posição e respectiva pontuação de cada país.

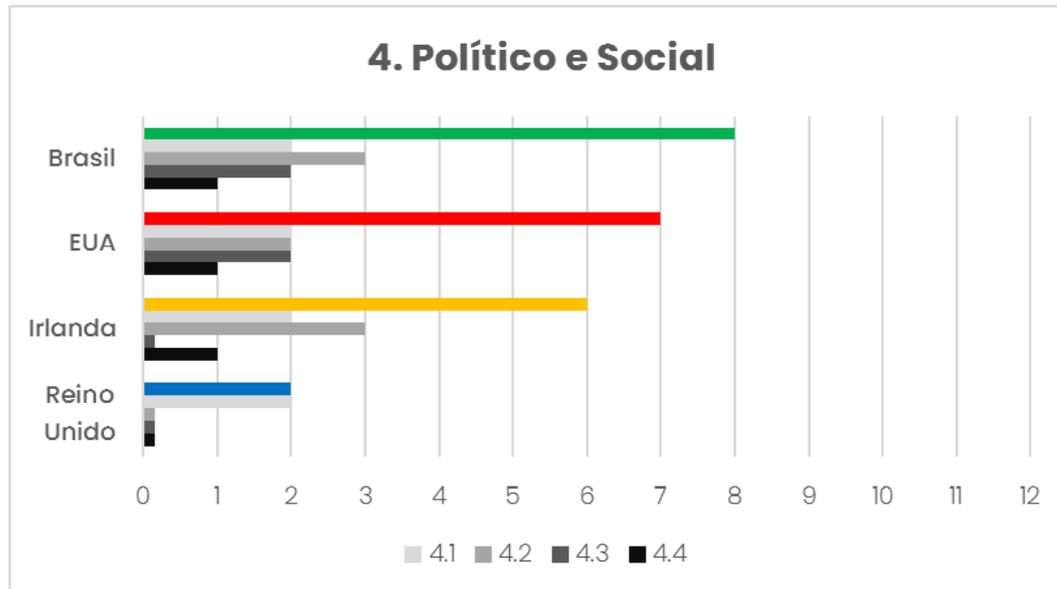
Quadro 19 - Resultados do indicador Político e Social

Político e Social							
Posição	País	Pontuação nas variáveis				Total	%
		4.1	4.2	4.3	4.4		
1º	Brasil	2	3	2	1	8	66,7%
2º	EUA	2	2	2	1	7	58,3%
3º	Irlanda	2	3	0	1	6	50,0%
4º	Reino Unido	2	0	0	0	2	16,7%

Fonte: Elaborado pelo autor (2021).

O Brasil assumiu o protagonismo no indicador Político e Social, ficando em primeiro lugar, seguido na sequência por EUA, Irlanda e Reino Unido. O Gráfico 4 exibe o desempenho comparativo entre os portais.

Gráfico 4 - Comparativo do indicador Político e Social



Fonte: Elaborado pelo autor (2021).

A performance geral foi muito insatisfatória, com apenas o Brasil e EUA acima de 50% de conformidade com o indicador, e ainda sim com valores relativos muito baixos. Irlanda atingiu 50% de conformidade e o Reino Unido ficou em último e muito abaixo com 16,7%. São resultados que preocupam para uma categoria tão importante e muito visada no material coletado a partir no referencial teórico.

Ficaram muito aquém em todas as iniciativas os quesitos de promoção de iniciativas de reutilização que estão acarretando benefícios sociais, divulgação de iniciativas de reutilização de dados em benefício de comunidades marginalizadas e transparência de abertura das agências do governo.

Com relação as consultas sobre quais *datasets* devem ter prioridade de abertura, todos os portais atingiram um nível estabelecido para a variável com oferta de pesquisas abertas, mas ainda faltando apresentar a avaliação e implementação dos *datasets* sugeridos. A conclusão que fica é que, ainda há muito o que se trabalhar sobre transparência nos portais, devendo os mecanismos de governança de dados atuarem com muita atenção nessa área, com estratégias e ações específicas. Também deve-se aumentar o envolvimento com outros setores da sociedade.

5.1.5 Indicador Econômico

Nos EUA, a variável 5.1 é atendida em nível 1, com o Plano de Ação apenas mencionando a importância dos dados para economicidade do governo e para a economia em geral, mas sem formalizar uma estratégia, metas e ações de ativação (ESTADOS UNIDOS, 2020b). Na variável 5.2, não foram localizados na página de Eventos iniciativas atualizadas de concursos ou *hackathons* em benefício da inovação no serviço público (ESTADOS UNIDOS, 2021a), ficando a variável no nível 1. A variável 5.3 pode ser acessada na página de Destaques, com exemplos de impactos econômicos pelo uso de dados abertos governamentais e relatórios agregados (ESTADOS UNIDOS, 2021a), atingindo a variável o nível 2.

No Reino Unido, a variável 5.1 é atendida em nível 3, com a Estratégia Nacional de Dados abordando planos para o aumento da produtividade, comércio, criação de novos negócios, empregos e fomento do uso sustentável dos dados. O plano deixa claro o escopo, as ações, as responsabilidades e o tempo de execução (REINO UNIDO, 2020). As variáveis 5.2 e 5.3 ficaram no nível 0, com o portal não divulgando nenhuma iniciativa ou impacto (REINO UNIDO, 2021a).

No Brasil, a variável 5.1 é atendida em nível 1, com o Plano de Dados Abertos apenas mencionando a importância da ativação de novos negócios, mas não formalizando uma estratégia de governança para atingir os objetivos (BRASIL, 2021a). Para a variável 5.2 foi analisado a página de Concursos, sendo possível identificar um concurso de criação de aplicações voltadas à gestão pública em 2021, mas

envolvendo apenas o setor público em sua organização (BRASIL, 2021b), ficando a variável no nível 2. Com relação à variável 5.3, não foram identificados exemplo de reutilização de dados que estão impactando a economia e os negócios do país (BRASIL, 2021b), ficando a variável no nível 0.

Na Irlanda, a variável 5.1 atingiu o nível 3, indicando que o Órgão de Governança de Dados Abertos (*Open Data Governance Board* - ODGB) detalha no documento de Estratégia de Dados Abertos, o seu plano de implementação (IRLANDA, 2017). O portal também relata uma iniciativa de engajamento para abertura de dados que se constitui de um fundo que visa financiar uma grande variedade de atividades relacionadas a dados abertos (IRLANDA, 2021a). Na variável 5.2, o último concurso ou competição para promoção de soluções para o serviço público consta de agosto de 2020 (IRLANDA, 2021a), ficando a variável no nível 1. Já a variável 5.3 atingiu o nível 3, podendo ser explorada pela Estratégia de Dados Abertos e pelo link Vitrines, destacando-se aplicações como a *Wripl Technologies*, que fornece serviços digitais de valor agregado baseado em dados abertos governamentais para outros negócios (IRLANDA, 2017), e o *Monetising MetOcean Data*, um *software* útil para que governos e empreendedores percebam o potencial financeiro dos recursos renováveis do país (IRLANDA, 2021a). O Quadro 20 apresenta os resultados alcançados por cada portal.

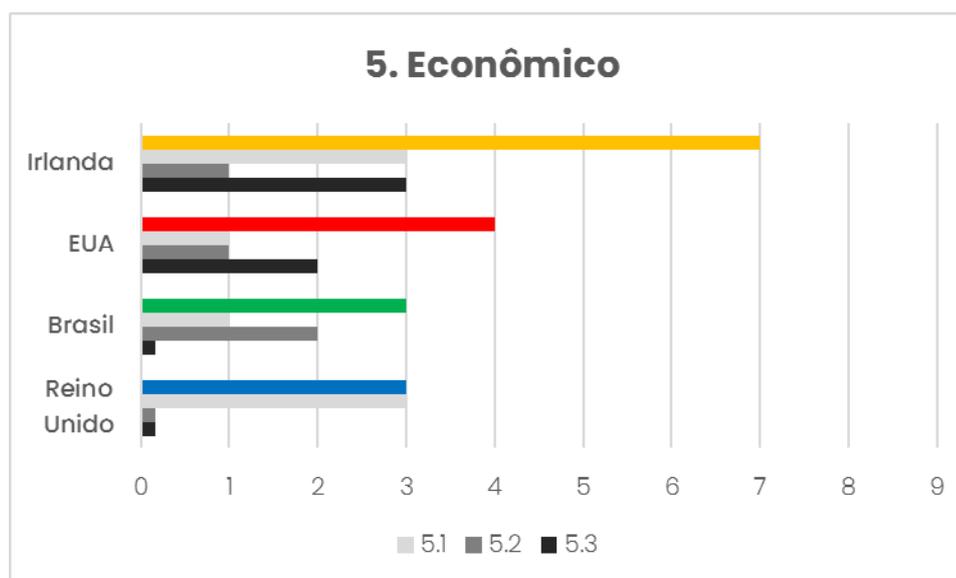
Quadro 20 - Resultados do indicador Econômico

Econômico						
Posição	País	Pontuação nas variáveis			Total	%
		5.1	5.2	5.3		
1º	Irlanda	3	1	3	7	77,8%
2º	EUA	1	1	2	4	44,4%
3º	Brasil	1	2	0	3	33,3%
3º	Reino Unido	3	0	0	3	33,3%

Fonte: Elaborado pelo autor (2021).

A Irlanda ficou em primeiro lugar novamente, seguido dos EUA. Dividindo o terceiro lugar ficaram o Brasil e o Reino Unido. O Gráfico 5 exhibe o desempenho comparativo entre os portais.

Gráfico 5 - Comparativo do indicador Econômico



Fonte: Elaborado pelo autor (2021).

O portal da Irlanda liderou com folga o indicador, especialmente por contemplar em suas políticas as estratégias de sustentabilidade econômica, exemplificando várias aplicações que já estão ativando novos negócios a partir da reutilização dos dados abertos. Os demais países ficaram abaixo dos 50% de conformidade, precisando trabalhar a promoção de iniciativas de dados e publicizar melhor as soluções que se propõem a gerar valor a partir dos dados.

Concluindo o indicador, é necessário incluir nos planos de ação de governança de dados a prospecção em nível nacional dos exemplos de reuso de dados que já estão transformando a economia no país, divulgando os resultados no portal. A política deve agregar estratégias para a ativação de novos negócios e apresentar os resultados. A promoção de iniciativas de reutilização de dados abertos deve ter maior periodicidade, com envolvimento de membros da sociedade.

5.1.6 Indicador Organização e Processos Internos

Na avaliação do portal norte-americano, a variável 6.1 é atendida pela constatação no Plano de Ação (ESTADOS UNIDOS, 2020b) de uma sistematização de diversas práticas a serem executadas para a construção de uma cultura administrativa que valorize e promova o uso dos dados, ficando a variável, portanto, no nível 3. A variável 6.2 também atinge o nível 3 a partir da análise da Ação 13 contida no documento formal do Plano, que define que as agências precisam gerenciar dados

como um ativo estratégico e tomar decisões baseadas em dados. O escopo, a responsabilidade e as datas estão expressas no documento para a ação citada (ESTADOS UNIDOS, 2020b). A variável 5.3 não pôde ser identificada no portal, ficando no nível 0.

No Reino Unido, as variáveis 6.1 e 6.2 atingiram o nível 3, estando contempladas na Estratégia Nacional de Dados, que prevê a evolução da cultura de dados e o seu reúso como ativos para a melhoria da administração pública (REINO UNIDO, 2020). A variável 6.3 não pôde ser identificada no portal, ficando no nível 0.

No Brasil, a variável 6.1 atinge o nível 3, com o Plano de Ação do INDA (BRASIL, 2021f) definindo a ação, as responsabilidades e as metas de criação da cultura de dados no serviço público, sistematizando o processo. A variável 6.2 não foi identificada na política de dados abertos, ficando no nível 0, enquanto a variável 6.3 não foi localizada no portal, compartilhando o mesmo nível 0 de maturidade.

Na Irlanda, as variáveis 6.1 e 6.2 foram elucidadas a partir da análise do documento de Estratégia de Dados do Serviço Público 2019-2023. O documento transparece uma governança de dados, que define o escopo, as responsabilidades, as funções e as ações sistematizadas para a reutilização dos dados. Entre os vários objetivos está o fomento de uma cultura de dados na administração, além da reutilização dos dados baseados em evidências como suporte na tomada de decisão (IRLANDA, 2019), logrando ambas as variáveis o nível 3. A variável 6.3 não pôde ser identificada no portal, ficando no nível 0 de maturidade. A seguir, o Quadro 21 apresenta os resultados alcançados por cada portal no indicador.

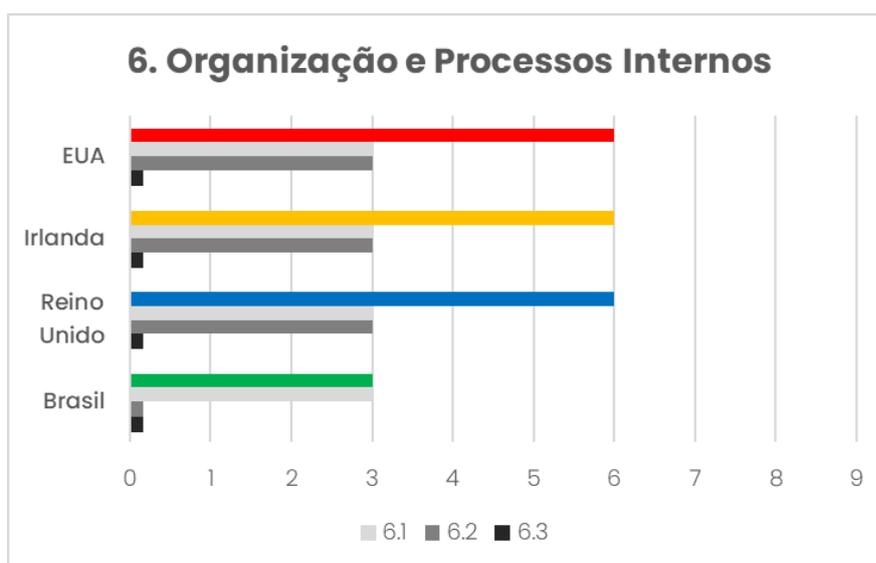
Quadro 21 - Resultados do indicador de Organização e Processos Internos

Organização e Processos Internos						
Posição	País	Pontuação nas variáveis			Total	%
		6.1	6.2	6.3		
1º	EUA	3	3	0	6	66,7%
1º	Irlanda	3	3	0	6	66,7%
1º	Reino Unido	3	3	0	6	66,7%
2º	Brasil	3	0	0	3	33,3%

Fonte: Elaborado pelo autor (2021).

Os EUA, a Irlanda e o Reino Unido ficaram empatados em primeiro lugar, seguido pelo Brasil. O Gráfico 6 exibe o desempenho comparativo entre os portais.

Gráfico 6 - Comparativo do indicador de Organização e Processos Internos



Fonte: Elaborado pelo autor (2021).

EUA, Irlanda e Reino Unido atingiram máxima conformidade nas variáveis que objetivam compreender a maturidade da governança de dados, contemplando em seus planos e ação e política de dados estratégias para melhorar a cultura e uso dos dados. O Brasil apresentou lacunas a serem tratadas pela gestão de dados.

Como elucidado na revisão bibliográfica, não basta apenas elaborar um documento que detenha uma governança impecável, com ações bem redigidas e formalizadas, é necessário colocar em prática todas as atividades estabelecidas e publicizá-las ao escrutínio público utilizando-se o portal para isso. Os portais estão falhando no aspecto transparência, não exibindo exemplos, impactos ou resultados advindos dessas ações, indiciando um perigoso obscurantismo sobre a real efetividade da política na gestão do recurso de dados dentro das organizações.

5.1.7 Indicador de Aspectos Legais e de Privacidade

No EUA, a variável 7.1 é atendida em nível 3, com a documentação da legislação que regulariza a abertura de dados disponível no *link* de Recursos (ESTADOS UNIDOS, 2021b). Na variável 7.2, a informação sobre a licença de uso encontra-se nos *datasets*, porém, tiveram ocorrências de conjuntos sem o detalhe da licença (ESTADOS UNIDOS, 2021a), ficando a variável no nível 1. Na variável 7.3, o portal oferece o *e-mail* do editor e um formulário de contato genérico (ESTADOS

UNIDOS, 2021a) para possíveis ações de denúncia para *datasets* que estejam em desacordo com aspectos legais e de privacidade, ficando a variável no nível 1.

No Reino Unido, a legislação pode ser acessada no rodapé do portal (REINO UNIDO, 2021a), estando a variável 7.1 no nível 3. Na variável 7.2, a licença de uso está sempre presente na descrição dos *datasets* e localizada com destaque (REINO UNIDO, 2021a), dessa forma a variável também atinge o nível 3. Na variável 7.3, o portal oferece o *e-mail* do editor ou o formulário de Suporte para os casos de não se tratar o elemento privacidade (REINO UNIDO, 2021a), ficando no nível 1.

No Brasil, a variável 7.1 é atendida em nível 3, com toda documentação pertinente ao tema podendo ser acessada na página da INDA (BRASIL, 2021d). Na variável 7.2, a licença de uso está presente e com detalhamento de sua abrangência, mas localizado de forma muito oculta na página (BRASIL, 2021b), o que rebaixa a variável para nível 2. A variável 7.3 pode ser atendida pelo canal de ouvidoria oferecido no portal (BRASIL, 2021e), ficando no nível 3; todavia, tal recurso poderia estar mais visível e destacado nas páginas dos *datasets*.

Na Irlanda, a variável 7.1 ficou no nível 3, com as informações sobre a licença de dados abertos e a legislação de reutilização dos dados disponíveis na página de Recursos e Publicações (IRLANDA, 2021a). A variável 7.2 atingiu o nível 3, estando a informação sobre a licença disposta de forma visível e com detalhes de sua abrangência através de *link* externo (IRLANDA, 2021a). Na variável 7.3, excetuando o *e-mail* do editor de dados, os usuários não possuem formas de alertar de forma mais assertiva erros de privacidade nos *datasets* (IRLANDA, 2021a), ficando a variável no nível 1. O Quadro 22 apresenta os resultados do indicador.

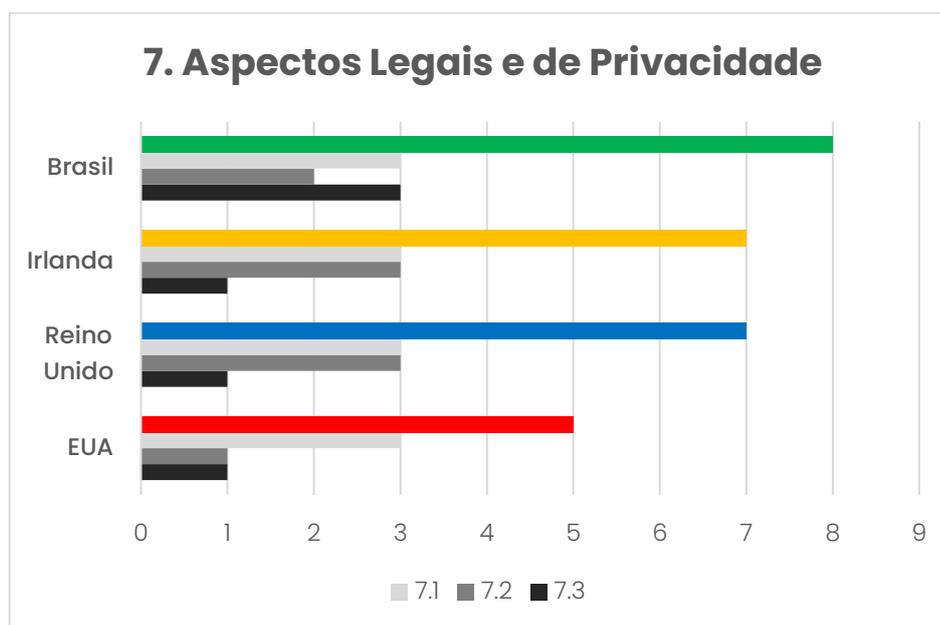
Quadro 22 - Resultados do indicador de Aspectos Legais e de Privacidade

Aspectos Legais e de Privacidade						
Posição	País	Pontuação nas variáveis			Total	%
		7.1	7.2	7.3		
1º	Brasil	3	2	3	8	88,9%
2º	Irlanda	3	3	1	7	77,8%
2º	Reino Unido	3	3	1	7	77,8%
3º	EUA	3	1	1	5	55,6%

Fonte: Elaborado pelo autor (2021).

O Brasil liderou esse indicador, seguido de Irlanda e Reino Unido empatados em segunda posição, com os EUA em último. O Gráfico 7 exhibe o desempenho comparativo entre os portais.

Gráfico 7 - Comparativo do indicador de Aspectos Legais e de Privacidade



Fonte: Elaborado pelo autor (2021).

Os portais tiveram boa conformidade nas variáveis, ficando os EUA em derradeira posição com relação à exibição da licença e oferta de canal de comunicação de erros. Todo o conteúdo a respeito do ordenamento legal que rege a abertura de dados foi identificado nas iniciativas. A exibição da licença precisa ser trabalhada de forma mais transparente nos *datasets*, sendo um exemplo de alta capacidade as experiências da Irlanda e Reino Unido.

A parte de abertura de chamado para erros de legalidade e privacidade nos *datasets* ficou a desejar, excetuando-se o Brasil, que propõe um interessante canal de ouvidoria que engloba vários tipos de denúncias. O ponto discutível na realidade brasileira é a necessidade de exigir-se, durante o acesso, a efetivação de um cadastro prévio para chamados que demandem acompanhamento. No entanto, torna-se uma boa solução para trabalhar no canal de denúncias, devendo ser um modelo a ser considerado em ações de melhoria da dos processos dirigidos à formalização de denúncias nos portais.

5.1.8 Indicador de Capacitação e Treinamento

Avaliando o portal dos EUA, a variável 8.1 atinge o nível 1 com a área de promoção de eventos existente, mas desatualizada (ESTADOS UNIDOS, 2021a). Na variável 8.2, o Plano de Ação faz apenas algumas menções sobre o uso de dados na

educação (ESTADOS UNIDOS, 2020b), ficando no nível 1. A variável 8.3 atingiu o nível 3, com a seção de governança de dados do diretório de Recursos apresentando guias, cursos, e aplicações destinadas a capacitação (ESTADOS UNIDOS, 2021b).

No Reino Unido, não foram identificados no portal a promoção de eventos com a temática de Dados Abertos, como também não foram localizados conteúdos que visam capacitar os funcionários em competência de dados, portanto, as variáveis 8.1 e 8.3 ficarem no nível 0. A variável 8.2 atingiu o nível 3, com o documento da Estratégia Nacional de Dados elencando a decisão do governo de inserir nos currículos escolares e nas universidades a disciplina de Ciência de Dados. Tal decisão partiu do envolvimento de agências do governo e organizações representativas na sociedade, como a *Royal Society*, instituição destinada à promoção do conhecimento científico, e *National Endowment for Science, Technology and the Arts (Nesta)*, instituição independente que trabalha para aumentar a inovação no país (REINO UNIDO, 2020).

No Brasil, a variável 8.1 atingiu o nível 2 a partir da identificação no portal de eventos recentes para promoção da temática de dados abertos, mas organizadas apenas para o setor público (BRASIL, 2021b). Na variável 8.2, o portal não divulga as ações que porventura estejam executando em prol da educação de dados nos currículos, ficando no nível 0. E finalmente, a variável 8.3 atinge o nível 1 a partir da apreciação do conteúdo disponibilizado no *link* do INDA (BRASIL, 2021d).

No portal da Irlanda, foi possível identificar eventos com a temática de dados abertos no ano de 2021 envolvendo representantes da sociedade (IRLANDA, 2021a), dessa forma a variável 8.1 logrou o nível 3. O portal ou a Estratégia de Dados não forneceu informações sobre as competências de dados nos currículos educacionais, ficando a variável 8.2 no nível 0. Sobre treinamentos e cursos de capacitação para os funcionários, o portal fornece material de instrução e cursos de capacitação (IRLANDA, 2021a), estabelecendo o nível 2 para a variável 8.3. O Quadro 23 apresenta os resultados alcançados por cada portal no indicador.

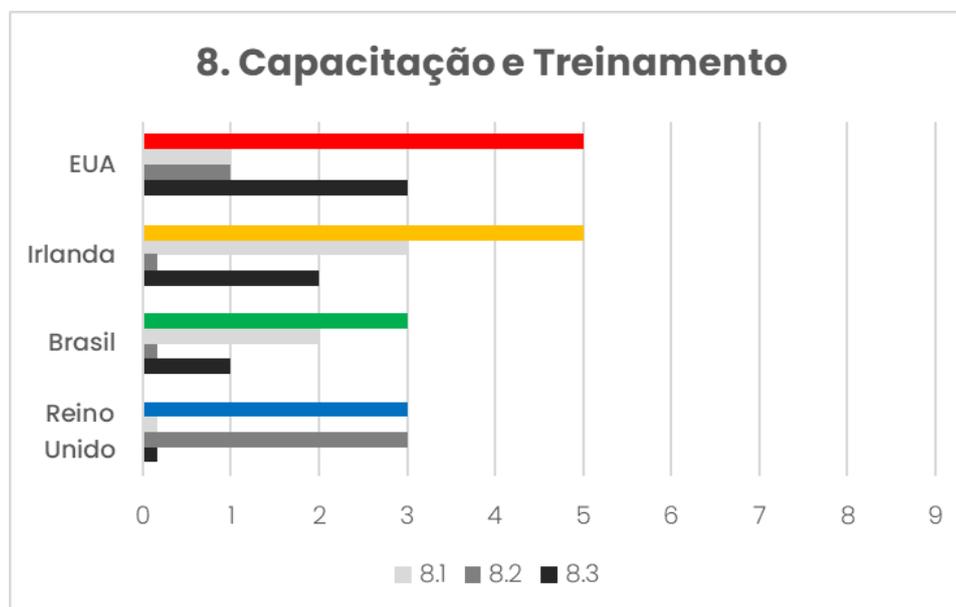
Quadro 23 - Resultados do indicador de Capacitação e Treinamento

Capacitação e Treinamento						
Posição	País	Pontuação nas variáveis			Total	%
		8.1	8.2	8.3		
1º	EUA	1	1	3	5	55,6%
1º	Irlanda	3	0	2	5	55,6%
2º	Brasil	2	0	1	3	33,3%
2º	Reino Unido	0	3	0	3	33,3%

Fonte: Elaborado pelo autor (2021).

Os EUA e a Irlanda lideraram o indicador mantendo-se empatados, seguidos por Brasil e Reino Unido que dividiram o segundo lugar. O Gráfico 8 exibe o desempenho comparativo entre os portais.

Gráfico 8 - Comparativo do indicador de Capacitação e Treinamento



Fonte: Elaborado pelo autor (2021).

O resultado indica que as iniciativas avaliadas ainda precisam trabalhar consideravelmente o aspecto de capacitação e treinamento, com todas abaixo de 60% de conformidade com o indicador. Pouca informação relacionada as ações para desenvolver habilidades de dados nos currículos escolares são compartilhadas, e, excetuando-se o exemplo irlandês, as iniciativas de promoção de eventos em dados abertos estão em níveis ainda pouco amadurecidos.

Com relação à disponibilização de conteúdo em dados o portal dos EUA destaca-se oferecendo recursos para a capacitação em competência de dados. Por ser aberta, ou seja, não demandando nenhum pré-cadastro para seu acesso, a capacitação pode ser usufruída tanto por funcionários do governo, quanto por qualquer outro indivíduo que esteja interessado em aprender mais sobre os dados abertos governamentais.

A seguir, a próxima Seção será destinada a apresentar o resultado consolidado da avaliação, contendo as respectivas posições dos portais dos EUA, Reino Unido, Brasil e Irlanda, todos validados a partir da aplicação do modelo de diagnóstico de maturidade.

5.1.9 Resultado consolidado da avaliação nos portais governamentais

A utilização do instrumento de diagnóstico serviu para levantar o nível de maturidade dos portais avaliados a partir dos indicadores pré-definidos. O resultado contendo a posição de cada iniciativa é apresentado no Quadro 24, possibilitando depreender um panorama apropriado acerca do estado de maturidade dos portais e promover considerações. O Quadro 24 também informa o nível de maturidade atingido pelos portais a partir da comparação com o Quadro 13, além da conformidade percentual de cada iniciativa com o diagnóstico de maturidade. Para se chegar ao nível de maturidade foi calculada a média da pontuação final de cada iniciativa em relação à quantidade de variáveis do diagnóstico.

Quadro 24 - *Ranking* dos portais avaliados a partir do diagnóstico

Posição	País	Pontuação final	Nível de maturidade	Conformidade com o diagnóstico
1º	Irlanda	66	2,20	73,3%
2º	EUA	54	1,80	60,0%
3º	Brasil	53	1,76	58,9%
4º	Reino Unido	50	1,66	55,5%

Fonte: Elaborado pelo autor (2021).

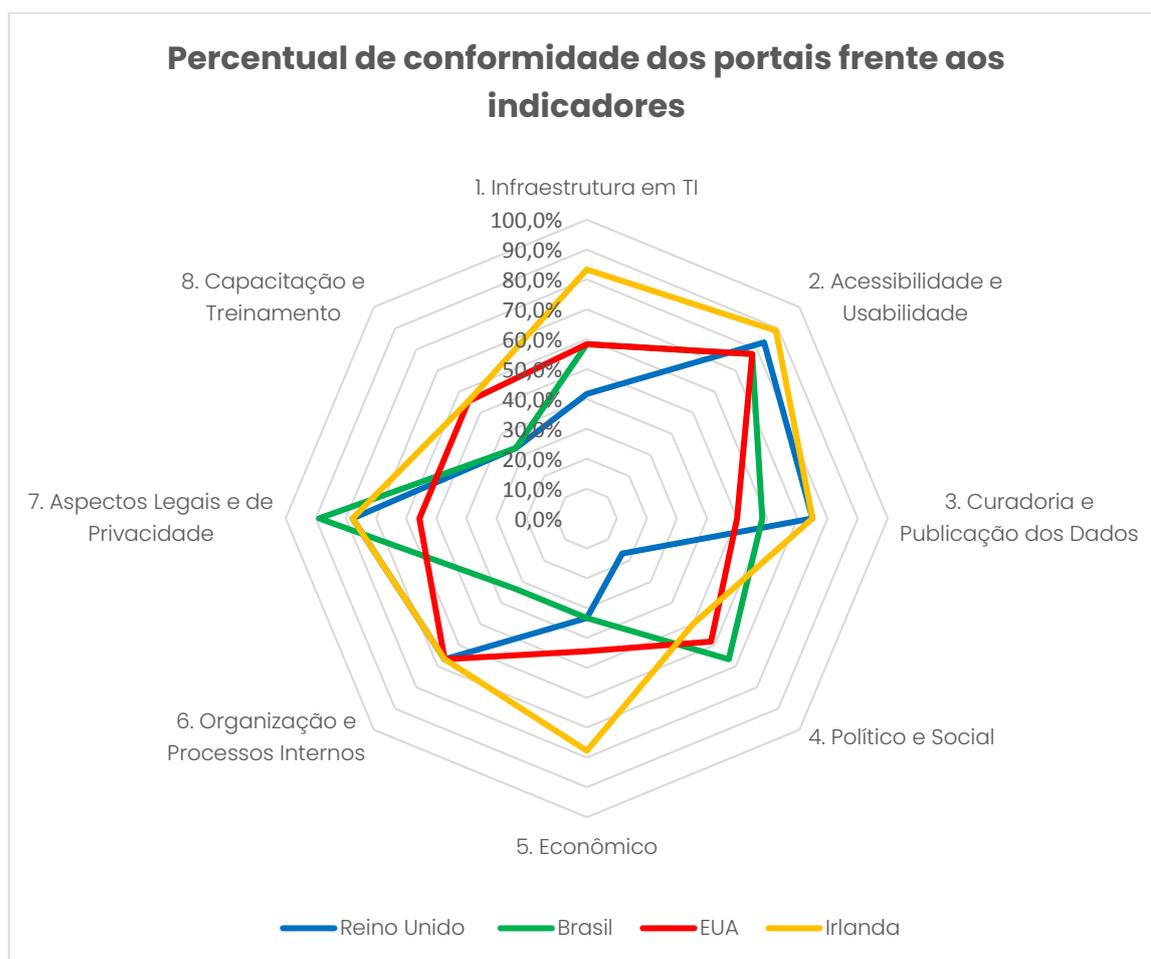
O primeiro lugar geral ficou com a Irlanda, seguindo de um empate técnico entre EUA e Brasil, ficando a iniciativa norte-americana em segundo lugar e a brasileira em terceiro, e o Reino Unido na última posição.

Com relação ao nível de maturidade, a Irlanda apresentou uma preocupação com a política de dados, a sustentabilidade e a modernidade do portal. A iniciativa consolida-se em um estágio estabelecido de maturidade, ficando no nível 2, devendo aperfeiçoar o aspecto social e a execução das ações de governança.

Os EUA possuem um portal com muitos recursos, mas que carece de maior engajamento social e de atualização. Falta abrangência às políticas, que não vêm sendo aplicadas de forma satisfatória, lançando dúvidas sobre o comprometimento do governo com uma boa gestão dos dados, ficando o portal no nível 1 de maturidade.

O Reino Unido também ficou estagnado no estágio 1, com o portal apresentando pouca preocupação social, sem explicitar os impactos dos dados, e pouca transparência nas ações desencadeadas por sua governança. O Gráfico 9 exhibe a relação de cada país com os respectivos indicadores da pesquisa.

Gráfico 9 - Desempenho dos portais frente aos indicadores



- **Irlanda**

Iniciando a discussão a partir do líder do *ranking*. A performance da Irlanda apontada no Gráfico 9 se sobrepõe às demais iniciativas nos indicadores de Infraestrutura em TI e Econômico. A julgar pelo fundo de investimentos em ações de reúso de dados abertos implantado pelo governo, sugere-se que as atividades para a geração de valor econômico, a partir dos dados, estejam logrando êxito pelo número de aplicações exemplificadas no portal com as quais novos negócios estão sendo ativados, o que torna a economia mais vibrante e pujante.

O retorno de divisas para os cofres públicos oportuniza o reinvestimento no setor de dados do governo garantindo recursos para a governança de dados aperfeiçoar o setor tecnológico, o que justificaria a boa capacidade em TI.

Os indicadores de Curadoria e Publicação dos Dados, Aspectos Legais e de Privacidade, e Acessibilidade e Usabilidade possuem em média uma boa capacidade de maturidade para publicação de dados abertos governamentais com qualidade em rede. No primeiro indicador citado, um aspecto a ser evoluído é o maior envolvimento de quaisquer membros da sociedade na publicação de dados no portal, uma questão que foi recorrente nos portais avaliados.

Também faltam investimentos no portal para denúncias nos *datasets* de erros técnicos e problemas com aspectos legais e de privacidade. O governo tem abertura para tornar a ouvidoria mais efetiva no portal, estabelecendo maior diálogo sobre a participação mais efetiva da sociedade no portal, especialmente na publicação.

Os três últimos indicadores da Irlanda apresentaram a pior performance no aspecto de designer gráfico: Político e Social, Organização e Processos Internos e Capacitação e Treinamento. No primeiro indicador, a Irlanda falha em não prover um painel de transparência com a finalidade de monitorar a publicação das agências, e não informar sobre as iniciativas de reuso de dados em benefício das comunidades menos favorecidas.

Existem exemplos bem-sucedidos elucidados na pesquisa, como o monitor de disponibilidade dos servidores de dados e o fundo de investimento econômico, que poderiam ser replicadas para sanar essas lacunas. Nos dois últimos indicadores mencionados, a chave para o aumento da conformidade estaria novamente no aumento da transparência do governo no portal, apresentando os benefícios que a reutilização dos dados produz à administração pública, além do andamento dos investimentos em competência de dados no seu sistema educacional.

- **Estados Unidos**

Nos EUA, as melhores performances foram nos indicadores de Acessibilidade e Usabilidade e Organização e Processos Internos, com uma conformidade entre 77,8% e 66,7% respectivamente. O Gráfico 9 tratando da performance norte-americana demonstra uma certa estabilidade, não exibindo grandes picos, estando os demais indicadores em uma conformidade de 60% e 50%, excetuando-se o indicador Econômico com 44,4%. A performance em Organização e Processos Internos permitiu aos EUA liderar o indicador, que apesar da conformidade de 66% apresentou alguns ensinamentos.

As estratégias de governança de dados dos EUA são bem formuladas e apresentam grande detalhamento técnico, especialmente seu Plano de Ação. O escopo, as atividades, as ações, as atribuições de responsabilidade e as metas costumam estar contempladas e bem documentadas, mas, como apontado pela revisão de literatura, a avaliação do seu portal tem piorado e regredido nos *rankings* de abertura de dados governamentais.

A percepção é de que as políticas ainda não vêm sendo aplicadas de forma satisfatória, o que poderia explicar o estabelecimento dos indicadores de Infraestrutura em TI, Curadoria e Publicação de Dados, Políticos e Sociais, Econômico, Legais e de Privacidade, e Capacitação e Treinamento em níveis medianos de maturidade, flertando entre uma capacidade inicial ou estabelecida.

Por exemplo, ao contrário dos outros portais avaliados, a estratégia norte-americana não contemplou o LOD, o que lança dúvidas sobre o compromisso do atual governo em aderir as boas práticas em governança de dados.

Também é perceptível a falta de um maior engajamento e participação cívica no portal, com muitos poucos pontos de contato interativo entre os membros da sociedade e a plataforma digital. As áreas de eventos e impactos estão refletindo o pouco engajamento, estando desatualizadas em sua maioria e com pouca participação de setores da sociedade em sua promoção.

- **Brasil**

Já na avaliação da experiência brasileira, o Gráfico 9 apresentou um grande desbalanceamento, com um pico no indicador de Aspectos Legais e de Privacidade, em contraste com as baixas performances em Organização e Processos Internos, Econômico, e Capacitação e Treinamento. O pré-teste no portal já havia indicado grande preocupação com o aspecto de legislação, o que refletiu na pontuação desse indicador. Os demais indicadores tiveram as variáveis com capacidade inicial ou estabelecida de média.

No indicador de Curadoria e Publicação dos Dados, o Brasil, a exemplo dos EUA, tem uma abrangência limitada da execução da sua política de dados à esfera federal. Mesmo sendo perceptível algumas publicações de outros entes em seu portal, a padronização da política de dados entre todas as esferas é um ponto que precisa de maior discussão e consenso para uma maior qualidade de publicação,

especialmente para o compartilhamento de padrões de interoperabilidade, como é o caso do ePING.

Semelhante a maioria das iniciativas investigadas, o Brasil apresenta em seu plano de dados a referência ao LOD para qualidade de dados, além de uma boa documentação técnica no portal referente as questões de publicação. No entanto, conforme elucidado por Macedo e Lemos (2019), o que se vê na prática é uma baixa adesão ao LOD, com os governos se conformando em atingir o nível de três estrelas. Essa situação escancara o baixo compromisso do governo em tornar realidade os próprios preceitos formalizados em seus instrumentos de política. Não é efetivado um plano de ação claro, com as responsabilidades, metas e prazo para colocar em prática todo o potencial do LOD. Apenas mencionar a adoção de boas práticas pode induzir as pessoas à interpretarem erroneamente que o governo está de fato aplicando essas medidas de forma sistemática.

No indicador de Organização e Processos Internos, o Brasil teve a pior pontuação entre os portais avaliados. Mesmo contemplando, na sua política, a criação de uma cultura de dados, não foram identificados no documento as estratégias para a reutilização dos dados no âmbito da administração pública. A exemplo das outras iniciativas analisadas, o portal precisa tornar mais transparente as ações governamentais que estão sendo tomadas nesse sentido, e publicizar os resultados alcançados em área apropriada do portal para apreciação pública.

O que se viu também no indicador de Capacitação e Treinamento foi uma baixa performance do Brasil. Apesar de ser perceptível no portal eventos recentes para promoção da temática de dados abertos, não foi possível identificar nos documentos de política e plano de ação, e tampouco no portal, as ações que porventura o governo esteja adotando em prol da educação de dados nos currículos nacionais. Mesmo o conteúdo relativo à capacitação em competências de dados abertos aos funcionários precisa de melhorias. Esses aspectos devem ser discutidos conjuntamente com órgãos que representam a sociedade e a academia, para que se fomente não apenas a cultura de dados no país, mas que se alcance um maior patamar de qualidade nos dados publicados em rede.

O indicador Econômico brasileiro foi afetado por uma manifestação pouco clara ou inexistente das ações de governança para ativação de novos negócios a partir do uso dos dados. O governo brasileiro e seus comitês de governança precisam estar mais atentos aos processos de geração de valor estratégico a partir da reutilização

dos dados abertos. Gerar atividades de valor pode alavancar o portal contribuindo para torná-lo sustentável, o que seria muito bem-vindo para o constante aprimoramento e reinvestimento de recursos em prol da plataforma.

As boas referências identificadas na avaliação do portal do Brasil foram a maior preocupação social, levando a iniciativa brasileira a liderar o indicador Político e Social. Destaca-se também o recurso de ouvidoria, que concentra em apenas um lugar importantes ações de interação com o portal e o governo, tais como: espaço para a sugestão de novos *datasets*, denúncias de erros técnicos no portal, e, denúncias de problemas legais e de privacidade.

- **Reino Unido**

Por fim, a análise do gráfico do Reino Unido apontou três indicadores que se destacam na sua análise individual: Acessibilidade e Usabilidade, Curadoria e Publicação de Dados e Aspectos Legais e de Privacidade. O portal foi um dos líderes no indicador de Curadoria e Publicação de Dados, se destacando com uma política de dados contendo estratégias de governança de dados bem detalhados e uma boa documentação técnica. No entanto, esses resultados conflitam com a revisão de literatura que apontam a nação como uma das piores no quesito de publicação, com o pré-teste sugerindo *datasets* em baixa conformidade com os melhores formatos.

A explicação pode ser dada a partir da constatação na revisão de literatura sobre a realidade organizacional do governo, e confirmada na presente pesquisa através da aplicação do diagnóstico de maturidade e com exposição dos resultados pelo Gráfico 9, ou seja: uma baixa maturidade dos indicadores de Organização e Processos, Capacitação e Treinamento, e uma baixa maturidade na questão política e social. A baixa maturidade nesses níveis indica que as estratégias e planos de ação desenhadas na política de dados não estão sendo aplicadas de forma devida, lançando dúvidas sobre a efetividade da governança de dados nacional.

O portal precisa dar maior transparência aos resultados e impactos do reuso dos dados econômicos, sociais e políticos. Como frisado pela pesquisa, não basta definir ações e métodos para se alcançar a maturidade de publicação ao nível LOD, se um grande número de *datasets* continua aberto em formatos inadequados como demonstra a literatura e os pré-testes durante a pesquisa.

A partir da avaliação dos portais frente aos 8 indicadores e dos resultados alcançados conclui-se que as iniciativas de abertura de dados abertos governamentais carecem de políticas de dados mais amplas e efetivas, com muitas iniciativas não abordando dimensões importantes para a qualidade dos dados.

Adicionalmente, está faltando comprometimento político para colocar em prática os planos de ação formalizados, observando-se pouca ou nenhuma transparência na execução das ações. Finalmente, verificou-se pouco envolvimento social na elaboração das políticas de dados, além de uma interação social ainda muito tímida nas atividades ofertadas pelos portais.

Essa conjuntura refletiu nos resultados de avaliação das iniciativas de abertura de dados governamentais que tiveram resultados medianos no diagnóstico, provocando decepção por se tratar, em sua maioria, de iniciativas pioneiras no movimento aberto. O portal de dados abertos do governo é a principal porta de entrada para o universo dos dados em âmbito nacional, tendo a responsabilidade de transparecer os principais benefícios e impactos da reutilização dos dados, e promover um maior envolvimento da sociedade com os dados.

A maneira como os governos lidam com a reutilização dos dados não permite a melhoria de processos internos, prejudicando a capacidade transformadora dos dados e comprometendo a eficiência governamental. Essa situação institucional reflete na dificuldade operacional que os editores do governo enfrentam ao publicarem dados com maior valor semântico (com o uso de padrões de metadados, vocabulários controlados, modelos ontológicos e aderência ao LOD), ilustrando bem os desafios que ainda precisam ser dirimidos.

Tais práticas teóricas e metodológicas tanto da CI quanto da CC viabilizam possíveis estratégias na busca de soluções inteligentes endereçadas à descrição de recursos de informação para a Web (MACHADO; SOUZA; DA GRAÇA SIMÕES, 2019; LEMOS; SOUZA, 2020a). A valorização no uso de padrões para a descrição de recursos digitais em nível de estrutura, valor, conteúdo e comunicação de dados (GILLILAND, 2016) reduz problemas ligados à normalização, à qualidade e a troca de descrições que em certa medida poderiam se valer das potencialidades da Web Semântica e do LOD (MACHADO; SOUZA; DA GRAÇA SIMÕES, 2019). De certa maneira, o que se pode considerar como dados de qualidade oriundos de um processo de curadoria digital são os descritos por esses princípios.

Nesse sentido, o emprego de sistemas de organização do conhecimento, a exemplo de taxonomias, tesouros, e até mesmo artefatos de representação mais sofisticados como ontologias, deveria ser considerado como pré-requisito nas ações de organização da informação, conforme advogam Lemos e Souza (2020a), propiciando, assim, enriquecimento semântico aos metadados, organização mais significativa das bases do governo e, conseqüentemente, melhorias nos sistemas de recuperação da informação que buscam se alinhar com os preceitos cinco estrelas do padrão LOD. Logo, tal prática ajudaria a amenizar um dos maiores problemas para dados abertos governamentais na quinta estrela, qual seja o uso de vocabulários que permitam relacionar dados de diversas bases.

Os dados devem ser posicionados dentro da administração pública como recursos imprescindíveis para a boa gestão governamental, portanto, as ações para se atingir esse intento devem ser claras e monitoradas para que obtenham real efetividade. Ao tornar os dados em ativos estratégicos para a tomada de decisão, permite-se o fomento da cultura de dados, com um ganho de eficiência a partir do uso de recursos informacionais em evidências. Com uma gestão mais capacitada na cultura de dados, espera-se o desenvolvimento de ações mais proativas na ativação de novos negócios orientados à reutilização dos dados, o que é salutar para toda a dinâmica de inovação nacional.

Dessa forma, gera-se capacidade de prover sustentabilidade econômica a iniciativa de dados através de reinvestimentos sistemáticos, beneficiando os setores que necessitam trabalhar com dados abertos governamentais de alta qualidade. Logo, todo o ciclo seria beneficiado.

Portanto, é proposto na próxima seção um modelo de governança visando publicar dados mais qualificados em rede, guiando a governança de dados na proposição de políticas mais abrangentes e eficazes, e mitigando as barreiras impostas a disponibilização dos dados.

Espera-se com o modelo garantir que os indivíduos tirem o melhor proveito dos dados abertos publicados nos portais governamentais, permitindo que se cumpra os benefícios esperados pela abertura de dados abertos do governo, quais sejam: o compartilhamento de conhecimento e ampliação democrática, a responsabilidade, a transparência, a eficiência governamental, a inovação, o combate a corrupção, o empoderamento cívico, além de incentivar o uso e a reutilização da informação para as mais diversas aplicações.

5.2 APRESENTAÇÃO DO MODELO DE GOVERNANÇA

A partir dos resultados da aplicação do diagnóstico de maturidade para dados abertos nos portais avaliados foi possível realizar inferências e interpretações necessárias à proposição do modelo de governança voltado à qualidade de dados para publicação em rede. As iniciativas de abertura que pretendem obter um grau elevado de maturidade na publicação de dados em rede precisam considerar três elementos na elaboração de suas políticas, estratégias e atividades a saber: i) transparência na execução das ações (WORLD WIDE WEB FOUNDATION, 2018; DIGITAL NATIONS, 2019); ii) divulgação dos impactos (WORLD WIDE WEB FOUNDATION, 2018); e iii) envolvimento social (DAVIES, 2010). A Figura 6 apresenta como os elementos se organizam:

Figura 6 - Elementos relacionados a uma publicação de dados com qualidade



Fonte: Elaborado pelo autor (2021) a partir de dados obtidos em agosto de 2021.

Abaixo segue a descrição dos três elementos:

- **transparência na execução das ações:** as ações determinadas pelo instrumento de Política de Dados, bem como, as práticas de responsabilidade do portal demandam monitoramento sistemático durante a sua execução, sendo o portal o principal instrumento de transparência. Espera-se que o monitoramento aponte a área do serviço público com falhas no cumprimento das metas e os respectivos motivos,

para que dessa forma a própria administração pública e a sociedade possam acompanhar e cobrar esclarecimentos e melhorias;

- **divulgação dos benefícios e impactos:** todos os benefícios e impactos advindos da abertura de dados abertos pelo governo (sociais, políticos, econômicos, ambientais, educacionais etc.) devem ser periodicamente compartilhados no portal de dados. O objetivo é aumentar a percepção acerca das benesses da abertura de dados para a sociedade. O ganho de confiança para o processo de abertura acarretaria o desenvolvimento do setor tecnológico dentro do governo, ativação de novos negócios para a economia, fomento à cultura de dados, maior capacitação dos estudantes e entusiastas, dentre outros;
- **envolvimento social:** os elementos que compõem a sociedade, (Organizações não governamentais, academia, pesquisadores, ativistas, estudantes, setor privado, e qualquer outro indivíduo) devem estar em constante envolvimento e diálogo com a governança de dados para elaboração de ações mais eficazes, de forma que a política de dados se torne um mecanismo mais abrangente e diversificado para contemplar todos os elementos da sociedade. O portal deve ser um instrumento fundamental no provisionamento de alguns elementos de interação com o público.

Logo, o conjunto de atividades relacionadas ao modelo de governança precisa considerar esses três elementos na sua elaboração, constituindo-se um instrumento mais eficiente na condução da gestão dos dados pela governança responsável. Com relação ao modelo de governança, a administração pública precisa estabelecer e operacionalizar um Órgão de Governança de Dados Abertos (OGDA), que terá a incumbência de colocar em prática as ações estabelecidas. As práticas aqui apresentadas devem ser observadas e gerenciadas pelo OGDA no esforço de fazer cumprir os objetivos de uma publicação de dados abertos governamentais qualificados em rede beneficiando a sua reutilização para as mais variadas aplicações. O Quadro 25 apresenta o conjunto de atividades do modelo, divididas em 8 temas oriundos da definição das categorias de análise, perfazendo 29 ações ao todo. Para cada ação estão definidas a responsabilidade e a técnica de modo que o órgão possa adotá-las visando a execução de cada prática.

Quadro 25 - Conjunto de atividades relacionadas ao modelo de governança proposto

(continua)

Tema 1: Infraestrutura em TI			
Escopo	Adequar as TIC's da organização para que sustentem as aplicações e os sistemas de informação responsáveis por prover capacidade ao portal de dados abertos.		
Item	Ação	Técnica	Responsabilidade
1.1	Implantar uma plataforma de gerenciamento de dados gratuita, de código aberto, e com uma comunidade de prática atuante (ex.: CKAN, Sonata), definindo os critérios para a seleção da plataforma que melhor se adeque à realidade da administração.	Plano Diretor de Tecnologia da Informação (PDTI)	OGDA e unidade de TIC
1.2	Adicionar informações de proveniência nas páginas dos <i>datasets</i> , como: data de publicação original, data da última atualização, editor responsável, tempo de atraso entre a publicação dos dados na sua fonte original até a atualização no portal, dentre outras informações relevantes.	PDTI e auditorias	OGDA e unidade de TIC
1.3	Disponibilizar monitores <i>online</i> para acompanhamento da disponibilidade dos servidores que hospedam o portal de dados abertos.	PDTI	OGDA e unidade de TIC
Tema 2: Acessibilidade e Usabilidade			
Escopo	Garantir que qualquer indivíduo possa acessar e usar os dados, além de oferecer recursos que permitam a navegação pelo portal sem obstrução, com o emprego de uma linguagem acessível em favor de uma experiência plena de uso.		
Item	Ação	Técnica	Responsabilidade
2.1	Criar um projeto de designer gráfico padronizado e consistente para todo o portal, com desejável compatibilidade com o domínio governamental principal.	Políticas de acessibilidade	OGDA e unidade de TIC
2.2	Oferecer no portal uma navegação destacada, e implantar recursos abrangentes e sofisticados de localização dos <i>datasets</i> .	Políticas de acessibilidade	OGDA e unidade de TIC
2.3	Melhorar a apresentação dos dados através de pré-visualização e gráficos, e implementar recursos variados de acesso aos <i>datasets</i> (formatos variados, API's e coleções de dados), pluralizando o uso do portal.	Políticas de acessibilidade	OGDA e unidade de TIC

Quadro 25 - Conjunto de atividades relacionadas ao modelo de governança proposto

(continua)

Tema 2: Acessibilidade e Usabilidade			
Item	Ação	Técnica	Responsabilidade
2.4	Discutir a participação cidadã na interação com os <i>datasets</i> , devendo ser oferecido um canal de ouvidoria para denúncias de erros no portal, e a possibilidade dos <i>datasets</i> serem avaliados para melhoria da qualidade e posterior responsabilização dos órgãos públicos na prestação do serviço.	Pesquisas, consultas públicas e ouvidoria	OGDA, unidade de TIC e sociedade
Tema 3: Curadoria e Publicação de Dados			
Escopo	Definir na organização a autoridade que irá zelar e cuidar do repositório de dados, além de garantir que os <i>datasets</i> sejam disponibilizados em conformidade com as melhores práticas de publicação e qualidade nos dados.		
Item	Ação	Técnica	Responsabilidade
3.1	Expandir a abrangência da política de dados para as esferas federal, regional e local, fazendo com que os órgãos públicos sejam incentivados a participar da iniciativa de abertura, compartilhando os mesmos princípios de publicação de dados objetivando melhorar a padronização e a interoperabilidade.	Estruturas legais e regulamentações	OGDA, e toda a sociedade
3.2	Promover intercâmbio técnico entre as unidades de publicação federal, e as demais unidades regionais e locais. A capacidade técnica deve ser compartilhada e reaproveitada em outros órgãos que desejam publicar dados e que carecem de recursos em comparação com a infraestrutura federal.	Códigos de conduta de governança, normas e valores culturais	OGDA, unidade de TIC, pesquisadores / academia e profissionais
3.3	Utilizar metadados, ontologias, vocabulários controlados, modelos de metadados (ex.: <i>Dublin Core</i> , <i>DCAT</i> , <i>Common Data Model</i>) buscando-se, nesses artefatos, uma representação semântica de dados abertos a serem publicados. O uso desses artefatos de representação deve ser normatizado, com toda documentação técnica bem descrita e publicada no portal de dados abertos para fins de consulta pública.	Organização e tratamento da informação; Documentação técnica	OGDA, Unidade de TIC

Quadro 25 - Conjunto de atividades relacionadas ao modelo de governança proposto

(continua)

Tema 3: Curadoria e Publicação de Dados			
Item	Ação	Técnica	Responsabilidade
3.4	Formalizar a necessidade de publicação de dados em aderência ao padrão LOD, esclarecendo e justificando seu uso nos documentos técnicos, planos e estratégias de dados abertos. Sugere-se integrar os princípios LOD aos planos de ação da governança de dados estabelecendo metas de conformidade ao padrão, de médio e longo prazo, e incentivando a alocação de recursos para pesquisa na área.	Normas e diretrizes para publicação de dados	OGDA e unidade de TIC
3.5	Discutir um maior envolvimento cívico no portal de dados abertos, devendo-se considerar a publicação de dados abertos por qualquer indivíduo.	Pesquisas, debates e chamadas públicas	OGDA, e toda a sociedade
Tema 4: Político e Social			
Escopo	Identificar e disponibilizar os <i>datasets</i> de maior impacto social, utilizando o portal de dados abertos como instrumento de captação dos anseios da sociedade, vitrine para divulgação dos impactos dos dados e monitoramento das ações de implementação dos dados abertos pelas agências governamentais.		
Item	Ação	Técnica	Responsabilidade
4.1	Diagnosticar quais dados governamentais devem ter prioridade de abertura pela administração pública.	Pesquisas e consultas públicas	OGDA, e toda a sociedade
4.2	Apresentar no portal de dados exemplos de impactos a partir da reutilização dos dados no campo político e social.	Política de auditoria	OGDA
4.3	Criar conselhos que monitorem a atuação das agências na publicação dos dados, fiscalizando o andamento da abertura.	Governança compartilhada	OGDA
4.4	Disponibilizar no portal de dados abertos painéis que monitorem o quantitativo de órgãos e agências que já abriram seus dados.	Políticas de tecnologia e auditorias	OGDA, unidade de TIC e publicadores
4.5	Realizar iniciativas que promovam o reuso dos dados abertos governamentais em benefício de comunidades marginalizadas, fazendo-o pelo menos de forma anual, e estabelecendo parcerias com outros setores da sociedade.	Concursos públicos e <i>hackathons</i>	OGDA, e toda a sociedade

Quadro 25 - Conjunto de atividades relacionadas ao modelo de governança proposto

(continua)

Tema 5: Econômico			
Escopo	Valorizar as iniciativas que permitam a viabilidade econômica e sustentável do portal de dados abertos, incluindo a promoção da participação de atores da iniciativa pública e privada e concomitantemente apontando os benefícios econômicos da reutilização dos dados para a melhoria do serviço público em geral.		
Item	Ação	Técnica	Responsabilidade
5.1	Criar e apresentar no plano de ação da governança de dados as iniciativas para ativação de novos negócios a partir do reuso dos dados abertos.	Estratégias de negócio	OGDA, profissionais, e terceiro setor
5.2	Catalogar e apresentar no portal de dados os exemplos que já estão ativando crescimento econômico no país a partir do reuso dos dados abertos.	Avaliações, classificações e auditoria	OGDA, profissionais, e terceiro setor
5.3	Promover periodicamente iniciativas de reutilização de dados com envolvimento de toda a sociedade para que aplicações inovadoras sejam desenvolvidas.	Concursos públicos e <i>hackathons</i>	OGDA e toda a sociedade
Tema 6: Organização e Processos Internos			
Escopo	Criar uma estratégia de gerenciamento, manutenção e uso dos dados dentro do governo constituindo esses dados em ativos estratégicos para a melhor tomada de decisão e fomentando uma cultura de dados na administração.		
Item	Ação	Técnica	Responsabilidade
6.1	Criar e apresentar no plano de ação da governança de dados as iniciativas para desenvolver uma cultura de dados institucional, bem como, o reuso dos dados como ativos estratégicos para a administração pública.	Estratégias de governança, normas e valores culturais	OGDA
6.2	Implantar monitores no portal de dados abertos com intuito de divulgar o andamento da implantação de todas as ações listadas na Política de Dados.	Políticas de tecnologia e auditorias	OGDA, unidade de TIC e publicadores
6.3	Catalogar e apresentar no portal de dados os exemplos que já estão transformando a administração pública a partir do reuso dos dados abertos.	Avaliações, classificações e auditoria	OGDA e publicadores

Quadro 25 - Conjunto de atividades relacionadas ao modelo de governança proposto

(conclusão)

Tema 7: Aspectos Legais e de Privacidade			
Escopo	Definir com clareza os termos relacionados a privacidade e segurança em que os dados estão apoiados e o tipo de licença de uso nos <i>datasets</i> disponibilizados no portal.		
Item	Ação	Técnica	Responsabilidade
7.1	Disponibilizar em local estratégico no portal de dados abertos e de forma atualizada a legislação que regulamenta a abertura dos dados, além das licenças de uso por meio de metadados dessa natureza. Este conteúdo deve estar bem localizado e atualizado para dirimir as dúvidas sobre o uso dos dados, resguardando a administração pública.	Códigos de conduta de governança, normas e valores culturais	OGDA
7.2	Oferecer no portal de dados abertos um canal de ouvidoria para denúncias de problemas nos <i>datasets</i> que estejam ferindo as questões de privacidade e demais aspectos legais estipulados pelo ordenamento jurídico local.	Ouvidoria	OGDA, unidade de TIC
Tema 8: Capacitação e Treinamento			
Escopo	Fomentar a cultura de dados entre a sociedade, além de aprimorar as competências de dados em alunos na rede de ensino e dos funcionários governamentais que trabalham diretamente com essa disciplina.		
Item	Ação	Técnica	Responsabilidade
8.1	Incluir no plano de ação da governança de dados iniciativas para melhor prover a capacitação e treinamento dos servidores do governo.	Estratégias de governança, normas e valores culturais	OGDA
8.2	Incluir no plano de ação da governança de dados iniciativas para melhor prover a capacitação de alunos da rede de ensino público e privada em competência de dados.	Estratégias de governança	OGDA, pesquisadores, academia, e sociedade
8.3	Promover e divulgar no portal de dados abertos eventos sobre a temática de dados abertos governamentais.	Política de governança	OGDA e toda a sociedade
8.4	Disponibilizar no portal de dados uma área específica para a capacitação de qualquer indivíduo em competência de dados abertos.	Cursos <i>online</i> , tutoriais e <i>softwares</i>	OGDA e unidade de TIC

Fonte: Elaborado pelo autor (2021).

O conjunto de atividades é fruto do levantamento bibliográfico e documental, um recurso de pré-análise que redundou nos indicadores e variáveis de requisitos, exploração, tratamento dos resultados do diagnóstico, a inferência e, finalmente, a interpretação das evidências. As ações apresentadas foram inferidas a partir do uso do método de análise de conteúdo e procuraram estabelecer relações com os três elementos ilustrados pela Figura 6.

Recomenda-se que o prazo para cumprimento de cada uma das ações também esteja contemplado no Quadro 25, porém não foram inseridos por vincularem-se a variáveis dependentes à realidade de cada governo. No entanto, espera-se celeridade para o cumprimento dos prazos.

Tanto as atividades listadas no Quadro 25, quanto o modelo de governança voltado a qualidade de dados para publicação em rede, devem estar devidamente formalizados nos documentos relacionados ao plano de ação da política de dados, compondo, dessa forma, a estratégia competitiva do órgão de governança responsável pelo gerenciamento de dados dentro da administração pública.

A Figura 7 apresenta o infográfico da proposta do modelo de governança voltado a qualidade de dados para publicação em rede, com a descrição de como se comporta seu fluxo operacional para que se atinja os objetivos de publicação aspirados.

A execução e o controle do modelo são de responsabilidade da OGDA, que precisa definir de forma prévia a visão da governança de dados a partir dos valores que inspiram a administração pública dentro da qual está submetida.

O modelo de governança compõe a estratégia para qualidade na publicação dos dados, contribuindo para que a OGDA cumpra as suas atribuições de identificar, gerenciar, avaliar, supervisionar e comunicar os dados, além de gerenciar os problemas relacionados a governança de dados.

Figura 7 - Modelo de governança proposto



Fonte: Elaborado pelo autor (2021).

Na estratégia estipulada pela governança, é necessário primeiramente o estabelecimento de um OGDA para o gerenciamento dos dados. Caberá ao órgão definir a missão, o valor, o escopo, as atividades, as ações, as técnicas, as metas e as responsabilidades que irão nortear o processo de gestão. Os deveres que o OGDA possui na gestão dos dados são: identificação, gerenciamento, avaliação, supervisão e comunicação dos dados, além do gerenciamento dos problemas que podem vir a ocorrer ao longo do processo.

O órgão precisará conduzir um conjunto de atividades para poder cumprir seus deveres. O modelo proposto contempla um conjunto de 29 ações, distribuídas em 8 temas que exploram a natureza complexa e heterogênea do escopo da gestão dos dados. Todas as ações estão baseadas no atendimento aos três pilares inferidos no processo metodológico: i) Transparência na execução das ações; ii) Envolvimento Social; e iii) Divulgação dos benefícios e impactos. O objetivo geral do modelo é claro: publicar dados abertos com qualidade em rede.

Aplicando o modelo de governança, o OGDA poderá atender aos anseios da administração pública e concomitantemente justificará a existência de sua criação. O resultado do modelo permitirá a disponibilização de dados com qualidade semântica, provendo um conteúdo com metadados mais bem organizados e representados, preservando assim o significado de um conteúdo tão essencial para a sociedade.

Além da questão da qualidade semântica com o uso de metadados, ontologias, vocabulários controlados e padrões de anotação, a execução das ações servirá para aprimorar toda a infraestrutura tecnológica que circunda a publicação dos dados e que reflete em sua qualidade.

No aspecto da organização e dos processos internos, espera-se com sua implantação um ganho na transparência administrativa e organizacional, com a redução da influência concentrada apenas sob uma pequena parcela de servidores da administração. A característica abrangente do modelo terá grande serventia na capacitação de servidores públicos e no cumprimento das metas de qualidade para publicação de dados em rede. As barreiras internas à administração podem ser enfrentadas de forma mais adequada a partir do efetivo empoderamento de servidores públicos e de gestão da reutilização dos dados abertos de forma estratégica dentro das instituições.

O grande envolvimento social, que é um dos pilares do modelo proposto, permitirá que o governo dialogue mais com a sociedade em um esforço conjunto para

o estabelecimento de políticas de dados mais abrangentes, inclusivas e que façam mais sentido com o movimento de dados abertos.

Espera-se que o modelo viabilize práticas orientadas à superação de barreiras críticas, com o provimento de maior acesso democrático aos dados, promoção de benefícios sociais a partir da reutilização dos dados, ativação de novos negócios para a economia, promoção de dados como verdadeiros ativos estratégicos dentro da administração e o fomento da cultura de dados em benefício de toda a sociedade.

Finalmente, muitas das ações promovidas pelo modelo possuem a característica de monitorar a execução das atividades governamentais contidas nas políticas, o que aumenta a transparência dessas ações. A divulgação dos impactos e benefícios, de forma atualizada, ajudará na percepção da real efetividade do conjunto de ações propostas pelo modelo: uma omissão nesse tipo de informação pode ser um forte indicador de algum problema em algum dos oito temas, levando a questionamentos a própria iniciativa de abertura de dados.

O próximo passo a ser consolidado no movimento de abertura é o maior envolvimento de qualquer pessoa ou organização na publicação de *datasets* através do portal de dados do governo. Países como a França já aceitam a publicação de *datasets* por parte de qualquer indivíduo, desde que sejam informações de utilidade pública (OECD, 2020; FRANÇA, 2021). O modelo proposto já contempla essa tendência, contribuindo para o fortalecimento do movimento de abertura de dados.

Sobre o engajamento cívico abordado pelo modelo, este exercerá um maior controle social sobre o governo, o que garantirá a discussão e reflexão acerca da abertura de dados. Com efeito, tal elemento permite aprimoramento contínuo do modelo de governança, produção de impacto positivo na atualização sistemática da política de dados e utilização dinâmica da política de dados abertos em rede como um instrumento representativo de boa governança em uma sociedade democrática.

6 DISCUSSÃO DOS RESULTADOS

A elaboração do modelo foi fundamentada nos resultados e inferências consolidadas oriundas de uma série de procedimentos metodológicos conduzidos durante as aplicações das técnicas de análise de conteúdo.

Em seu início, a pesquisa apresentou os antecedentes e uma abordagem histórica que subjaz o movimento de dados abertos em vários países levando-se a um panorama atual de esforços locais, evidenciando estágios e contribuições relevantes às discussões sobre desafios e oportunidades para os governos na busca de abertura de dados à sociedade contemporânea. A partir desse estado da arte foi possível sistematizar um conjunto de desafios e oportunidades mapeados nas ações de abertura de dados governamentais em alguns países, sendo essa contribuição reconhecida como insumo a ser trabalhada na fase inicial do método.

Na fase inicial, conhecida como pré-análise, foi conduzido um experimento de pré-teste nos portais de dados abertos governamentais dos EUA, Reino Unido, Brasil, e Irlanda, com objetivo de se obter melhor compreensão da realidade da governança de dados em iniciativas de abertura. Todo o material coletado tanto na pesquisa bibliográfica e documental, quanto nas fontes documentais dos portais investigados foram coletados e analisados a luz dos desafios levantados pelo estado da arte. Como resultado obteve-se a primeira grande entrega da pesquisa: a definição das categorias de análise. Foi uma etapa especialmente importante para conhecer a realidade de como se apresenta na prática a disponibilização de dados governamentais em rede.

Na fase intermediária, responsável pela exploração do material coletado tanto na pesquisa bibliográfica e documental, quanto nas fontes documentais dos portais elencados para o pré-teste, chegou-se a mais uma entrega da pesquisa: a elaboração de um diagnóstico de maturidade para avaliação de portais de dados abertos. Posteriormente, o diagnóstico foi validado nos quatro portais de dados abertos já investigados no pré-teste. Os dados advindos da avaliação foram submetidos às técnicas de análise de conteúdo (BARDIN, 2016), com suas conclusões semânticas sendo expostas através de gráficos comparativos e discussões. Ao final, chegou-se a importantes resultados para a condução do último procedimento metodológico.

Na etapa final do método foi realizado o tratamento dos resultados, a inferência e a interpretação, podendo-se elaborar a última entrega da pesquisa: o modelo de governança voltado à qualidade de dados para publicação em rede, além das

atividades relacionadas. O modelo emerge como um instrumento para atender os objetivos de governo e do movimento dados abertos, que anseiam por uma publicação de dados governamentais em rede com qualidade. De forma subjacente, a contribuição do modelo de governança poderá ser estendida para a elaboração de políticas de dados mais representativas e eficientes.

Faz-se necessária uma política de dados abrangente, especialmente quando se observa as diversas dimensões que envolvem uma publicação de dados em rede com qualidade. No entanto, uma avaliação de qualidade para publicações em rede deve olhar com restrição para as políticas: não basta um documento bem redigido e com ações bem assertivas, se persiste nas ações propostas a ausência de envolvimento com os diversos atores que compõem a sociedade. Mesmo que a política esteja repleta de boas intenções, ela não terá muita utilidade se não estabelecer em seu corpo uma governança de dados focada em atingir os objetivos da administração. Ainda que seja uma política abrangente e com uma governança de dados exemplar, não terá serventia se as ações não puderem ser monitoradas e percebidas pela sociedade, com o detalhamento do impacto da reutilização dos dados sendo esclarecido no portal. O modelo de governança proposto visa preencher essas lacunas.

Portanto, a presente dissertação cumpre seu papel em responder as questões de pesquisa, evidenciando no estado da arte como os governos mundiais estão conduzindo práticas de publicação de dados abertos e as barreiras que estão enfrentando, e finalmente, propondo como produto de dissertação um modelo de governança que visa mitigar tais desafios impostos aos governos na publicação de dados mais qualificados em rede.

O objetivo geral também foi cumprido, com a apresentação do modelo de governança devidamente fundamentado, acompanhado de seu conjunto de atividades. Os artefatos irão guiar a governança ao melhor gerenciamento da sua infraestrutura de dados, permitindo que se cumpra uma publicação mais qualificada de dados em rede. Os objetivos específicos também foram todos atingidos, através do estado da arte, dos pré-testes, e das discussões advindas a partir da validação do modelo de diagnóstico de maturidade para portais.

7 CONCLUSÕES E TRABALHOS FUTUROS

Apresentou-se um estudo do estado da arte contendo os antecedentes do movimento de abertura de dados, desde as discussões científicas até a propagação da abertura de dados pelos governos através de seus portais. Também foi externado um panorama atual do movimento, com informações sobre o progresso da abertura de dados governamentais destacando o Brasil, englobando os cinco continentes, além de indicações de abertura que são tendências pelo mundo.

Através da dissertação também foi possível entender a contribuição ao fomento da representação semântica de dados governamentais, ampliando a compreensão sobre a necessidade do uso de políticas e boas práticas de catalogação de dados, metadados, ontologias, vocabulários controlados e modelos conceituais de acordo com cada especificidade de abertura de dados governamentais, colocando em destaque as contribuições das Ciências da Informação e da Computação na organização e representação da informação e do conhecimento em contextos digitais objetivando aumentar o escopo da interoperabilidade e a integração plena dos conjuntos de dados disponíveis na Web.

A contribuição subsequente foi a sistematização de um conjunto de desafios e oportunidades, com a entrega de uma linha do tempo contendo os principais marcos de abertura no Brasil e no mundo. A literatura costuma abordar de forma sistemática as barreiras, porém, de forma isolada ou em grupos correlatos, sendo incomum abarcar indicadores tão abrangentes como os apresentados pela pesquisa. Essa importante inferência foi essencial para o desdobramento da pesquisa, constituindo-se como um item chave para a melhor compreensão das características e desafios que os governos vêm enfrentando para abrir seus dados em rede e cumprir com os preceitos estabelecidos pelo movimento aberto.

A partir dos indicadores, foi possível elaborar um modelo de governança voltado à qualidade de dados para publicação em rede que conseguiu capturar a realidade e os desafios de publicação pelos quais os governos mundiais atravessam na abertura de seus dados. Abordando aspectos tão amplos, o modelo permite que a administração pública, conjuntamente com a participação da sociedade civil e OGDA, elabore políticas de dados mais eficientes, abrangentes e transparentes.

De forma consonante, será possível capacitar os desenvolvedores e técnicos a publicarem os dados na rede com mais qualidade a partir do uso eficaz das ações

propostas pelo modelo de governança, também servindo para encorajar aqueles servidores que atuam nessa atividade e que queiram iniciar uma cultura de publicação de dados com qualidade dentro de seus órgãos.

Ao permitir a publicação de dados com qualidade, o modelo contribuirá com objetivos primários do movimento de abertura de dados, quais sejam, o compartilhamento de conhecimento e a ampliação democrática de acesso a esse conhecimento beneficiando tanto a sociedade quanto o governo no aspecto de acessibilidade a fontes de dados, aumentando a responsabilidade, a eficiência governamental, a transparência, a inovação, o combate a corrupção e o empoderamento cívico.

O modelo de governança proposto tem como premissa respeitar as dimensões e as capacidades que cada governo possui no enfrentamento dos desafios de abertura de dados. No entanto, o modelo está limitado a aplicação apenas em estados ou nações verdadeiramente democráticas, nas quais os indivíduos possam ter a plena capacidade de se expressar, opinar e interagir de forma cidadã com os governos, sem que para isso tenham cerceadas suas garantias e liberdades individuais.

Como trabalho futuro, o modelo será disponibilizado após a defesa da dissertação em um repositório *online* aberto ao público denominado *GitHub*, para livre acesso e reutilização por comunidades interessadas na temática de dados abertos governamentais. O modelo estará comentado e com componentes de hipertexto, além de acompanhar um manual técnico para aplicação.

Espera-se que a partir da reutilização do modelo, o mesmo possa ser discutido, expandido e constantemente aperfeiçoado por outros pesquisadores, dentre os quais profissionais da área acadêmica, ativistas, sociedade civil organizada, terceiro setor, estudantes, editores de dados, servidores e gestores da área pública, e especialmente os responsáveis pela gestão de dados abertos governamentais nas agências e órgãos públicos, de âmbito global.

REFERÊNCIAS

- AFFUL-DADZIE, E.; AFFUL-DADZIE, A. Open Government Data in Africa: a preference elicitation analysis of media practitioners. **Government Information Quarterly**, v. 34, n. 2, p. 244–255, 2017. Disponível em: <http://dx.doi.org/10.1016/j.giq.2017.02.005>. Acesso em: 10 jun. 2020.
- ALTAYAR, M. S. Motivations for open data adoption: An institutional theory perspective. **Government Information Quarterly**, v. 35, n. 4, p. 633–643, 2018. Disponível em: <https://doi.org/10.1016/j.giq.2018.09.006>. Acesso em: 10 jun. 2020.
- ALMEIDA, M. B. Uma abordagem integrada sobre ontologias: Ciência da Informação, Ciência da Computação e Filosofia. **Perspectivas em Ciência da Informação**, v. 19, n. 3, p. 242–258, set. 2014. Disponível em: <http://dx.doi.org/10.1590/1981-5344/1736>. Acesso em: 25 mar. 2021.
- ANNE, K. M.; CARLISLE, T.; DOMBROWSKI, Q.; GLASS, E.; GNIADY, T.; JONES, J. B. *et al.* Building Capacity for Digital Humanities: A Framework for Institutional Planning. **ECAR working group paper**. Louisville, CO: ECAR, v. 31, 2017.
- ANSI/NISO Z39.19-2005 (R2010). **Guidelines for the construction, format, and management of monolingual controlled vocabularies**. Baltimore: NISO Press, 2005.
- ATTARD, J.; ORLANDI, F.; SCERRI, S.; AUER, S. A systematic review of open government data initiatives. **Government Information Quarterly**, v. 32, n. 4, p. 399–418, 2015. Disponível em: <http://dx.doi.org/10.1016/j.giq.2015.07.006>. Acesso em: 10 jun. 2020.
- AUER, S. Introduction to LOD2. *In: Linked Open Data - Creating Knowledge Out of Interlinked Data*, [S.l.]: Springer, Cham, 2014. p. 1–17.
- BARDIN, L. **Análise de conteúdo**. São Paulo: Edições 70, 2016.
- BAUER, F.; KALTENBÖCK, M. **Linked open data: the essentials**. Vienna: Edition mono/monochrome, 2011.
- BERNERS-LEE, T. **Linked data principles**. 2006. Disponível em: <http://www.w3.org/DesignIssues/LinkedData.html>. Acesso em: 11 nov. 2019.
- BERNERS-LEE, T.; HENDLER, J.; LASSILA, O. The semantic web. **Scientific American**, v. 284, n. 5, p. 34–43, maio 2001. Disponível em: <http://www2.ic.uff.br/~bazilio/cursos/sistweb/material/Barners-Lee-Scientific-American-May-2001.pdf>. Acesso em: 15 mar. 2019.
- BIZER, C.; HEATH, T. **Linked data: evolving the web into a global data space**. [S.l.]: Morgan & Claypool, 2011. v. 1. (Synthesis Lectures on the Semantic Web: Theory and Technology).

BIZER, C.; HEATH, T.; BERNERS-LEE, T. Linked data - the story so far. **International Journal on Semantic Web and Information Systems**, [S.l.], v. 5, n. 3, p. 1–22, 2009.

BLANK, M. **Open Data Maturity Report 2019**, 2019. Disponível em: https://www.europeandataportal.eu/sites/default/files/open_data_maturity_report_2019.pdf. Acesso em: 10 nov. 2019.

BRASIL. Controladoria-Geral da União. **Plano de Dados Abertos vigência junho/2021 a junho/2023**, 2021a. Disponível em: <https://www.gov.br/cgu/pt-br/aceso-a-informacao/dados-abertos/plano-de-dados-abertos-da-cgu/arquivos/plano-de-dados-abertos-2018-dezembro.pdf>. Acesso em: 17 jul. 2021.

BRASIL. Controladoria-Geral da União. **Portal Brasileiro de Dados Abertos**, 2021b. Disponível em: <http://dados.gov.br/>. Acesso em: 02 maio 2021.

BRASIL. Controladoria-Geral da União. **Painel Monitoramento de Dados Abertos**, 2021c. Disponível em: <http://paineis.cgu.gov.br/dadosabertos/index.htm>. Acesso em: 21 ago. 2021.

BRASIL. Controladoria-Geral da União. **Infraestrutura Nacional de Dados Abertos (INDA)**, 2021d. Disponível em: <https://wiki.dados.gov.br/>. Acesso em: 10 maio 2021.

BRASIL. Controladoria-Geral da União. **Fala.BR - Plataforma Integrada de Ouvidoria e Acesso à Informação**, 2021e. Disponível em: <https://falabr.cgu.gov.br/publico/Manifestacao/SelecionarTipoManifestacao.aspx?ReturnUrl=%2f>. Acesso em: 10 ago. 2021.

BRASIL. Comitê Gestor da Infraestrutura Nacional de Dados Abertos. **Plano de ação 2021-2022**, 2021f. Disponível em: <https://dados.gov.br/wp/wp-content/uploads/2021/03/Plano-de-A%c3%a7%c3%a3o-INDA-vf.pdf>. Acesso em: 10 ago. 2021.

BRASIL. Programa de Governo Eletrônico Brasileiro. **Padrões de Interoperabilidade de Governo Eletrônico - ePING**, 2018. Disponível em: <http://eping.governoeletronico.gov.br/>. Acesso em: 12 maio 2021.

BRASIL. TRIBUNAL DE CONTAS DA UNIÃO. **Cinco Motivos para a Abertura de Dados na Administração Pública**, 2015. Disponível em: <https://portal.tcu.gov.br/biblioteca-digital/cinco-motivos-para-a-abertura-de-dados-na-administracao-publica.htm>. Acesso em: 11 jul. 2020.

CAMPOS, M. L. DE A. Modelização de domínios de conhecimento: uma investigação de princípios fundamentais. **Ciência da Informação**, Brasília, v. 33, n. 1, p. 22–32, abr. 2004. Disponível em: <https://doi.org/10.1590/S0100-19652004000100003>. Acesso em: 26 set. 2019.

CHATFIELD, A. T.; REDDICK, C. G. The role of policy entrepreneurs in open government data policy innovation diffusion: an analysis of Australian Federal and State Governments. **Government Information Quarterly**, v. 35, n. 1, p. 123–134,

2018. Disponível em: <http://dx.doi.org/10.1016/j.giq.2017.10.004>. Acesso em: 10 jun. 2020.

CHIGNARD, S. A brief history of open data. **Paris Tech Review**, v. 29, 2013. Disponível em: <http://parisinnovationreview.com/articles-en/a-brief-history-of-open-data>. Acesso em: 13 jan. 2020.

CKAN. **Open Source Data Management System**, 2021. Disponível em: <https://ckan.org/>. Acesso em: 02 maio 2021.

CORDEIRO, K. DE F.; FARIA, F. F. DE.; PEREIRA, B. DE O.; FREITAS, A.; RIBEIRO, C. E.; FREITAS, J. V. V. B. *et al.* An approach for managing and semantically enriching the publication of Linked Open Governmental Data. *In: Proceedings of the 3rd workshop in applied computing for electronic government (WCGE)*, SBBD. 2011. p. 82-95. Disponível em: <https://www.researchgate.net/publication/284686112>. Acesso em: 15 nov. 2019.

CRUSOE, J. R.; AHLIN, K. Users' activities for using open government data—a process framework. **Transforming government: People, process and policy**, v. 13, n. 3/4, p. 213–236, 2019. Disponível em: <http://dx.doi.org/10.1108/TG-04-2019-0028>. Acesso em: 10 jun. 2020.

DAVIES, T. **Open data, democracy and public sector reform**, 2010. Disponível em: <http://www.opendataimpacts.net/report/>. Acesso em: 20 jan. 2020.

DAVIES, T.; BAWA, Z. A. The promises and perils of open government data (OGD). **The Journal of Community Informatics**, [S.l.], v. 8, n. 2, p. 1–6, 2012. Disponível em: <https://openjournals.uwaterloo.ca/index.php/JoCI/article/view/3035/3937>. Acesso em: 12 jun. 2020.

DAVIES, T.; FRANK, M. 'There's no such thing as raw data' exploring the socio-technical life of a government dataset. *In: Proceedings of the 5th annual ACM web science conference*, 2013. p. 75-78. Disponível em: <https://doi.org/10.1145/2464464.2464472>. Acesso em: 18 jun. 2020.

DAVIES, T.; PERINI, F. Researching the emerging impacts of open data: revisiting the ODDC conceptual framework. **The Journal of Community Informatics**, [S.l.], v. 12, n. 2, 2016. Disponível em: <https://doi.org/10.15353/joci.v12i2.3246>. Acesso em: 20 mar. 2021.

DIGITAL NATIONS. **Digital Nations Data 360° Declaration**, 2019. Disponível em: <https://www.canada.ca/en/government/system/digital-government/digital-nations/digital-nations-data-360-declaration.html>. Acesso em: 20 jan. 2020.

DING, L; PERISTERAS, V; HAUSENBLAS, M. Linked open government data [Guest editors' introduction]. **IEEE Intelligent systems**, [S.l.], v. 27, n. 3, p. 11–15, 2012.

DOERR, M. The CIDOC conceptual reference module: an ontological approach to semantic interoperability of metadata. **AI magazine**, [S.l.], v. 24, n. 3, p. 75–75, 2003. Disponível em: <https://doi.org/10.1609/aimag.v24i3.1720>. Acesso em: 24 out. 2019.

DUVAL, E.; HODGINS, W.; SUTTON, S.; WEIBEL, S. L. Metadata principles and practicalities. **D-lib Magazine**, [S.l.], v. 8, n. 4, p. 1082-9873, 2002. Disponível em: <https://www.dlib.org/dlib/april02/weibel/04weibel.html>. Acesso em: 12 out. 2019.

EUROPEAN UNION. **European data portal**, 2021. Disponível em: <https://data.europa.eu/en>. Acesso em: 03 mar. 2021.

ESTADOS UNIDOS. U.S. CONGRESS. **H.R.1770 - OPEN Government Data Act**, 2017. Disponível em: <https://www.congress.gov/bill/115th-congress/house-bill/1770>. Acesso em: 10 jul. 2021.

ESTADOS UNIDOS. FEDERAL DATA STRATEGY. **Data Governance Playbook**. 2020a. Disponível em: <https://resources.data.gov/assets/documents/fds-data-governance-playbook.pdf>. Acesso em: 22 maio 2021.

ESTADOS UNIDOS. FEDERAL DATA STRATEGY. **2020 Action Plan**. 2020b. Disponível em: <https://strategy.data.gov/action-plan/>. Acesso em: 10 jul. 2021.

ESTADOS UNIDOS. U.S. GENERAL SERVICES ADMINISTRATION. **U.S. Government's Open Data Portal**, 2021a. Disponível em: <https://www.data.gov/>. Acesso em: 10 maio 2021.

ESTADOS UNIDOS. OFFICE OF MANAGEMENT AND BUDGET. **Federal Enterprise Data Resources**, 2021b. Disponível em: <https://resources.data.gov/>. Acesso em: 10 maio 2021.

ESTADOS UNIDOS. OFFICE OF MANAGEMENT AND BUDGET. **Federal Data Strategy**, 2021c. Disponível em: <https://strategy.data.gov/>. Acesso em: 10 maio 2021.

ESTADOS UNIDOS. U.S. GENERAL SERVICES ADMINISTRATION. **Open Data Stack Exchange**, 2021d. Disponível em: <https://opendata.stackexchange.com/>. Acesso em: 10 jul. 2021.

ESTADOS UNIDOS. U.S. GENERAL SERVICES ADMINISTRATION. **Project Open Data Dashboard**, 2021e. Disponível em: <https://labs.data.gov/dashboard/offices/qa>. Acesso em: 10 jul. 2021.

FRANÇA. **Documentação da plataforma aberta para dados públicos franceses**, 2021. Disponível em: <https://doc.data.gouv.fr/a-propos/que-publier-et-comment-le-publier/>. Acesso em: 20 jul. 2021.

G8. **G8 Open Data Charter and Technical Annex**, 2013. Disponível em: <https://www.gov.uk/government/publications/open-data-charter/g8-open-data-charter-and-technical-annex>. Acesso em: 20 jan. 2020.

GIL, A. C. **Como elaborar projetos de pesquisa**. 6. ed. São Paulo: Atlas, 2019.

GILLILAND, A. J. Setting the Stage. In: BACA, M. (Ed.). **Introduction to metadata**. 3. ed. Los Angeles: Getty Research Institute, 2016.

GRAY, J. Towards a genealogy of open data. *In: General Conference of the European Consortium for Political Research*, Glasgow, 2014.

GUARINO, N (Ed.). **Formal ontology in information systems**: Proceedings of the first international conference (FOIS'98), June 6-8, Trento, Italy. IOS press, 1998.

GUIZZARDI, G. **Ontological foundations for structural conceptual models**. Capítulo 1 e 9.2. 2005. 416 f. Tese (Doutorado em Ciência da Computação) - Universidade de Twente, Enschede, Holanda, 2005.

GUIZZARDI, G.; FALBO, R. A.; GUIZZARDI, R. S. S. A importância de Ontologias de Fundamentação para a Engenharia de Ontologias de Domínio: o caso do domínio de Processos de Software. **IEEE Latin America Transactions**, [S.l.], v. 6, n. 3, p. 244–251, 2008.

HJØRLAND, B. Domain analysis in information science: eleven approaches traditional as well as innovative. **Journal of Documentation**, [S.l.], v. 58, n. 4, p. 422–463, 2002.

HJØRLAND, B. Fundamentals of knowledge organization. **Knowledge organization**, [S.l.], v. 30, n. 2, p. 87–111, 2003.

HJØRLAND, B. Semantics and knowledge organization. **Annual Review of Information Science and Technology**, [S.l.], p. 367–405, 2007.

HORVATH, A. 'Governance'—in crisis? A cross-disciplinary critical review of three decades of 'governance' scholarship. **Centre for Global Higher Education**, [S.l.], p. 1-181, 2017.

INICIATIVA LATINOAMERICANA POR LOS DATOS ABIERTOS. **Portal de la iniciativa latinoamericana por los datos abiertos**, 2021. Disponível em: <https://idatosabiertos.org/>. Acesso em: 03 mar. 2021.

INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS (IFLA). **Digital libraries**: metadata resources, 2005. Disponível em: <https://www.ifla.org/node/9337>. Acesso em: 29 fev. 2020.

IRLANDA. DEPARTMENT OF PUBLIC EXPENDITURE AND REFORM. **Open Data Strategy 2017-2022**, 2017. Disponível em: https://data.gov.ie/uploads/page_images/2018-03-07-114306.063816Final-Strategy-online-version1.pdf. Acesso em: 25 jul. 2021.

IRLANDA. DEPARTMENT OF PUBLIC EXPENDITURE AND REFORM. **Public Service Data Strategy 2019-2023**, 2019. Disponível em:

<https://assets.gov.ie/7107/7ac4ae109cd944dc99c9e178962fa095.pdf>. Acesso em: 25 jul. 2021.

IRLANDA. DEPARTMENT OF PUBLIC EXPENDITURE AND REFORM. **Ireland's Open Data Portal**, 2021a. Disponível em: <https://data.gov.ie/>. Acesso em: 22 maio 2021.

IRLANDA. DEPARTMENT OF PUBLIC EXPENDITURE AND REFORM. **DGI Front Page SLA report**, 2021b. Disponível em: <https://synthetics.newrelic.com/report/R2MvK?view=monthly-sla-report>. Acesso em: 22 jul. 2021.

IRLANDA. DEPARTMENT OF PUBLIC EXPENDITURE AND REFORM. **Measuring Open Data Impact**, 2021c. Disponível em: <https://ec.europa.eu/eusurvey/runner/opendatasurvey>. Acesso em: 22 jul. 2021.

IRLANDA. DEPARTMENT OF PUBLIC EXPENDITURE AND REFORM. **Open Data Publication Guidelines**, 2021d. Disponível em: https://data.gov.ie/uploads/page_images/2021-05-11-144708.586537Open-Data-Publication-Guidelines-2021.pdf. Acesso em: 22 jul. 2021.

JETZEK, T.; AVITAL, M.; BJØRN-ANDERSEN, N. Generating sustainable value from open data in a sharing society. *In: International Working Conference on Transfer and Diffusion of IT*. Springer, Berlin, Heidelberg, 2014. p. 62–82.

JUANA-ESPINOSA, S. DE.; LUJÁN-MORA, S. Open government data portals in the European Union: Considerations, development, and expectations. **Technological Forecasting and Social Change**, v. 149, p. 119769, 2019. Disponível em: <https://doi.org/10.1016/j.techfore.2019.119769>. Acesso em: 12 jun. 2020.

KASSEN, M. Open data and its institutional ecosystems: A comparative cross-jurisdictional analysis of open data platforms. **Canadian Public Administration**, v. 61, n. 1, p. 109–129, 2018. Disponível em: <https://doi.org/10.1111/capa.12251>. Acesso em: 22 jul. 2020.

KITCHIN, R. **Four critiques of open data initiatives**. Programmable City blog, 2013. Disponível em: <http://progcity.maynoothuniversity.ie/2013/11/four-critiques-of-open-data-initiatives/>. Acesso em: 13 jan. 2020.

KUČERA, J.; CHLAPEK, D.; KLÍMEK, J.; NEČASKÝ, M. Methodologies and Best Practices for Open Data Publication. *In: Proceedings of the Dateso 2015 annual international workshop on databases, texts, specifications and objects in Czech Republic*. [S.l.], 2015. p. 52-64.

LANCASTER, F. W. **Vocabulary control for information retrieval**. Information Resources Press, Arlington, VA, 1986.

LANCASTER, F. W. **Indexação e resumos: teoria e prática**. 2. ed. Brasília: Briquet de Lemos, 2004.

LAUFER, C. Guia de web semântica. **Governo do Estado de São Paulo e Governo do Reino Unido**, 2015.

LE COADIC, Y-F. **A Ciência da Informação**. 2.ed. Brasília: Briquet de Lemos, 1996.

LEMOS, D. L. DA S.; MENDONÇA, F. M.; SOUZA, R. R. Ontologias no suporte semântico na organização de acervos digitais em rede. *In: M. B. Almeida, Representação do Conhecimento, Ontologias e Linguagem: pesquisa aplicada em Ciência da Informação*, p. 161–191, Curitiba: CRV, 2020.

LEMOS, D. L. DA S.; SÁ, A. M.; SOUZA, R. R. Padrões para documentação de réplicas digitais em 3d: o caso de esculturas modernistas no espaço público do Rio de Janeiro. *In: Encontro Nacional de Pesquisa em Ciência da Informação*, n. XX ENANCIB, 2019. Disponível em: <https://brapci.inf.br/index.php/res/v/122745>. Acesso em: 22 mar. 2020.

LEMOS, D. L. DA S.; SOUZA, R. R. Organização de recursos bibliográficos e multimídia na web: contribuições interdisciplinares. **Informação & Informação**, Londrina, v. 23, n. 2, p. 98–126, set. 2018. Disponível em: <https://dx.doi.org/10.5433/1981-8920.2018v23n2p98>. Acesso em: 22 mar. 2020.

LEMOS, D. L. DA S.; SOUZA, R. R. Ontologias na representação de documentos: um panorama atual para descrição de conteúdo multimídia em rede. **Informação & Sociedade: Estudos**, João Pessoa, v. 29, n. 4, p. 103–134, 9 dez. 2019. Disponível em: <https://brapci.inf.br/index.php/res/v/147946>. Acesso em: 22 mar. 2020.

LEMOS, D.L. DA S.; SOUZA, R. R. Knowledge Organization Systems for the Representation of Multimedia Resources on the Web: A Comparative Analysis. **Knowledge Organization**, v. 47, n. 4, p. 300–319, 2020a. Disponível em: <https://dx.doi.org/10.5771/0943-7444-2020-4-300>. Acesso em: 15 mar. 2021.

LEMOS, D.L. DA S.; SOUZA, R. R. Representação de recursos multimídia na web: uso e reúso de padrões de anotação. **Perspectivas em Ciência da Informação (on line)**, Belo Horizonte, v. 25, número especial, p. 202–232, fev. 2020b. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/135735>. Acesso em: 24 jan. 2021.

LUNA-REYES, L. F.; NAJAFABADI, M. M. The US open data initiative: The road ahead. **Information Polity**, v. 24, n. 2, p. 163–182, 2019. Disponível em: <https://dx.doi.org/10.3233/IP-180106>. Acesso em: 10 jun. 2020.

MACEDO, D. F.; LEMOS, D. L. DA S. Dados governamentais na rede linked open data: iniciativas nas esferas públicas brasileiras. *In: Encontro Nacional de Pesquisa em Ciência da Informação*, n. XX ENANCIB, 2019. Disponível em: <https://www.brapci.inf.br/index.php/res/v/122809>. Acesso em: 20 jan. 2020.

MACHADO, L. M. O.; SOUZA, R. R.; DA GRAÇA SIMÕES, M. Semantic web or web of data? a diachronic study (1999 to 2017) of the publications of tim berners-lee and the world wide web consortium. **Journal of the Association for Information**

Science and Technology, v. 70, n. 7, p. 701–714, 2019. Disponível em: <https://doi.org/10.1002/asi.24111>. Acesso em: 14 out. 2020.

MENDONÇA, F. M.; ALMEIDA, M. B. Modelos e teorias para representação: uma teoria ontológica sobre o sangue humano. *In: Encontro Nacional de Pesquisa em Ciência da Informação*, n. XIII ENANCIB, 2012. Disponível em: https://mba.eci.ufmg.br/downloads/ModTeoEnancib2012_v8_MBA_web.pdf. Acesso em: 15 out. 2020.

MYLOPOULOS, J. **Conceptual modeling and telos**. *In: LOUCOPOULOS, P.; ZICARI, R. (Ed.). Conceptual modeling, databases, and CASE*. [S.l.]: Wiley, 1992. cap. 2, p. 49–68.

OBERLE, D; ANKOLEKAR, A; HITZLER, P.; CIMIANO, P.; SINTEK, M.; KIESEL, M. *et al.* DOLCE ergo SUMO: On foundational and domain models in the SmartWeb Integrated Ontology (SWIntO). **Journal of Web Semantics**, [S.l.], v. 5, n. 3, p. 156–174, 2007. Disponível em: <https://doi.org/10.1016/j.websem.2007.06.002>. Acesso em: 15 jul. 2020.

OPEN DATA. **5-Star**, 2010. Disponível em: <https://5stardata.info/pt-BR/>. Acesso em: 17 jul. 2019.

OPEN DATA CHARTER. **International Open Data Charter**, 2015. Disponível em: <https://opendatacharter.net/principles/>. Acesso em: 20 jan. 2020.

OPEN DATA MONITOR. **Open Data Monitor Portal**, 2021. Disponível em: <https://opendatamonitor.eu/>. Acesso em: 20 jan. 2021.

OPEN GOVERNMENT PARTNERSHIP. **Open Government Declaration**, 2011. Disponível em: <https://www.opengovpartnership.org/process/joining-ogp/open-government-declaration/>. Acesso em: 9 jan. 2020.

OPEN GOVERNMENT WORKING GROUP. **Eight principles of open government data**, 2007. Disponível em: https://public.resource.org/8_principles.html. Acesso em: 11 jan. 2020.

OPEN KNOWLEDGE FOUNDATION. **The Open Definition**, 2021. Disponível em: <http://opendefinition.org/>. Acesso em: 20 ago. 2021.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT (OECD). Open, Useful and Re-usable data (OURdata) Index: 2019. **OECD Public Governance Policy Papers**. Paris, 2020. Disponível em: <http://www.oecd.org/gov/digital-government/ourdata-index-policy-paper-2020.pdf>. Acesso em: 20 ago. 2021.

PAVLÍK, J.; HRNČÍROVÁ, M.; STOČES, M.; MASNER, J.; VANĚK, J. Usability of IoT and Open Data Repositories for Analyzing Water Pollution. A Case Study in the Czech Republic. **ISPRS International Journal of Geo-Information**, v. 9, n. 10, p.

591, 2020. Disponível em: <https://dx.doi.org/10.3390/ijgi9100591>. Acesso em: 21 jan. 2021.

PINTO, J. A.; ALMEIDA, M. B. Ontologias públicas sobre governo eletrônico. **Brazilian Journal of Information Science: Research Trends**, v. 14, n. 3, p. e020003, 2020. Disponível em: <https://doi.org/10.36311/1940-1640.2020.v14n3.10105>. Acesso em: 18 mar. 2021.

RAUTENBERG, S.; SOUZA, L. DE.; DALL'AGNOL, J. M. H.; MICHELON, G. A. **Guia prático para publicação de dados abertos conectados na web**. Curitiba: Appris, 2018.

REINO UNIDO. DEPARTMENT FOR DIGITAL, CULTURE, MEDIA & SPORT. **U.K National Data Strategy**, 2020. Disponível em: <https://www.gov.uk/government/publications/uk-national-data-strategy/national-data-strategy>. Acesso em: 16 ago. 2021.

REINO UNIDO. GOVERNMENT DIGITAL SERVICE. **U.K Government's Open Data Portal**, 2021a. Disponível em: <https://data.gov.uk/>. Acesso em: 16 maio 2021.

REINO UNIDO. GOVERNMENT DIGITAL SERVICE. **Data.gov.uk technical documentation datasets**, 2021b. Disponível em: <https://guidance.data.gov.uk/>. Acesso em: 16 maio 2021.

REINO UNIDO. GOVERNMENT DIGITAL SERVICE. **Using metadata to describe data shared within government**, 2021c. Disponível em: <https://www.gov.uk/government/publications/recommended-open-standards-for-government/using-metadata-to-describe-data-shared-within-government>. Acesso em: 16 maio 2021.

REINO UNIDO. GOVERNMENT DIGITAL SERVICE. **Data.gov.uk feedback**, 2021d. Disponível em: <https://www.smartsurvey.co.uk/s/3SEXD/>. Acesso em: 16 jul. 2021.

RIBEIRO, C. J. S.; ALMEIDA, R. F. DE. Dados abertos governamentais (Open Government Data): instrumento para exercício de cidadania pela sociedade. *In: Encontro Nacional de Pesquisa em Ciência da Informação*, n. XII ENANCIB, 2011. Disponível em: <http://arq.3rengtt.com.br/wp-content/uploads/2015/09/EnancibXII-RibeiroAlmeida.pdf>. Acesso em: 10 out. 2019.

RILEY, J. **Understanding metadata: what is metadata, and what is it for?** National Information Standards Organization (NISO), 2017.

ROCHA, L. DE L.; CAMPOS, M. L. DE A.; COSTA, L. C. DA. Ontologia como ferramenta para a interligação de dados governamentais abertos: uma análise comparativa. *In: Encontro Nacional de Pesquisa em Ciência da Informação*, n. XVII ENANCIB, 2016. Disponível em: <http://repositorios.questoesemrede.uff.br/repositorios/handle/123456789/3380>. Acesso em: 15 out. 2019.

SÁEZ MARTÍN, A.; ROSARIO, A. H. D.; PÉREZ, M. D. C. C. An international analysis of the quality of open government data portals. **Social science computer review**, v. 34, n. 3, p. 298–311, 2016. Disponível em: <https://doi.org/10.1177/0894439315585734>. Acesso em: 10 ago. 2020.

SANT'ANA, R. C. G. **Tecnologia e Gestão Pública Municipal: mensuração da interação com a sociedade**. São Paulo: Cultura Acadêmica Editora - UNESP, 2009. v. 1.

SARACEVIC, T. Ciência da informação: origem, evolução e relações. **Perspectivas em Ciência da Informação**, Belo Horizonte/MG, v. 1, n. 1, p. 41–62, jan./jun. 1996.

SARACEVIC, T. Information science. In: Marcia J. Bates and Mary Niles Maack (eds.) **Encyclopedia of Library and Information Science**. New York: Taylor & Francis. p. 2570–2586, 2009.

SAYÃO, L. F. Modelos teóricos em ciência da informação - abstração e método científico. **Ciência da Informação**, Brasília, v. 30, n. 1, p. 82–91, abr. 2001. Disponível em: <http://revista.ibict.br/ciinf/article/view/941>. Acesso em: 20 out. 2020.

SANTAREM SEGUNDO, J. E. Web semântica, dados ligados e dados abertos: uma visão dos desafios do brasil frente às iniciativas Internacionais. In: **Encontro Nacional de Pesquisa em Ciência da Informação**, n. XVI ENANCIB, 2015. Disponível em: <https://www.brapci.inf.br/index.php/article/download/43838>. Acesso em: 20 out. 2020.

SAXENA, S. Open public data (OPD) and the Gulf Cooperation Council (GCC): challenges and prospects. **Contemporary Arab Affairs**, v. 10, n. 2, p. 228–240, 2017. Disponível em: <http://dx.doi.org/10.1080/17550912.2017.1297565>. Acesso em: 28 jul. 2020.

SCHROCK, A. R. Civic hacking as data activism and advocacy: A history from publicity to open government data. **New media & society**, v. 18, n. 4, p. 581–599, 2016. Disponível em: <https://doi.org/10.1177/1461444816629469>. Acesso em: 10 jun. 2020.

UNIÃO EUROPÉIA. **Serviço de Publicações da União Europeia**, 2021. Disponível em: <https://op.europa.eu/en/web/eu-vocabularies/ontologies>. Acesso em: 20 jul. 2021.

SHEPHERD, E.; BUNN, J.; FLINN, A.; LOMAS, E.; SEXTON, A; BRIMBLE, S. *et al.* Open government data: critical information management perspectives. **Records Management Journal**, v. 29, n. 1/2, p. 152–167, 2019. Disponível em: <https://dx.doi.org/10.1108/RMJ-08-2018-0023>. Acesso em: 10 jun. 2020.

SHERIDAN, J.; TENNISON, J. Linking UK Government Data. In: **Linked Data on the Web (LDOW2010)**, Raleigh, North Carolina, 2010.

SILVA, P. N.; PINHEIRO, M. M. K. Dados governamentais abertos e a lei de acesso à informação: diagnóstico nas universidades públicas federais brasileiras. *In: Encontro Nacional de Pesquisa em Ciência da Informação*, n. XVI ENANCIB, 2015. Disponível em: <http://repositorios.questoesemrede.uff.br/repositorios/handle/123456789/2900>. Acesso em: 20 nov. 2020.

SILVA, P. N.; PINHEIRO, M. M. K. Execução da política de dados abertos no Brasil. *Tendências da Pesquisa Brasileira em Ciência da Informação*, v. 12, n. 2, 2019. Disponível em: <https://revistas.ancib.org/index.php/tpbci/article/view/495>. Acesso em: 22 out. 2020.

SIQUEIRA, I. C. P.; SILVA, J. F. M. DA. Metadados: o fio de Ariadne ou a coragem de Teseu? *Bibl. Univ.*, Belo Horizonte, v. 1, n. 1, p. 11–18, jun. 2011. Disponível em: <https://periodicos.ufmg.br/index.php/revistarbu/article/view/3059>. Acesso em: 21 out 2020.

SOUZA, R. R; ALVARENGA, L. A Web Semântica e suas contribuições para a ciência da informação. *Ciência da Informação*, Brasília, v. 33, n. 1, p. 132–141, 2004. Disponível em: <https://doi.org/10.1590/S0100-19652004000100016>. Acesso em: 21 out. 2020.

VALENTIM, M. L. P. (org.). **Métodos qualitativos de pesquisa em Ciência da Informação**. São Paulo: Polis, 2005.

VASSILAKIS, C; LEPOURAS, G. Ontology for E-government public services. *In: Encyclopedia of E-commerce, E-government, and mobile commerce*. IGI Global, 2006. p. 865–870.

VICKERY, B. C. Knowledge representation: a brief review. *Journal of Documentation*, [S.l.], v. 42, n. 3, p. 145–159, 1986.

VETRÒ, A.; CANOVA, L.; TORCHIANO, M.; MINOTAS, C. O.; IEMMA, R.; MORANDO, F. Open data quality measurement framework: Definition and application to Open Government Data. *Government Information Quarterly*, v. 33, n. 2, p. 325–337, 2016. Disponível em: <http://dx.doi.org/10.1016/j.giq.2016.02.001>. Acesso em: 10 jun. 2020.

WANG, D.; CHEN, C.; RICHARDS, D. A prioritization-based analysis of local open government data portals: A case study of Chinese province-level governments. *Government Information Quarterly*, v. 35, n. 4, p. 644–656, 2018. Disponível em: <https://doi.org/10.1016/j.giq.2018.10.006>. Acesso em: 10 jun. 2020.

WANG, V.; SHEPHERD, D. Exploring the extent of openness of open government data—A critique of open government datasets in the UK. *Government Information Quarterly*, v. 37, n. 1, p. 101405, 2020. Disponível em: <https://doi.org/10.1016/j.giq.2019.101405>. Acesso em: 10 jun. 2020.

WANG, V., SHEPHERD, D.; BUTTON, M. The barriers to the opening of government data in the UK: A view from the bottom. **Information Polity**, v. 24, n. 1, p. 59–74, 2019. Disponível em: <https://dx.doi.org/10.3233/IP-180107>. Acesso em: 10 jun. 2020.

WIRTZ, B. W. Resistance of public personnel to open government: A cognitive theory view of implementation barriers towards open government data. **Public Management Review**, v. 18, n. 9, p. 1335–1364, 2016. Disponível em: <https://dx.doi.org/10.1080/14719037.2015.1103889>. Acesso em: 21 jan. 2021.

W3C. **Linked data cookbook**. W3C Government Linked Data Working Group Charter, 2013. Disponível em: http://www.w3.org/2011/gld/wiki/Linked_Data_Cookbook. Acesso em: 7 jan. 2020.

W3C. **Data on the web best practices**. W3C, 2017. Disponível em: <https://www.w3.org/TR/dwbp/>. Acesso em: 7 jan. 2020.

W3C. **Data Catalog Vocabulary (DCAT) - Version 2**. W3C Recommendation, 2020a. Disponível em: <https://www.w3.org/TR/vocab-dcat/>. Acesso em: 3 maio 2021.

W3C. **RDFa Core Initial Context**. W3C Digital Publishing Activity Lead, 2020b. Disponível em: <https://www.w3.org/2011/rdfa-context/rdfa-1.1>. Acesso em: 3 maio 2021.

WORLD WIDE WEB FOUNDATION. **Open Data Barometer - Leaders Edition**. Washington DC: World Wide Web Foundation. 2018. Disponível em: <https://opendatabarometer.org/doc/leadersEdition/ODB-leadersEdition-Report.pdf>. Acesso em: 13 jan. 2020.

YI, M. Exploring the quality of government open data: Comparison study of the UK, the USA and Korea. **The Electronic Library**, v. 37, n. 1, p. 35–48, 2019. Disponível em: <https://dx.doi.org/10.1108/EL-06-2018-0124>. Acesso em: 10 jun. 2020.

YOUNG, A.; VERHULST, S. **The global impact of open data: Key findings from detailed case studies around the world**. O'Reilly Media, Inc., 2016.

APÊNDICES

APÊNDICE A – REFERÊNCIAS DA AMOSTRA DE DOCUMENTOS ESCOLHIDOS PARA ANÁLISE ORGANIZADOS EM ORDEM CRONOLÓGICA DE PUBLICAÇÃO

ATTARD, J.; ORLANDI, F.; SCERRI, S.; AUER, S. A systematic review of open government data initiatives. **Government Information Quarterly**, v. 32, n. 4, p. 399-418, 2015. Disponível em: <http://dx.doi.org/10.1016/j.giq.2015.07.006>. Acesso em: 10 jun. 2020.

BRASIL. TRIBUNAL DE CONTAS DA UNIÃO. **Cinco Motivos para a Abertura de Dados na Administração Pública**, 2015. Disponível em: <https://portal.tcu.gov.br/biblioteca-digital/cinco-motivos-para-a-abertura-de-dados-na-administracao-publica.htm>. Acesso em: 11 jul. 2020.

SCHROCK, A. R. Civic hacking as data activism and advocacy: A history from publicity to open government data. **New media & society**, v. 18, n. 4, p. 581–599, 2016. Disponível em: <https://doi.org/10.1177/1461444816629469>. Acesso em: 10 jun. 2020.

VETRÒ, A.; CANOVA, L.; TORCHIANO, M.; MINOTAS, C. O.; IEMMA, R.; MORANDO, F. Open data quality measurement framework: Definition and application to Open Government Data. **Government Information Quarterly**, v. 33, n. 2, p. 325–337, 2016. Disponível em: <http://dx.doi.org/10.1016/j.giq.2016.02.001>. Acesso em: 10 jun. 2020.

WIRTZ, B. W. Resistance of public personnel to open government: A cognitive theory view of implementation barriers towards open government data. **Public Management Review**, v. 18, n. 9, p. 1335–1364, 2016. Disponível em: <https://dx.doi.org/10.1080/14719037.2015.1103889>. Acesso em: 21 jan. 2021.

YOUNG, A.; VERHULST, S. **The global impact of open data: Key findings from detailed case studies around the world**. O'Reilly Media, Inc., 2016.

AFFUL-DADZIE, E.; AFFUL-DADZIE, A. Open Government Data in Africa: a preference elicitation analysis of media practitioners. **Government Information Quarterly**, v. 34, n. 2, p. 244–255, 2017. Disponível em: <http://dx.doi.org/10.1016/j.giq.2017.02.005>. Acesso em: 10 jun. 2020.

ESTADOS UNIDOS. U.S. CONGRESS. **H.R.1770 - OPEN Government Data Act**, 2017. Disponível em: <https://www.congress.gov/bill/115th-congress/house-bill/1770>. Acesso em: 10 jul. 2021.

IRLANDA. DEPARTMENT OF PUBLIC EXPENDITURE AND REFORM. **Open Data Strategy 2017-2022**, 2017. Disponível em: https://data.gov.ie/uploads/page_images/2018-03-07-114306.063816Final-Strategy-online-version1.pdf. Acesso em: 25 jul. 2021.

SAXENA, S. Open public data (OPD) and the Gulf Cooperation Council (GCC): challenges and prospects. **Contemporary Arab Affairs**, v. 10, n. 2, p. 228–240,

2017. Disponível em: <http://dx.doi.org/10.1080/17550912.2017.1297565>. Acesso em: 28 jul. 2020.

ALTAYAR, M. S. Motivations for open data adoption: An institutional theory perspective. **Government Information Quarterly**, v. 35, n. 4, p. 633–643, 2018. Disponível em: <https://doi.org/10.1016/j.giq.2018.09.006>. Acesso em: 10 jun. 2020.

CHATFIELD, A. T.; REDDICK, C. G. The role of policy entrepreneurs in open government data policy innovation diffusion: an analysis of Australian Federal and State Governments. **Government Information Quarterly**, v. 35, n. 1, p. 123–134, 2018. Disponível em: <http://dx.doi.org/10.1016/j.giq.2017.10.004>. Acesso em: 10 jun. 2020.

KASSEN, M. Open data and its institutional ecosystems: A comparative cross-jurisdictional analysis of open data platforms. **Canadian Public Administration**, v. 61, n. 1, p. 109–129, 2018. Disponível em: <https://doi.org/10.1111/capa.12251>. Acesso em: 22 jul. 2020.

WANG, D.; CHEN, C.; RICHARDS, D. A prioritization-based analysis of local open government data portals: A case study of Chinese province-level governments. **Government Information Quarterly**, v. 35, n. 4, p. 644–656, 2018. Disponível em: <https://doi.org/10.1016/j.giq.2018.10.006>. Acesso em: 10 jun. 2020.

BLANK, M. **Open Data Maturity Report 2019**, 2019. Disponível em: https://www.europeandataportal.eu/sites/default/files/open_data_maturity_report_2019.pdf. Acesso em: 10 nov. 2019.

CRUSOE, J. R.; AHLIN, K. Users' activities for using open government data—a process framework. **Transforming government: People, process and policy**, v. 13, n. 3/4, p. 213–236, 2019. Disponível em: <http://dx.doi.org/10.1108/TG-04-2019-0028>. Acesso em: 10 jun. 2020.

IRLANDA. DEPARTMENT OF PUBLIC EXPENDITURE AND REFORM. **Public Service Data Strategy 2019-2023**, 2019. Disponível em: <https://assets.gov.ie/7107/7ac4ae109cd944dc99c9e178962fa095.pdf>. Acesso em: 25 jul. 2021.

JUANA-ESPINOSA, S. DE.; LUJÁN-MORA, S. Open government data portals in the European Union: Considerations, development, and expectations. **Technological Forecasting and Social Change**, v. 149, p. 119769, 2019. Disponível em: <https://doi.org/10.1016/j.techfore.2019.119769>. Acesso em: 12 jun. 2020.

LUNA-REYES, L. F.; NAJAFABADI, M. M. The US open data initiative: The road ahead. **Information Polity**, v. 24, n. 2, p. 163–182, 2019. Disponível em: <https://dx.doi.org/10.3233/IP-180106>. Acesso em: 10 jun. 2020.

SHEPHERD, E.; BUNN, J.; FLINN, A.; LOMAS, E.; SEXTON, A; BRIMBLE, S. *et al.* Open government data: critical information management perspectives. **Records**

Management Journal, v. 29, n. 1/2, p. 152–167, 2019. Disponível em: <https://dx.doi.org/10.1108/RMJ-08-2018-0023>. Acesso em: 10 jun. 2020.

SILVA, P. N.; PINHEIRO, M. M. K. Execução da política de dados abertos no Brasil. **Tendências da Pesquisa Brasileira em Ciência da Informação**, v. 12, n. 2, 2019. Disponível em: <https://revistas.ancib.org/index.php/tpbci/article/view/495>. Acesso em: 22 out. 2020.

WANG, V., SHEPHERD, D.; BUTTON, M. The barriers to the opening of government data in the UK: A view from the bottom. **Information Polity**, v. 24, n. 1, p. 59–74, 2019. Disponível em: <https://dx.doi.org/10.3233/IP-180107>. Acesso em: 10 jun. 2020.

YI, M. Exploring the quality of government open data: Comparison study of the UK, the USA and Korea. **The Electronic Library**, v. 37, n. 1, p. 35–48, 2019. Disponível em: <https://dx.doi.org/10.1108/EL-06-2018-0124>. Acesso em: 10 jun. 2020.

ESTADOS UNIDOS. FEDERAL DATA STRATEGY. **Data Governance Playbook**. 2020a. Disponível em: <https://resources.data.gov/assets/documents/fds-data-governance-playbook.pdf>. Acesso em: 22 maio 2021.

ESTADOS UNIDOS. FEDERAL DATA STRATEGY. **2020 Action Plan**. 2020b. Disponível em: <https://strategy.data.gov/action-plan/>. Acesso em: 10 jul. 2021.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT (OECD). Open, Useful and Re-usable data (OURdata) Index: 2019. **OECD Public Governance Policy Papers**. Paris, 2020. Disponível em: <http://www.oecd.org/gov/digital-government/ourdata-index-policy-paper-2020.pdf>. Acesso em: 20 ago. 2021.

PAVLÍK, J.; HRNČÍROVÁ, M.; STOČES, M.; MASNER, J.; VANĚK, J. Usability of IoT and Open Data Repositories for Analyzing Water Pollution. A Case Study in the Czech Republic. **ISPRS International Journal of Geo-Information**, v. 9, n. 10, p. 591, 2020. Disponível em: <https://dx.doi.org/10.3390/ijgi9100591>. Acesso em: 21 jan. 2021.

PINTO, J. A.; ALMEIDA, M. B. Ontologias públicas sobre governo eletrônico. **Brazilian Journal of Information Science: Research Trends**, v. 14, n. 3, p. e020003, 2020. Disponível em: <https://doi.org/10.36311/1940-1640.2020.v14n3.10105>. Acesso em: 18 mar. 2021.

REINO UNIDO. DEPARTMENT FOR DIGITAL, CULTURE, MEDIA & SPORT. **U.K National Data Strategy**, 2020. Disponível em: <https://www.gov.uk/government/publications/uk-national-data-strategy/national-data-strategy>. Acesso em: 16 ago. 2021.

WANG, V.; SHEPHERD, D. Exploring the extent of openness of open government data—A critique of open government datasets in the UK. **Government Information Quarterly**, v. 37, n. 1, p. 101405, 2020. Disponível em: <https://doi.org/10.1016/j.giq.2019.101405>. Acesso em: 10 jun. 2020.

BRASIL. Controladoria-Geral da União. **Plano de Dados Abertos vigência junho/2021 a junho/2023**, 2021a. Disponível em: <https://www.gov.br/cgu/pt-br/acesso-a-informacao/dados-abertos/plano-de-dados-abertos-da-cgu/arquivos/plano-de-dados-abertos-2018-dezembro.pdf>. Acesso em: 17 jul. 2021.

BRASIL. Comitê Gestor da Infraestrutura Nacional de Dados Abertos. **Plano de ação 2021-2022**, 2021f. Disponível em: <https://dados.gov.br/wp/wp-content/uploads/2021/03/Plano-de-A%c3%a7%c3%a3o-INDA-vf.pdf>. Acesso em: 10 ago. 2021.

ESTADOS UNIDOS. OFFICE OF MANAGEMENT AND BUDGET. **Federal Data Strategy**, 2021c. Disponível em: <https://strategy.data.gov/>. Acesso em: 10 maio 2021.

IRLANDA. DEPARTMENT OF PUBLIC EXPENDITURE AND REFORM. **DGI Front Page SLA report**, 2021b. Disponível em: <https://synthetics.newrelic.com/report/R2MvK?view=monthly-sla-report>. Acesso em: 22 jul. 2021.

IRLANDA. DEPARTMENT OF PUBLIC EXPENDITURE AND REFORM. **Open Data Publication Guidelines**, 2021d. Disponível em: https://data.gov.ie/uploads/page_images/2021-05-11-144708.586537Open-Data-Publication-Guidelines-2021.pdf. Acesso em: 22 jul. 2021.

REINO UNIDO. GOVERNMENT DIGITAL SERVICE. **Data.gov.uk technical documentation datasets**, 2021b. Disponível em: <https://guidance.data.gov.uk/>. Acesso em: 16 maio 2021.

APÊNDICE B – PRÉ-TESTE REALIZADO NO PORTAL DE DADOS ABERTOS DOS ESTADOS UNIDOS

O portal dos EUA é responsável por disponibilizar dados abertos tanto do governo federal, quanto de *datasets* de governos locais. A Figura 8 a seguir exibe como se apresenta a página inicial do portal norte-americano.

Figura 8 - Tela inicial do portal dos EUA



Fonte: Estados Unidos (2021a).

Sua estrutura principal de navegação está bem exposta na parte superior do cabeçalho, sendo o ponto de partida para a exploração de todo seu conteúdo e documentação institucional. A seção de recursos conta com diversas instruções orientadas aos profissionais das áreas de interesse na abertura de dados, categorizadas da seguinte forma: i) Gestão e governança de dados; ii) Ferramentas de dados; iii) Incubadora de dados; iv) Desenvolvimento de habilidades; v) Orientação; e vi) Estudos de caso e exemplos (ESTADOS UNIDOS, 2021a).

O portal americano é impulsionado por dois aplicativos de código aberto, CKAN v.2.8.7 e *WordPress*, com os códigos disponibilizados na plataforma *GitHub* (ESTADOS UNIDOS, 2021a). O CKAN é um sistema de gerenciamento de código aberto baseado na web usado para catalogar dados abertos, facilitando sua localização, organização, distribuição e documentação. Ele guarda os metadados,

podendo armazenar ou não os conjuntos de dados. Inicialmente ele fornece um conjunto de metadados, permitindo que alguns sejam customizados (CKAN, 2021).

A área de Recursos regulamenta que a normalização dos metadados publicados no portal devam ser no padrão do vocabulário DCAT-US *Schema v1.1 (Project Open Data Metadata Schema)* (ESTADOS UNIDOS, 2021b), sendo este derivado do *Data Catalog Vocabulary (DCAT)*, um vocabulário hierárquico específico para conjuntos de dados desenvolvido pela W3C (W3C, 2020a). Outros vocabulários com o uso encorajado para descrição de metadados são: *Dublin Core*, *Schema.org*, FGDC, ISO 19115 e NIEM (ESTADOS UNIDOS, 2021b).

Ainda na área de Recursos, é disponibilizado conteúdo sobre gestão e governança de dados, com variado material para instrução (*Playbooks*, orientações, modelos etc.) que visam apoiar a implementação de políticas, a criação de estruturas de governança de dados e o trabalho diário de gerenciamento de dados no governo federal (ESTADOS UNIDOS, 2021b).

O portal é linkado com o sítio da Estratégia Federal de Dados, que tem a incumbência de acelerar o uso de dados, servir ao público e administrar recursos enquanto protege a segurança, a privacidade e a confidencialidade das informações (ESTADOS UNIDOS, 2021c).

Com relação as licenças de uso, os trabalhos criados por funcionários do Governo em suas atividades são padronizados para Domínio Público dos Estados Unidos, não estando sujeitos à proteção de direitos autorais doméstica. No entanto, os trabalhos produzidos por terceiros que são criados ou obtidos para uso pelo Governo dos Estados Unidos podem precisar de licenças abertas aplicadas a eles (ESTADOS UNIDOS, 2021b).

Todo o conteúdo informacional produzido ou compartilhado pelas agências do governo americano é limitado à jurisdição dos EUA, tendo o status de domínio público. Caso uma agência promova internacionalmente os trabalhos que produziu, ela pode usar uma dedicação de domínio público mundial, como *Creative Commons Zero*⁹. Uma observação é que os dados não federais do portal são regidas por diferentes termos de serviço e, muitas vezes, por licenças diferentes dos dados federais (ESTADOS UNIDOS, 2021b).

⁹ A pessoa que associou uma obra a este documento dedicou a obra ao domínio público, renunciando a todos os seus direitos sobre a obra em todo o mundo sob a lei de direitos autorais, incluindo todos os direitos relacionados e conexos, na medida permitida por lei.

Concluindo o pré-teste americano, o portal possui uma boa documentação, compreendendo espectros importantes da abertura de dados, como governança, publicação, políticas, estratégias e capacitação. Além disso dispõe de vários recursos tecnológicos e aplicações úteis para publicadores de dados. Também possui boa acessibilidade aos dados, *design* funcional, uma busca eficiente, e com suporte de uma plataforma tecnológica com muitos recursos de catalogação e uso.

Exatamente por isso chama atenção tantos problemas para publicar dados em conformidade com os preceitos das melhores práticas internacionais, conforme elucidado no referencial teórico. Os fatores negativos ficam com a grande quantidade de dados abertos em formato PDF, que não é compatível com as boas práticas de abertura de dados. O portal também carece de uma exposição temática de *datasets* mais explicitada na sua página principal, o que melhoraria sobremaneira a experiência de uso, mas que não chega a comprometer o acesso.

APÊNDICE C – PRÉ-TESTE REALIZADO NO PORTAL DE DADOS ABERTOS DO REINO UNIDO

Com relação ao Reino Unido, o governo central, e os órgãos e autoridades locais podem utilizar o portal para catalogar seus dados e metadados (REINO UNIDO, 2021a). A Figura 9 exibe como se apresenta a página inicial do portal.

Figura 9 - Tela inicial do portal do Reino Unido

The screenshot shows the homepage of data.gov.uk. At the top, there is a navigation bar with the text 'data.gov.uk | Find open data' on the left and 'Publish your data', 'Documentation', and 'Support' on the right. Below this is a 'BETA' notice: 'This is a new service – your feedback will help us to improve it'. The main heading is 'Find open data'. Below the heading is a description: 'Find data published by central government, local authorities and public bodies to help you build products and services'. There is a search bar with the placeholder text 'Search data.gov.uk' and a magnifying glass icon. Below the search bar is a grid of nine data categories, each with a title and a brief description:

<p>Business and economy Small businesses, industry, imports, exports and trade</p>	<p>Government Staff numbers and pay, local councillors and department business plans</p>	<p>Towns and cities Includes housing, urban planning, leisure, waste and energy, consumption</p>
<p>Crime and justice Courts, police, prison, offenders, borders and immigration</p>	<p>Government spending Includes all payments by government departments over £25,000</p>	<p>Transport Airports, roads, freight, electric vehicles, parking, buses and footpaths</p>
<p>Defence Armed forces, health and safety, search and rescue</p>	<p>Health Includes smoking, drugs, alcohol,</p>	<p>Digital service performance Cost, usage, completion rate, digital</p>

Fonte: Reino Unido (2021a).

A página inicial possui uma organização temática dos *datasets*, com destaque para as bases econômicas, sociais, de saúde, além de um campo de busca por *datasets*. A navegação principal fica destacada no topo, sendo o ponto de partida para a exploração da documentação (REINO UNIDO, 2021a). A plataforma tecnológica adotada é o CKAN v.2.8.3 (REINO UNIDO, 2021b).

Para publicação, recomenda-se o uso do padrão de metadados *Dublin Core* (REINO UNIDO, 2021c), as exceções seriam: i) Para publicação dos dados diretamente no domínio do Reino Unido, deve-se utilizar o padrão *Schema.org*; ii) para descrever como formatar dados em um arquivo específico, deve-se utilizar os padrões RFC 4180 e *CSV on the Web standards*; iii) para criar e gerenciar dados geográficos espaciais, deve-se utilizar o padrão de metadados espaciais do Reino Unido denominado *GEMINI*, atentando-se as recomendações do *INSPIRE*; e iv) para criar e

gerenciar dados geográficos não espaciais, deve-se utilizar o padrão de metadados ISO 19139 (REINO UNIDO, 2021b).

A área de documentação do Reino Unido não abordou questões de governança e política de dados. Esse tipo de documentação deve ser explorado no domínio central, o que compromete a consistência da navegação e localização da documentação pertinente a governança de dados no portal de dados abertos.

Os termos de uso dos dados do Reino Unido estão sob a tutela da *Open Government Licence* (OGL) em sua versão 3.0 (aderente a Definição Aberta), que segundo a documentação oficial, contemplam tanto o licenciamento de direitos autorais, quanto o de banco de dados (REINO UNIDO, 2021a).

Portanto, pode-se redistribuir as Informações, podendo ser exploradas comercialmente. No entanto, é necessário dar o crédito apropriado, fornecendo um *link* para a licença e indicar se alterações foram feitas. A licença OGL não permite o uso de dados pessoais, utilização de informações que transgridam a legislação de acesso à informação, logotipos de organizações do setor público, insígnias militares, dentre outros direitos de propriedade intelectual (REINO UNIDO, 2021a).

O pré-teste britânico aponta uma boa acessibilidade aos dados, com um processo de busca simples, mas eficiente, porém preocupa a carência de documentação a respeito de governança, estratégias e política de dados. A grande quantidade de dados abertos no formato PDF também preocupa, já que é incompatível com as boas práticas de abertura.

APÊNDICE D – PRÉ-TESTE REALIZADO NO PORTAL DE DADOS ABERTOS DO BRASIL

No Brasil, o Portal Brasileiro de Dados Abertos objetiva catalogar dados abertos de diversas áreas da administração pública, sendo um facilitador para a busca e uso de dados publicados pelos órgãos do governo (BRASIL, 2021b). A Figura 10 a seguir exhibe como se comporta a página inicial do portal brasileiro.

Figura 10 - Tela inicial do portal brasileiro

Conjunto de dados	Data	Responsável
Roteiros de audiodescrição elaborados	30/04/2021	Instituto Benjamin Constant - IBC
Convênios e acordos	28/04/2021	Instituto Benjamin Constant - IBC
Extensão	27/04/2021	Universidade Federal de Minas Gerais...
Pós-graduação	27/04/2021	Universidade Federal de Minas Gerais...
Terceirização	27/04/2021	Universidade Federal de Minas

Fonte: Brasil (2021b).

A realidade brasileira apontou *datasets* publicados tanto por agências governamentais, ministérios, órgãos de controle, quanto por estados e municípios. O portal não hospeda os dados, apenas armazena os metadados e *links* para os dados que são catalogados nele pelas entidades públicas responsáveis, cabendo a cada um deles providenciar a própria hospedagem dos dados (BRASIL, 2021b).

A página inicial contém a navegação principal, um campo de busca por *datasets*, e área de publicações e notícias, além de uma navegação secundária no rodapé. Sua navegação principal se localiza de forma discreta no rodapé do topo (BRASIL, 2021b).

A partir da exploração do repositório INDA foi possível identificar a documentação relativa aos aspectos técnicos, políticos e legais da iniciativa brasileira.

Percebe-se uma preocupação com a criação de uma cultura de dados institucional, com planos, manuais, cartilhas técnicas, e *kits* de abertura para serem usadas pela administração pública (BRASIL, 2021d).

O conjunto de tecnologias documentadas pelo portal são: HTML, CSS, *Javascript*, *Python*, *WordPress*, *PostgreSQL*, todas estruturadas na plataforma CKAN v.2.6.8 (BRASIL, 2021b). O padrão de metadados adotado pela iniciativa baseia-se no padrão do governo chileno, este por sua vez utiliza o padrão *Dublin Core* (BRASIL, 2021d).

Com relação ao uso de vocabulários controlados, o portal suporta campos para uso de um vocabulário próprio, desenvolvido para assuntos do governo que utiliza termos comuns e é voltada para a sociedade, denominado: Vocabulário Controlado de Governo Eletrônico (VCGE) (BRASIL, 2021d).

Os demais padrões de vocabulários, taxonomias, ontologias e outros métodos de organização e recuperação de informações podem ser utilizados seguindo a documentação preconizada pelos Padrões de Interoperabilidade de Governo Eletrônico (ePING) (BRASIL, 2018). O padrão ePING segue orientação da W3C (BRASIL, 2018), esta por sua vez sugere muitos padrões, tais como: *RDF Schema 1.1*, *SKOS*, *D-CAT*, entre outros (W3C, 2020b).

Os documentos referentes a condução da política de dados podem ser localizados a partir da exploração do repositório INDA, onde foi possível minerar documentos sobre a política, plano de ação, e modelo de referência de abertura (BRASIL, 2021d). Com relação a licenças de uso, a INDA não concluiu se o ordenamento jurídico é suficiente para tratar os dados governamentais de forma aberta, sem a necessidade de licenças, se outras licenças podem ser utilizadas, ou se é o caso de criar uma licença própria (BRASIL, 2021d).

Concluindo, O portal brasileiro possui grande capilaridade documental, transparecendo uma preocupação relativa à interoperabilidade de dados, além de uma atenção as documentações que conduzem as políticas de dados. Os *datasets* não são organizados de forma temática, o que pode prejudicar a busca. Também foi percebido alguns problemas de padronização entre as páginas, o que pode atrapalhar a navegação.

APÊNDICE E – PRÉ-TESTE REALIZADO NO PORTAL DE DADOS ABERTOS DA IRLANDA

No portal da Irlanda, os dados abertos são publicados por departamentos e órgãos públicos, sendo que muitos *datasets* são publicados e atualizados individualmente por eles, já outros são coletados pelo *data.gov.ie* diariamente de catálogos de dados específicos. O portal não hospeda os *datasets*, sendo um catálogo de metadados (IRLANDA, 2021a). A Figura 11 a seguir exibe a tela inicial do portal irlandês.

Figura 11 - Tela inicial do portal da Irlanda

The screenshot shows the homepage of the Ireland's Open Data Portal. At the top, there is a navigation menu with links for Home, Datasets, Publishers, Suggest Data, Showcases, Contact, About, More, and EN/GA. Below the navigation is a search bar with the text 'Search Datasets...' and a magnifying glass icon. To the left of the search bar, there are two statistics: '10411 Datasets' and '122 Publishers'. Below these statistics is a banner for 'IRELAND'S OPEN DATA PORTAL' with the tagline 'Promoting innovation and transparency through the publication of Irish Public Sector data in open, free and reusable formats.' and a button labeled 'EXPLORE DATASETS'. Below the banner is a yellow section for 'COVID-19 Resources' with a link to 'COVID-19 DATA HUB & MORE' and a button 'SHOW DATASETS'. Below this is a grid of thematic categories: Agriculture, Fisheries, Forestry & Food; Arts, Culture and Heritage; Justice, Legal System and Public Safety; Economy and Finance; Education and Sport; Energy; Environment; Government and Public Sector; Health; Housing and Zoning; Population and Society; Science and Technology; Regions and Cities; and Transport. At the bottom, there is a blue section for 'Subscribe to our mailing list*' with an input field for 'Enter your email to signup' and a 'Sign up' button.

Fonte: IRLANDA (2021a).

É apresentada uma organização temática de *datasets* na página inicial, identificando-se bases sobre agricultura, arte, justiça, economia, educação, energia, meio ambiente, governo, saúde, habitação, sociedade, ciência, transporte e um grande destaque aos *datasets* relacionados a Covid-19. Também é possível identificar claramente a navegação principal do portal, o campo de busca, notícias e outras publicações de cunho institucional (IRLANDA, 2021a).

A plataforma tecnológica adotada é o CKAN v.2.8.7, e de acordo com a Seção de Desenvolvedores, os dados podem ser publicados de forma manual, com auxílio de formulário, ou usando uma API (*Application Programming Interface*) do CKAN para

programar as publicações, ou automatizando a publicação usando uma colheitadeira para extrair os dados de uma fonte pré-definida de forma periódica (IRLANDA, 2021a).

As recomendações técnicas para a publicação estão devidamente localizadas na Seção de Guia para Editores. A iniciativa irlandesa segue os preceitos definidos a partir *Open Data Technical Framework*, que estabelece a estrutura técnica para alavancar as iniciativas de abertura de dados na Irlanda, visando garantir que as publicações no país tenham consistência e uniformidade (IRLANDA, 2021a).

A iniciativa preconiza uma normatização para descrição de diversos domínios, como por exemplo no domínio de metadados em que a padronização é o uso do padrão *Dublin Core* com o uso do vocabulário *Application Profile for European Data Portals* (DCAT-AP), sendo este derivado do DCAT. Para outros domínios existem outros padrões específicos e pré-determinados (IRLANDA, 2021a).

Além da compatibilidade com o padrão DCAT-AP, o *Open Data Technical Framework* orienta o uso de formatos de dados a partir de 3 estrelas (CSV, JSON ou XML), além do uso do formato JSON para guardar a representação escolhida para os *datasets* (IRLANDA, 2021a).

A parte de governança e gestão de dados está documentada na Seção de Recursos e Publicações, contendo a estratégia de dados abertos do governo, estrutura para treinamento, manuais de boas práticas, guias e publicações diversas (IRLANDA, 2021a). De acordo com a Seção de Licença de Dados Abertos, os dados e metadados devem ser no mínimo associados a licença *Creative Commons Attribution 4.0*¹⁰, devendo estar claramente identificada nos metadados. Se for apropriado, é facultada a entidades públicas renunciarem a direitos autorais e associar seus *datasets* a licença (IRLANDA, 2021a).

O portal irlandês demonstra um comprometimento com as boas práticas internacionais. Essa característica é percebida nos planos e políticas de abertura, preocupadas com os aspectos políticos, técnicos e de gestão. Também é muito simbólico a participação de consultas públicas para algumas importantes definições, como foi o caso das licenças a serem adotadas pelo governo.

¹⁰ Sob a licença é possível copiar e redistribuir o material em qualquer meio ou formato, além de remixar, transformar e construir sobre o material para qualquer finalidade, mesmo comercialmente.