

Dados abertos no contexto brasileiro: uma exploração da rede de autores e dos temas correlatos

José Nivaldo Balbino

Doutorando em Gestão da Informação pela Universidade Federal do Paraná (UFPR) -Curitiba, PR – Brasil.
Mestre em Ciência, Gestão e Tecnologia da Informação pela.

Universidade Federal do Paraná (UFPR) -PR - Brasil. Administrador do Instituto Federal de Educação, Ciência e Tecnologia do Paraná (IFPR)- Curitiba, PR – Brasil.

<http://lattes.cnpq.br/2711486822172560>

E-mail: jnbalbino@gmail.com

Helena de Fátima Nunes da Silva

Doutora em Engenharia de Produção pela Universidade Federal de Santa Catarina – Florianópolis – SC
Professora Sênior do Programa de Pós-Graduação em Ciência, Gestão e Tecnologia da Informação, da Universidade Federal do Paraná – Curitiba - PR

<http://lattes.cnpq.br/1216712773575109>

E-mail: helenanunes@gmail.com

Flávia Roberta Fernandes

Doutoranda em Gestão da Informação pela Universidade Federal do Paraná (UFPR) -Curitiba, PR – Brasil.

Mestre em Ciência, Gestão e Tecnologia da Informação pela Universidade Federal do Paraná (UFPR) - Brasil. Bolsista pesquisadora da Universidade Federal do Paraná (UFPR) - Brasil. Pesquisadora do Instituto Brasileiro de Estudos e Pesquisas Sociais (IBEPES) - Brasil.

<http://lattes.cnpq.br/4979799552470346>

E-mail: flaroberta@gmail.com

João Augusto Wendt Mischiatti

Doutorando em Gestão da Informação pela Universidade Federal do Paraná (UFPR) – Curitiba, PR – Brasil. Mestrado profissional em Planejamento e Governança Pública pela

Universidade Tecnológica Federal do Paraná (UTFPR) – PR - Brasil. Analista de Gestão em Saúde da Fundação Oswaldo Cruz (Fiocruz) – Brasil.

<http://lattes.cnpq.br/4451724166753942>

E-mail: jomischiatti@gmail.com

Submetido em: 11/08/2019. Aprovado em: 19/02/2020. Publicado em: 03/08/2020.

RESUMO

A iniciativa de dados em formato aberto contribui para a ampliação da transparência entre a sociedade e o governo. Devido à crescente importância do tema, este estudo tem o objetivo de explorar os conceitos de dados abertos e os temas correlatos e de construir uma rede das principais referências utilizadas na elaboração de publicações científicas sobre dados abertos no contexto brasileiro. Foram realizadas buscas nas bases de dados da EBSCO, SciELO, Science Direct, Scopus e Web of Science utilizando o termo “dados abertos”. Após os critérios de exclusão e com o enfoque em artigos aplicados ao contexto brasileiro, foram identificados 17 artigos. Na sequência, procedeu-se à codificação e categorização dos conceitos com auxílio do software NVivo 12 Plus. Para melhor entendimento, utilizou-se a análise qualitativa baseada na metodologia da teoria fundamentada em dados. O processo de codificação resultou em oito categorias consolidadas a partir de 45 codificações iniciais. As categorias encontradas e classificadas pela quantidade de artigos vinculados foram: i) conceitos de dados abertos; ii) governo; iii) sociedade e controle; iv) tecnologias e ferramentas; v) benefícios e oportunidades; vi) regulação; vii) abertura de dados; e, viii) barreiras e desafios. Os resultados demonstram que sete referências são responsáveis por conectar (rede) os 17 artigos analisados que apresentam conceitos sobre dados abertos. Pela análise dos textos foi possível perceber a necessidade de um estudo que seja capaz de identificar e analisar o uso de dados abertos para a produção de novos conhecimentos.

Palavras-chave: Dados abertos. Revisão sistemática. Produção científica.

Open data in the brazilian context: an exploration of the network of authors and related themes

ABSTRACT

The open data initiative contributes to the expansion of transparency between society and government. Due to the growing importance of the theme, this study aims to explore the concepts of open data and related themes and to build a network of the main references used in the preparation of scientific publications on open data in the brazilian context. Searches were performed in the databases of EBSCO, SciELO, Science Direct, Scopus and Web of Science using the term “Open data”. After the exclusion criteria and with the focus on articles applied to the brazilian context, 17 articles were identified. Subsequently, the concepts were codified and categorized with the help of the NVivo 12 Plus software. For a better understanding, we used the qualitative analysis based on the methodology of grounded theory. The coding process resulted in eight consolidated categories from 45 initial encodings. The categories found and classified by the number of articles linked were: i) concepts of open data; ii) government; iii) society and control; iv) technologies and tools; v) benefits and opportunities; vi) Regulation; vii) opening of data; and, viii) barriers and challenges. The results show that seven references are responsible for connecting (network) The 17 analyzed articles that present concepts about open data. By analyzing the texts it was possible to perceive the need for a study that is able to identify and analyze the use of open data for the production of new knowledge.

Keywords: Open data. Systematic review. Scientific production.

Datos abiertos en el contexto brasileño: una exploración de la red de autores y temas relacionados

RESUMEN

La iniciativa de datos en formato abierto contribuye a la expansión de la transparencia entre la sociedad y el gobierno. Debido a la creciente importancia del tema, este estudio tiene como objetivo explorar los conceptos de datos abiertos y temas relacionados y construir una red de las principales referencias utilizadas en la preparación de publicaciones científicas sobre datos abiertos en el contexto brasileño. Las búsquedas se realizaron en las bases de datos de EBSCO, SciELO, Science Direct, Scopus y Web of Science utilizando el término "Open data". Después de los criterios de exclusión y con el enfoque en los artículos aplicados al contexto brasileño, se identificaron 17 artículos. Posteriormente, los conceptos fueron codificados y categorizados con la ayuda del software NVivo 12 Plus. Para una mejor comprensión, utilizamos el análisis cualitativo basado en la metodología de la teoría fundamentada. El proceso de codificación dio lugar a ocho categorías consolidadas a partir de 45 codificaciones iniciales. Las categorías encontradas y clasificadas por el número de artículos vinculados fueron: i) conceptos de datos abiertos; ii) gobierno; iii) sociedad y control; iv) tecnologías y herramientas; v) Beneficios y oportunidades; vi) reglamento; vii) apertura de datos; y, viii) barreras y desafíos. Los resultados muestran que siete referencias son responsables de conectar (red) los 17 artículos analizados que presentan conceptos sobre datos abiertos. Analizando los textos fue posible percibir la necesidad de un estudio que sea capaz de identificar y analizar el uso de datos abiertos para la producción de nuevos conocimientos.

Palabras clave: Datos abiertos. Revisión sistemática. Producción científica.

INTRODUÇÃO

O acesso à informação e ao conhecimento passa pelas discussões sobre o acesso aos dados abertos. Os protocolos, os formatos, as plataformas que permitem maior interoperabilidade, processamento, cruzamento e reúsos da informação estão em amplo debate, já que o acesso aberto tradicional, dos últimos 10 ou 15 anos, ficou ultrapassado, sendo insuficiente para dar conta dos avanços do uso e reuso de uma base de dados aberta (MACHADO, 2015).

O entendimento sobre o termo dado pode variar consideravelmente entre pesquisadores e áreas do conhecimento. A constatação é que os dados são gerados para diferentes propósitos, por distintas comunidades acadêmicas e científicas e por meio de diversos processos (SAYÃO; SALES, 2013).

O compartilhamento de dados em formato aberto concede amplo leque de oportunidades e benefícios. Sua utilização possibilita o desenvolvimento de novos serviços e produtos, o que favorece o desenvolvimento social e econômico, assim como apoia a ampliação da transparência e a prestação de contas públicas ao aproximar o governo da sociedade.

Nessa perspectiva, as pesquisas científicas cooperam com a divulgação dos conceitos, a sensibilização a respeito da importância do tema e a disseminação de boas práticas. As contribuições das produções científicas relacionadas a dados abertos, no contexto brasileiro, destacam o percurso realizado no movimento para a abertura dos dados, bem como seus principais estímulos e potenciais.

Logo, o presente estudo destina-se a apresentar os conceitos de dados abertos e os temas correlatos a partir da codificação de publicações científicas relacionadas ao tema e no contexto brasileiro. Este estudo também foca em construir uma rede das principais referências utilizadas na elaboração das publicações científicas analisadas. Para a realização desses objetivos, o próximo tópico apresenta a definição das bases conceituais sobre dados abertos e temas correlatos. Por conseguinte, a metodologia destina-se a explicitar as etapas para o desenvolvimento do estudo. Na sequência, são expostos os resultados e as discussões sobre as codificações realizadas e a rede de referências construída a partir dos artigos analisados.

As considerações finais sintetizam os principais elementos integradores do estudo, bem como, as limitações e trabalhos futuros.

CONCEITOS E TEMAS RELACIONADOS A DADOS ABERTOS

Apresenta-se uma revisão sobre os conceitos de dados abertos e de termos a eles relacionados, quais sejam, dados conectados, dados governamentais abertos, transparência e prestação de contas públicas, encontrados em publicações científicas.

Conceitualmente, os dados abertos são compreendidos como dados disponíveis a qualquer pessoa, seja física ou jurídica, para acessar, utilizar, reutilizar, manipular, compartilhar, assim como gerar novos produtos ou serviços. Independentemente da finalidade, os dados abertos estão sujeitos, no máximo, à atribuição da fonte e à manutenção da licença original (ALBANO; CRAVEIRO, 2015; SANTAREM SEGUNDO, 2013; W3C, 2009).

A Open Knowledge Foundation, organização internacional, sem fins lucrativos, com a finalidade de gerar valor para a sociedade por meio dos dados abertos, apresenta três características que são fundamentais neste caso: i) disponibilidade e acesso: os dados devem ser disponibilizados de forma completa, conveniente e modificável e contendo apenas custos de reprodução, preferencialmente disponibilizados pela internet; ii) reutilização e redistribuição: os dados devem ser fornecidos mediante licença que permita a reutilização, a redistribuição e a combinação com outros conjuntos de dados; e, iii) participação universal: todos devem ser capazes de usar, reutilizar e redistribuir, sem discriminação das áreas de atuação, pessoas ou grupos (ISOTANI; BITTENCOURT, 2015; OKF, 2019).

A jeito mais simples e eficaz de disponibilizar os dados na internet é publicando-os em sua forma bruta. Contudo, para facilitar sua reutilização é necessário que estes estejam devidamente estruturados e em formatos que permitam o uso automatizado, evitando estruturas que concedam apenas a visualização das informações disponibilizadas (W3C, 2009).

A importância da conexão entre os dados abertos refere-se ao fato de que novas informações somente poderão ser obtidas a partir do relacionamento em diferentes contextos ou valores semânticos. Com a evolução dos serviços disponíveis na internet, a quantidade de informações aumentou exponencialmente. Todavia, a maioria dos dados não está estruturada de forma adequada para o acesso, compreensão, manipulação e para geração de novas informações e conhecimentos. Nessa perspectiva e devido ao aumento da quantidade de dados sem a estrutura adequada, a obtenção de novas informações exige meios cada vez mais complexos (KLEIN; LUCIANO; MACADAR, 2015; SILVA FILHO; BRANDI, 2019).

Os dados abertos estendem-se por esferas sociais, como educação, saúde, economia, governo, entre outros. No contexto governamental, para incentivar a abertura dos dados, Bernes-Lee (2006) propôs uma classificação de dados abertos conectando-os por meio de um Sistema de 5 Estrelas¹. Esse sistema classifica a etapa de abertura de dados mediante a quantidade de estrelas, sendo: (i) uma estrela para os dados disponíveis na internet com licença aberta; (ii) duas estrelas para dados abertos estruturados e legíveis por máquinas; (iii) três estrelas para dados abertos estruturados, legíveis por máquinas e disponibilizados em um formato livre; (iv) quatro estrelas para dados que atendem às regras da terceira estrela e que utilizem os formatos abertos do W3C; e (v) cinco estrelas para os dados que atendem às regras das estrelas anteriores e ainda vinculem-se seus dados a dados de terceiros para melhor contextualizá-los.

Os dados abertos são fomentados pelo governo brasileiro. Nessa situação, recebe nomenclatura específica, consistente no termo “*dados abertos governamentais*”. Isto decorre por conta da perspectiva normativa nacional, pois a Constituição da República Federativa do Brasil de 1988 assegura à sociedade o direito de receber informações originárias das instituições públicas.

¹ A publicação inicial do documento “Linked Data” por Tim Bernes-Lee remete-se ao ano de 2006, no entanto o “Sistema de 5 Estrelas” foi acrescentado ao documento original no ano de 2010.

No plano infraconstitucional, tem-se a Lei de Acesso à Informação – LAI (Lei n. 12.527, publicada em 2011). Com essa publicação normativa o Brasil se tornou o décimo terceiro país da América Latina e o nonagésimo primeiro do mundo a permitir mecanismos concretos que assegurem o direito de acesso a informações públicas. A partir de 2012, com a entrada em vigor da referida lei, as instituições públicas federais, estaduais e municipais têm a obrigação de disponibilizar informações de forma ativa. A LAI determina que as informações de interesse coletivo ou geral devem ser, obrigatoriamente, acessíveis nos sítios oficiais da rede mundial de computadores, com ferramentas que permitam a obtenção a informação de forma objetiva, transparente e de fácil compreensão. O acesso automatizado por sistemas externos em formatos abertos também é garantido pela lei (BRASIL, 2011; CAROSI; TEIXEIRA FILHO, 2017).

Em decorrência da implementação normativa, uma quantidade significativa de informações públicas tornou-se disponível, sem restrições para o uso e a redistribuição. Os portais que disponibilizam os dados abertos governamentais (DAG) colaboram na publicação de elementos essenciais do conhecimento, compartilhando os dados localizáveis, acessíveis, reutilizáveis e legíveis por máquina, contribuindo com a eficiência das análises e de novas oportunidades.

Com isto, novos modelos de relacionamento entre as instituições públicas e a sociedade têm emergido, transformando a relação entre governos e cidadãos, na qual os governantes buscam novas formas de gestão e de ferramentas para atender às exigências da sociedade (ALBANO; CRAVEIRO, 2015; KLEIN; KLEIN; LUCIANO, 2018).

Os dados abertos governamentais, no plano nacional brasileiro, apresentam-se como uma resposta do governo as demandas sociais por transparência e controle das ações praticadas. No que é pertinente à transparência, às demandas da sociedade giram em torno da utilização e destinação dos recursos públicos e controle sobre os serviços prestados.

Nesse contexto, os dados abertos propiciam ambiente benéfico a todos, possibilitando maior e melhor transparência sobre a gestão pública, gerando novos produtos e serviços por meio da interação entre os diversos setores da sociedade e o governo. O movimento pela transparência é fortalecido com as iniciativas de governo aberto e com a Lei de Acesso à Informação e ampliam de modo considerável a quantidade de dados abertos disponibilizados para a sociedade (ALBANO; REINHARD, 2015; FREITAS *et al.*, 2018).

Apesar dos grandes avanços na ampliação da transparência, as iniciativas governamentais para prestação de contas com a disponibilidade de dados abertos ainda são esporádicas. Assim sendo, vários setores da sociedade ainda anseiam por soluções efetivas para consolidar informações claras, acessíveis e a interação entre governantes e cidadãos. Portanto, a transparência aferida por meio dos dados abertos governamentais pode ser aprimorada a partir das considerações de especialistas e usuários, ampliando a prestação pública de contas. Esta evolução na forma de prestar contas poderá, inclusive, responsabilizar servidores por suas ações ou omissões (KLEIN; KLEIN; LUCIANO, 2019; MOREIRA *et al.*, 2017; SOARES *et al.*, 2017).

Vale ressaltar que com aplicação de técnicas e ferramentas adequadas é possível gerar informações úteis e novos conhecimentos a partir dos dados disponibilizados, além de contribuir com a melhoria dos próprios conjuntos de dados. A abertura dos dados contribui para a melhoria dos serviços públicos, com a gestão mais efetiva dos recursos, com o aumento da responsabilização e da prestação de contas, da mesma forma que seu uso pode ser realizado pela sociedade em novos projetos ou integrados a novos produtos ou serviços. (KLEIN; KLEIN; LUCIANO, 2019; OLIVEIRA; LÓSCIO, 2014; VICTORINO *et al.*, 2017).

Embora existam alguns estudos sobre a temática, a realidade brasileira ainda é incipiente. Tendo em vista esta premissa, este artigo utiliza como fonte de evidência a produção científica publicada em revistas nacionais e internacionais com foco no contexto brasileiro, conforme é detalhado na metodologia.

PROCEDIMENTOS METODOLÓGICOS

Trata-se de uma pesquisa exploratória utilizando uma revisão sistemática da literatura, em publicações disponíveis nas bases de dados da EBSCO, SciELO, Science Direct, Scopus e Web of Science utilizando o termo “dados abertos”, sem restrição temporal. A seleção dessas bases foi feita pelos critérios escopo e cobertura. As buscas resultaram na identificação de 41 artigos, dos quais sete foram descartados por duplicidade, restando 34 artigos. Posteriormente, foi realizada a leitura completa dos artigos com o objetivo de identificar a existência de uma definição do termo de “dados abertos” e a vinculação com o contexto brasileiro. Nessa etapa foram descartados 17 artigos, que não se enquadraram no critério de seleção. A busca na base de dados *Science Direct* resultou em dois artigos, no entanto eles não foram utilizados na análise por não se enquadrar no contexto brasileiro. Assim, foram selecionados 17 artigos para os procedimentos de codificação e análises, conforme apresentado no quadro 1.

A etapa seguinte de análise consistiu na identificação, codificação e categorização dos conceitos vinculados ao termo “dados abertos”. Esta etapa foi realizada com o auxílio do software NVivo 12 Plus, que facilita a organização, a investigação e a busca de informações em dados não estruturados (QSR INTERNATIONAL, 2019). Apesar de o software utilizado possibilitar a codificação automática, a análise qualitativa baseou-se na metodologia da teoria fundamentada em dados, identificando primeiramente os conceitos em foco e categorizando cada segmento dos textos relacionados ao tema. Posteriormente, os códigos iniciais foram reclassificados e sintetizados em uma codificação focalizada com termos mais significativos, representando grande quantidade de dados. Esta etapa definiu a estrutura para as análises e reflexões sobre o tema em estudo, realizada por meio do processo de descoberta e codificação dos termos relacionados aos dados abertos (CHARMAZ, 2009).

Quadro 1 – Publicações por ano e base de dados

Ano	Quantidade Artigos	Autores	Bases de Dados pesquisadas			
			EBSCO	SciELO	Scopus	Web of Science
2013	1	Santarem Segundo	X		X	
2014	1	Oliveira e Lóscio	X			
2015	4	Albano e Craveiro	X			
		Albano e Reinhard	X			
		Klein, Luciano e Macadar	X			
		Lobo, Maia e Parreiras	X			
2016	1	Carossi e Teixeira Filho	X			
2017	4	Moreira et al		X	X	X
		Rautenberg et al		X	X	
		Soares et al	X			
		Victorino et al			X	
2018	3	Freitas et al			X	
		Gonçalves e Gama	X			
		Klein, Klein e Luciano		X		
2019	3	Klein, Kleine Luciano	X			
		Penteado, Bittencourt e Isotani	X			
		Silva Filho e Brandi	X			

Fonte: elaborado pelos autores (2019)

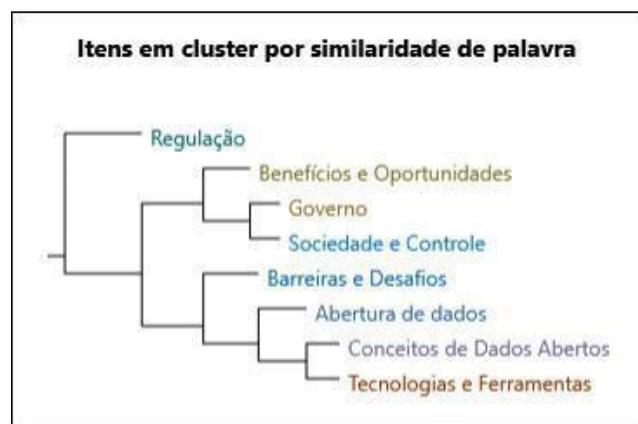
- 2) Governo: abertura de dados na esfera pública; governo aberto e governo eletrônico; informações governamentais; incentivos de integração com a sociedade; transparência e prestação de contas públicas;
- 3) Sociedade e controle: controle social das contas públicas; manifestação de interesses da sociedade nos dados públicos; participação efetiva na sociedade para aprimorar a abertura de dados governamentais;
- 4) Tecnologia e ferramentas: importância das tecnologias de informação e comunicação no contexto de dados abertos; tecnologias ligadas à abertura e disponibilização dos dados; formatos para divulgação e disponibilização dos dados;
- 5) Benefícios e oportunidades: as possibilidades de melhoria e inovação por meio da disponibilização dos dados abertos; oportunidade de novos produtos e serviços; a eficiência nas prestações de serviços público; a construção de conhecimento aberto;
- 6) Regulação: o papel da Lei de Acesso à Informação na abertura de dados governamentais; as informações disponibilizadas pelo Sistema Eletrônico do Serviço de Informações ao Cidadão; levantamento de normativas que influenciam nas questões de abertura de dados;
- 7) Abertura de dados: movimentos voltados para a abertura da dados; a função dos intermediários na disponibilização das informações para a sociedade; as políticas e estratégias adotadas na abertura dos dados; e
- 8) Barreiras e desafios: a complexidade do processo para disponibilização dos dados; as limitações institucionais; dúvidas ou ausências de normas que orientem claramente sobre dados abertos; qualidade da informação disponibilizada; criação de técnicas adequadas; reconhecimento da utilidade dos dados abertos.

O quadro 2 apresenta a quantidade e a identificação dos autores dos artigos utilizados na consolidação de cada uma das categorias de acordo com os temas identificados.

A análise de *cluster* das oito categorias identificadas e o conteúdo dos 45 nós elaborados pelo software NVivo utilizando a métrica de similaridade do coeficiente de correlação de Pearson apresentaram como resultado um diagrama de similaridade de palavras (gráfico 1), no qual as categorias com maior grau de similaridade (com base na frequência de palavras codificadas nos artigos) estão agrupadas com maior proximidade, enquanto aquelas com grau de semelhança menor estão mais distantes no diagrama.

Assim, evidencia-se a similaridade encontrada entre as categorias de “conceitos de dados abertos”, de “tecnologias e ferramentas” e de “abertura de dados”. Esse agrupamento fortalece a importância dos avanços nas tecnologias da informação, da abertura e disponibilização dos dados abertos, assim como a necessidade de compreender as interações entre os recursos disponíveis e as políticas de abertura de dados (ALBANO; CRAVEIRO, 2015; FREITAS *et al.*, 2018).

Gráfico 1 – Diagrama de similaridade de palavras elaborado a partir das principais codificações relacionadas aos dados abertos



Fonte: Elaborado pelos autores (2019)

Quadro 2 – Artigos utilizados para formação das categorias

Categorias consolidadas	Quantidade de artigos	Artigos codificados
Conceitos de dados abertos	17	Albano e Craveiro (2015); Albano e Reinhard (2015); Carossi e Teixeira Filho (2017); Freitas et al (2018); Gonçalves e Gama (2018); Klein, Klein e Luciano (2018); Klein, Klein e Luciano (2019); Klein, Luciano e Macadar (2015); Lobo, Maia e Parreiras (2015); Moreira et al (2017); Oliveira e Loscio (2014); Penteado, Bittencourt e Isotani (2019); Rautenberg et al (2017); Santarem Segundo (2013); Silva Filho e Brandi (2019); Soares et al (2017); Victorino et al (2017).
Governo	16	Albano e Craveiro (2015); Albano e Reinhard (2015); Carossi e Teixeira Filho (2017); Freitas et al (2018); Gonçalves e Gama (2018); Klein, Klein e Luciano (2018); Klein, Klein e Luciano (2019); Klein, Luciano e Macadar (2015); Lobo, Maia e Parreiras (2015); Moreira et al (2017); Oliveira e Loscio (2014); Penteado, Bittencourt e Isotani (2019); Santarem Segundo (2013); Silva Filho e Brandi (2019); Soares et al (2017); Victorino et al (2017).
Sociedade e controle	14	Albano e Craveiro (2015); Albano e Reinhard (2015); Carossi e Teixeira Filho (2017); Freitas et al (2018); Klein, Klein e Luciano (2018); Klein, Klein e Luciano (2019); Klein, Luciano e Macadar (2015); Moreira et al (2017); Oliveira e Loscio (2014); Penteado, Bittencourt e Isotani (2019); Santarem Segundo (2013); Silva Filho e Brandi (2019); Soares et al (2017); Victorino et al (2017).
Tecnologia e ferramentas	13	Albano e Craveiro (2015); Albano e Reinhard (2015); Freitas et al (2018); Klein, Klein e Luciano (2018); Klein, Luciano e Macadar (2015); Lobo, Maia e Parreiras (2015); Moreira et al (2017); Oliveira e Loscio (2014); Penteado, Bittencourt e Isotani (2019); Rautenberg et al (2017); Santarem Segundo (2013); Silva Filho e Brandi (2019); Victorino et al (2017).
Benefícios e oportunidades	13	Albano e Craveiro (2015); Albano e Reinhard (2015); Freitas et al (2018); Gonçalves e Gama (2018); Klein, Klein e Luciano (2018); Klein, Klein e Luciano (2019); Klein, Luciano e Macadar (2015); Moreira et al (2017); Oliveira e Loscio (2014); Penteado, Bittencourt e Isotani (2019); Santarem Segundo (2013); Soares et al (2017); Victorino et al (2017).
Regulação	10	Albano e Craveiro (2015); Albano e Reinhard (2015); Carossi e Teixeira Filho (2017); Freitas et al (2018); Gonçalves e Gama (2018); Luciano e Macadar (2015); Moreira et al (2017); Oliveira e Loscio (2014); Penteado, Bittencourt e Isotani (2019); Santarem Segundo (2013).
Abertura de dados	09	Albano e Craveiro (2015); Carossi e Teixeira Filho (2017); Freitas et al (2018); Klein, Klein e Luciano (2018); Moreira et al (2017); Oliveira e Loscio (2014); Penteado, Bittencourt e Isotani (2019); Santarem Segundo (2013); Soares et al (2017).
Barreiras e desafios	06	Albano e Craveiro (2015); Albano e Reinhard (2015); Klein, Klein e Luciano (2019); Penteado, Bittencourt e Isotani (2019); Santarem Segundo (2013); Soares et al (2017).

Fonte: Elaborado pelos autores (2019).

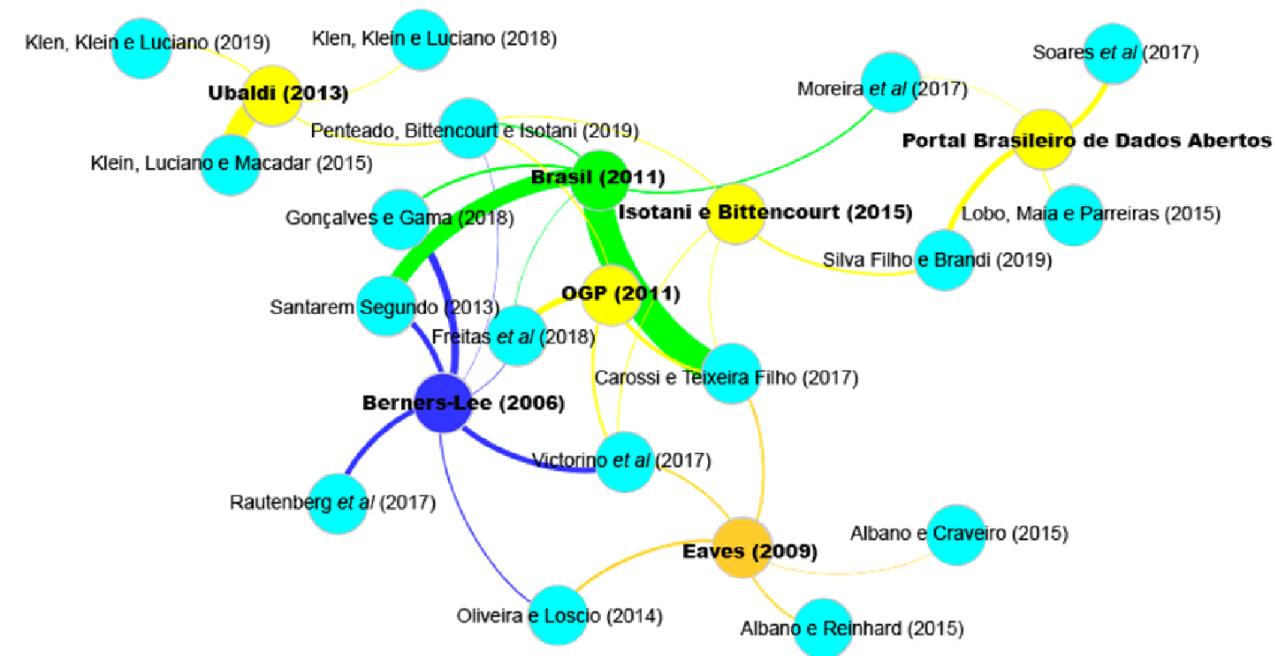
Ainda com relação ao gráfico 1, destaca-se a aproximação entre as categorias “sociedade e controle”, “governo” e “benefícios e oportunidades”, correspondendo às possibilidades de geração de novos produtos e serviços, úteis tanto para o governo quanto para a sociedade (ALBANO; REINHARD, 2015; PENTEADO; BITTENCOURT; ISOTANI, 2019).

Observa-se que os benefícios apresentados a partir da abertura de dados aumenta a eficiência das atividades governamentais e fortalecem o desenvolvimento do setor econômico com a geração de novos produtos, trazendo vantagens para as entidades públicas e a sociedade como um todo (SANTAREM SEGUNDO, 2013; VICTORINO *et al.*, 2017).

O gráfico 2 representa a rede estabelecida entre artigos estudados ligados por meio de referências comuns, nos quais sete referências são responsáveis por conectar os dezessete artigos estudados que apresentaram conceitos sobre dados abertos. A rede completa de referências utilizadas para contextualizar sobre o tema relacionado aos dados abertos é composta por 17 artigos e 165 referências. No entanto, no gráfico 2 estão destacadas apenas as sete referências necessárias para interligar todos os artigos deste estudo, as quais foram analisadas de forma detalhada devido à importância para o tema.

O principal nó entre os artigos na rede apresentada é o autor Berners-Lee (2006), referenciado em sete artigos, que propôs uma classificação de dados abertos conectados por meio do Sistema de 5 Estrelas para incentivar a abertura dos dados, especialmente aqueles governamentais. Os autores Santarem Segundo (2013), Oliveira e Lóscio (2014), Rautenberg *et al.* (2017), Victorino *et al.*, (2017), Freitas *et al.* (2018), Gonçalves e Gama (2018) e Penteadó, Bittencourt e Isotani (2019) utilizaram o autor Berners-Lee (2006) como fonte de fundamentação em suas pesquisas. Nos artigos Rautenberg *et al.* (2017), Victorino *et al.*, (2017), Freitas *et al.* (2018) e Gonçalves e Gama (2018), Berners-Lee é utilizado para fundamentar o modelo de classificação do Sistema de 5 Estrelas.

Gráfico 2 – Rede de artigos e principais referências ao tema de dados abertos



Legendas:

- Referência utilizada em sete artigos analisados
- Referência utilizada em seis artigos analisados
- Referência utilizada em cinco artigos analisados
- Referência utilizada em quatro artigos analisados
- Artigos analisados

Fonte: Os Autores com base nos dados da pesquisa (2019)

Penteado, Bittencourt e Isotani (2019) comparam brevemente o sistema de classificação proposto por Berners-Lee (2006) com o modelo Melhores Práticas para Dados na Web (DWBP). Os autores Santarem Segundo (2013) e Oliveira e Lóscio (2014) não mencionam o sistema de classificação dos dados abertos e utilizam Berners-Lee para fundamentar os conceitos de dados conectados.

O segundo nó com maior destaque na rede estabelecida entre artigos estudados refere-se à Lei de Acesso à Informação – LAI (BRASIL, 2011). A LAI foi utilizada em seis artigos para fundamentar as questões relacionadas à transparência e obrigatoriedade das instituições governamentais em fornecer as informações solicitadas pela sociedade. A LAI foi utilizada pelos autores Santarem Segundo (2013), Carossi e Teixeira Filho (2017), Moreira *et al.* (2017), Freitas *et al.* (2018), Gonçalves e Gama (2018) e Penteado, Bittencourt e Isotani (2019). Nota-se que quatro autores que utilizaram Berners-Lee (2006) também estão no grupo daqueles que utilizaram a Lei de Acesso à Informação.

Eaves (2009) é a terceira fonte de referência, sendo utilizado em cinco artigos: Oliveira e Lóscio (2014), Albano e Craveiro (2015), Albano e Reinhard (2015), Carossi e Teixeira Filho (2017) e Victorino *et al.*, (2017). Albano e Craveiro (2015) e Albano e Reinhard (2015) utilizam o autor para fundamentar a importância dos dados abertos estarem disponíveis na internet e com a possibilidade de serem reutilizados livremente por terceiros. Os autores Oliveira e Lóscio (2014), Carossi e Teixeira Filho (2017) e Victorino *et al.*, (2017) baseiam-se em três leis desenvolvidas por Eaves (2009) constituídas como: i) se o dado não pode ser encontrado e indexado na internet ele não existe; ii) se o dado não possuir uma licença aberta e disponível em formato compreensível por máquina, o dado não poderá ser reaproveitado; e, iii) se o dado apresentar qualquer restrição de uso não poderá ser considerado como dado aberto.

As quatro últimas referências responsáveis por interligar todos os artigos estudados são *Open Government Partnership* - OGP (2011), Ubaldi (2013), Isotani e Bittencourt (2015) e Portal Brasileiro de Dados Abertos (Brasil, 2019), e cada uma está ligada a também quatro artigos. A OGP (2011) é referenciada pelos autores nos temas relacionados à declaração internacional de abertura de dados governamentais. O Brasil integra essa iniciativa desde 2011, e com isso assumiu o compromisso para o aumento da disponibilidade de informações sobre as atividades governamentais.

Isotani e Bittencourt (2015) apresentam a maior diversidade de fundamentações utilizadas pelos autores. Carossi e Teixeira Filho (2017) utilizam Isotani e Bittencourt (2015) para fundamentar que quaisquer pessoas, físicas ou jurídicas, podem usar e redistribuir as informações.

Victorino *et al.* (2017) recorrem aos autores para trazer o conceito de dados conectados. Já Silva Filho e Brandi (2019) ressaltam que a maioria dos dados não estão estruturados, e como consequência a extração de informações e a produção de novos conhecimentos não ocorrem com a agilidade e eficácia necessárias. Já Penteado, Bittencourt e Isotani (2019) focam no fato que a liberação de dados em formatos abertos e sem restrições pelo governo para a sociedade resulta em novas análises produtos e serviços que beneficiam a própria sociedade.

Em relação ao Portal Brasileiro de Dados Abertos (BRASIL, 2019), Lobo, Maia e Parreiras (2015) utilizam uma definição de dado abertos da *Open Knowledge Foundation*, enquanto Moreira *et al.* (2017) afirmam que o Portal de Brasileiro de Dados Abertos é fruto do compromisso firmado com a *Open Government Partnership*. Os autores Soares *et al.* (2017) e Silva Filho e Brandi (2019) indicam a ferramenta e a diversidade de dados disponíveis no portal.

Por fim, Ubaldi (2013) é utilizado por Klein, Luciano e Macadar (2015) para apontar questões relacionadas à transparência, prestação de contas públicas, princípios sobre dados abertos governamentais, os benefícios do uso dos dados para gerar inovações e desenvolver a economia. Klein, Klein e Luciano (2018 e 2019) destacam a importância dos dados abertos para projetos desenvolvidos pela sociedade, integrando novos produtos, aplicativos ou serviços. Completando, Penteadó, Bittencourt e Isotani (2019) afirmam que os dados abertos podem auxiliar a sociedade a compreender as atividades do governo e identificar seu desempenho.

CONSIDERAÇÕES FINAIS

Este trabalho buscou apresentar uma revisão de publicações científicas sobre dados abertos, com foco no contexto brasileiro, sem restrição temporal, com o intuito de mostrar os conceitos, a rede de autores e as principais referências utilizadas.

Para a realização desses objetivos, o artigo reuniu bases conceituais e temas correlatos. Os conceitos de dados abertos e de termos a eles relacionados, que foram identificados mediante busca nas bases de dados sobre esses temas foram: dados abertos; dados conectados; dados governamentais abertos; e transparência e prestação de contas públicas. Utilizou-se a análise de conteúdo qualitativo das publicações científicas relacionadas à abertura de dados e empregou-se a codificação baseada na teoria fundamentada em dados.

A contribuição deste estudo reside na identificação dos conceitos de dados abertos, os quais podem ser compreendidos como dados disponíveis a quaisquer pessoas, físicas ou jurídicas, para o acesso, utilização, manipulação, novos compartilhamento e elaboração de novos produtos ou serviços. Uma possível condição para a utilização dos dados abertos é a atribuição da fonte e a manutenção da licença original. Destaca-se que o compartilhamento de dados em formato aberto favorece inúmeras oportunidades e benefícios, para o desenvolvimento de novos serviços e produtos, contribuindo com o desenvolvimento social e econômico e a ampliação da transparência.

O presente estudo também traz os principais temas correlatos aos dados abertos obtidos a partir da codificação das publicações analisadas, apresentando as palavras mais frequentes e as principais categorias utilizadas pelos autores nas pesquisas, as quais podem ser definidas como: i) conceitos de dados abertos; ii) governo; iii) sociedade e controle; iv) tecnologias e ferramentas; v) benefícios e oportunidades; vi) regulação; vii) abertura de dados; e, viii) barreiras e desafios. A outra contribuição se refere à associação da rede de referências (autores) utilizadas na fundamentação dos estudos analisados. Identificou-se que sete referências são responsáveis por conectar os 17 artigos estudados que apresentaram conceitos sobre dados abertos.

Verificou-se também que somente após a promulgação da LAI em 2011 iniciaram-se as publicações de artigos sobre o tema, sendo o primeiro em 2013. Passou-se de uma publicação por ano em 2013 e 2014, para uma média de três artigos nos anos subsequentes. Apesar do aumento de publicações nessa área, ainda existem lacunas a serem preenchidas, principalmente estudos que sejam capazes de identificar e analisar a produção de conhecimento a partir do uso de dados abertos.

A partir da elaboração da rede de referências estabelecida entre artigos analisados, constata-se a existência de um referencial em comum utilizado na elaboração dos artigos, sendo que com apenas sete referências torna-se possível por conectar os 17 artigos que apresentaram conceitos sobre dados abertos. As sete principais referências encontradas a partir da codificação das publicações analisadas foram: i) Berners-Lee (2006); ii) Lei de Acesso à Informação – LAI (BRASIL, 2011); iii) Eaves (2009); iv) Open Government Partnership - OGP (2011); v) Ubaldi (2013); vi) Isotani e Bittencourt (2015), e; vii) Portal Brasileiro de Dados Abertos (Brasil, 2019).

Outra lacuna a ser preenchida em estudos futuros seria o desenvolvimento de uma política pública de informação que permita a implementação de práticas científicas orientadas à gestão de dados de pesquisa. Com isto seria possível maximizar os recursos tão escassos para a pesquisa no Brasil.

REFERÊNCIAS

- ALBANO, C. S.; CRAVEIRO, G. da S. Lições Aprendidas com a utilização de Dados Orçamentários em Formato Aberto: Um estudo exploratório no ecossistema Brasileiro. *Revista de Gestão e Projetos*, [S.l.], v.6, n.3, p.17–27, 2015. Disponível em: <http://www.revistagep.org/ojs/index.php/gep/article/view/370>. Acesso em: 2 maio. 2019.
- ALBANO, C. S.; REINHARD, N. Desafios para Governos e Sociedade no Ecossistema Brasileiro de Dados Governamentais Abertos. *Cadernos Gestão Pública e Cidadania*, [S.l.], v.20, n.67, p.214–235, 2015. Disponível em: <http://bibliotecadigital.fgv.br/ojs/index.php/cgpc/article/view/41150>. Acesso em: 29 maio. 2019.
- BERNERS-LEE, T. *Linked Data*. [S.l.], 27 July 2006. Disponível em: <https://www.w3.org/DesignIssues/LinkedData.html>. Acesso em: 4 jun. 2019.
- BRASIL. *Lei nº 12.527, de 18 de novembro de 2011*. Regula o acesso a informações previsto no inciso XXXIII do art. 5º, no inciso II do § 3º do art. 37 e no § 2º do art. 216 da Constituição Federal; altera a Lei nº 8.112, de 11 de dezembro de 1990; revoga a Lei nº 11.111, de 5 de maio de 2005, e dispositivos da Lei nº 8.159, de 8 de janeiro de 1991; e dá outras providências. Brasília, DF: Presidência da República, 2011. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/l12527.htm. Acesso em: 4 jun. 2019.
- BRASIL. *Portal Brasileiro de Dados Abertos*. [S.l.], 2019. Disponível em: <http://dados.gov.br/>. Acesso em: 4 jun. 2019.
- CAROSI, D. F.; TEIXEIRA FILHO, J. G. de A. Uma Análise dos Pedidos de Acesso à Informação Encaminhados a uma Instituição de Ensino Superior. *Gestão. Org.*, [S.l.], v.14, p.255–264, 2017. Edição especial. Disponível em: <https://periodicos.ufpe.br/revistas/gestaoorg/article/view/22551>. Acesso em: 20 maio. 2019.
- CHARMAZ, K. *A construção da teoria fundamentada*: guia prático para análise qualitativa. Porto Alegre: Artmed, 2009.
- EAVES, D. *The three laws of open government data*. [S.l.], 30 Sept. 2009. Disponível em: <http://eaves.ca/2009/09/30/three-law-of-open-government-data/>. Acesso em: 4 jun. 2019.
- FREITAS, J. A. C. et al. O ecossistema de dados abertos do governo federal: Um estudo sobre a composição e desafios. *Ciência da Informação*, Brasília, v.47, n.2, p.110–132, 2018. Disponível em: <http://revista.ibict.br/ciinf/article/view/3952>. Acesso em: 2 maio. 2019.
- GEPHI.ORG. *Features*. [S.l.], 2019. Disponível em: <https://gephi.org/features/>.
- GONÇALVES, B. A.; GAMA, K. S. Transparência e dados abertos do Recife: Uma estratégia bem sucedida de publicação. *Cadernos BAD*, Lisboa, n.1, p.157–164, 2018. Disponível em: <https://www.bad.pt/publicacoes/index.php/cadernos/article/view/1901/pdf>. Acesso em: 2 maio. 2019.
- ISOTANI, S.; BITTENCOURT, I. I. *Dados abertos conectados*. Núcleo de informação e coordenação do ponto br – NIC.br Centro de Estudos sobre Tecnologia Web – CeWeb.br, São Paulo, v.1, p.175, 2015. Disponível em: https://www.researchgate.net/publication/283569633_Dados_Abertos_Conectados. Acesso em: 4 jun. 2019.
- ISOTANI, S.; BITTENCOURT, I. I. **Dados abertos conectados [recurso eletrônico]**. 1. ed. São Paulo: Novatec, 2015. Disponível em: https://www.researchgate.net/publication/282218981_Dados_Abertos_Conectados_em_Busca_da_Web_do_Conhecimento. Acesso em: 4 jun. 2019.
- KLEIN, R. H.; KLEIN, D. C. B.; LUCIANO, E. M. Identificação de mecanismos para a ampliação da transparência em portais de dados abertos: uma análise no contexto brasileiro. *Cadernos EBAPE.BR*, [S.l.], v.16, n.4, p.692–715, 2018. Disponível em: <http://bibliotecadigital.fgv.br/ojs/index.php/cadernosebape/article/view/73241>. Acesso em: 5 maio 2019.
- KLEIN, R. H.; KLEIN, D. C. B.; LUCIANO, E. M. Ampliação da transparência de dados abertos governamentais: a percepção dos observatórios sociais brasileiros. *Contextus – Revista Contemporânea de Economia e Gestão*, [S.l.], v.17, n.1, p.8-36, 2019. Disponível em: <http://www.periodicos.ufc.br/contextus/article/view/32834>. Acesso em: 19 maio. 2019.
- KLEIN, R. H.; LUCIANO, E. M.; MACADAR, M. A. Grau de transparência de dados abertos governamentais do site dados.rs.gov.br. *Revista Economia & Gestão*, [S.l.], v.15, n.41, p.256–285, 2015. Disponível em: <http://periodicos.pucminas.br/index.php/economiaegestao/article/view/9965>. Acesso em: 29 maio. 2019.
- LOBO, A. S. M.; MAIA, L. C. G.; PARREIRAS, F. S. O impacto do uso de ferramenta de visualização de dados abertos como recurso de ensino e aprendizagem. *Caderno de Geografia*, [S.l.], v.25, n.44, p.97–117, 2015. Disponível em: <http://periodicos.pucminas.br/index.php/geografia/article/view/9057>. Acesso em: 2 maio. 2019.
- MACHADO, J. Dados abertos e ciência aberta. In: ALBAGLI, S.; MACIEL, M. L.; ABDO, A.H. (org.). *Ciência aberta*: questões abertas. Brasília: IBICT; Rio de Janeiro: UNIRIO, 2015. cap. 9, p. 201-227.
- MOREIRA, F. M. et al. Metadados para descrição de datasets e recursos informacionais do “Portal Brasileiro de Dados Abertos”. *Perspectivas em Ciência da Informação*, Belo Horizonte, v.22, n.3, p.158–185, 2017. Disponível em: <http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/2947>. Acesso em: 2 maio. 2019.
- OPEN KNOWLEDGE FOUNDATION - OKF. *Guia de Dados Abertos*. [S.l.], 2019. Disponível em: http://opendatahandbook.org/guide/pt_BR/. Acesso em: 4 jun. 2019.
- OLIVEIRA, L. E. R. de A.; LÓSCIO, B. F. Uma abordagem para captura de informações sobre aplicações que fazem uso de dados abertos. *Revista Brasileira de Administração Científica*, [S.l.], v.5, n.2, p.127–140, 2014. Disponível em: <http://www.sustenera.co/index.php/rbadm/article/download/SPC2179-684X.2014.002.0010/471>. Acesso em: 28 maio 2019.

OPEN GOVERNMENT PARTNERSHIP - OGP. *Open government declaration*. [S.l.], 2011. Disponível em: <http://www.opengovpartnership.org/about/open-government-declaration>. Acesso em: 4 jun. 2019.

PENTEADO, B. E.; BITTENCOURT, I. I.; ISOTANI, S. Análise exploratória sobre a abertura de dados educacionais no Brasil: como torná-los prontos para o ecossistema da Web? *Revista Brasileira de Informática na Educação*, [S.l.], v.27, n.01, p.175-195, 2019. Disponível em: <http://br-ie.org/pub/index.php/rbie/article/download/7801/6039>. Acesso em: 29 maio. 2019.

QSR INTERNATIONAL. *NVivo*: o software n1 para análise qualitativa de dados. [S.l.], 2019. Disponível em: <http://www.qsrinternational.com/nvivo-portuguese>. Acesso em: 2 maio. 2019.

RAUTENBERG, S. *et al.* Dados abertos conectados e gestão do conhecimento: estudos de caso cientométricos em uma universidade brasileira. *Perspectivas em Ciência da Informação*, [S.l.], v.22, n.3, p.116-142, 2017. Disponível em: <http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/2885>. Acesso em: 2 maio. 2019.

SANTAREM SEGUNDO, J. E. Tecnologias de informação e comunicação para disponibilização de dados abertos em formato semântico. *Ibersid*, [S.l.], v.7, p.33-40, 2013. Disponível em: <https://www.ibernid.eu/ojs/index.php/ibernid/article/view/4075/3744>. Acesso em: 2 maio. 2019.

SAYÃO, L. F.; SALES, L. F. Dados de pesquisa: contribuição para o estabelecimento de um modelo de curadoria digital para o país. *Tendências da Pesquisa Brasileira em Ciência da Informação*, Belo Horizonte, v.6, n.1, 2013.

SILVA FILHO, V. B. Da; BRANDI, L. S. N. Um estudo focado ao PROUNI através da análise de dados abertos: período de 2005 até 2016. *Prisma.com*, [S.l.], n.38, p.37-53, 2019. Disponível em: <http://ojs.letras.up.pt/index.php/prismacom/article/view/5204>. Acesso em: 29 maio. 2019.

SOARES, E. *et al.* Sistema web para mapeamento de dados de crimes letais no Estado de Pernambuco. *Gestão.Org*, Pernambuco, v.14, p.288-295, 2017. Edição especial. Disponível em: <https://periodicos.ufpe.br/revistas/gestaoorg/article/view/22554/18703>. Acesso em: 2 maio 2019.

UBALDI, B. Open government data: towards empirical analysis of open government data initiatives. *OECD iLibrary*, [S.l.], n. 22, p.1-60, 2013. Disponível em: https://www.oecd-ilibrary.org/governance/open-government-data_5k46bj4f03s7-en. Acesso em: 4 jun. 2019.

VICTORINO, M. de C. *et al.* Uma proposta de ecossistema de big data para a análise de dados abertos governamentais conectados. *Informação e Sociedade*, João Pessoa, v.27, n.1, p.225-242, 2017. Disponível em: <http://www.periodicos.ufpb.br/ojs/index.php/ies/article/view/29299>. Acesso em: 2 maio. 2019.

W3C. *Publishing Open Government Data*. [S.l.], 2009. Disponível em: <https://www.w3.org/TR/gov-data/>. Acesso em: 29 maio. 2019.