

AVALIANDO A PRODUÇÃO CIENTÍFICA E A PUBLICAÇÃO DE DADOS CONECTADOS NA AMÉRICA LATINA E CARIBE

EVALUATING SCIENTIFIC PRODUCTION AND THE PUBLICATION OF DATA CONNECTED IN LATIN AMERICA AND THE CARIBBEAN

Larissa Mariany Freiburger Pereira¹, Guilherme Bertoni Machado², José Leomar Todesco³

Recebido em: 12/10/2018. Aprovado em: 14/10/2018.

Resumo

Desde 2006, quando Tim Berner-Lee criou o conceito de Dados Abertos Conectados, muitas pesquisas e iniciativas neste contexto começaram a ser desenvolvidas. Alguns países destacam-se na produção científica e publicação de dados governamentais utilizando tecnologia de Dados Abertos Conectados, como Estados Unidos, Inglaterra, Alemanha, etc. Esta pesquisa teve como objetivo fazer um levantamento da produção científica e publicação de Dados Abertos Conectados Governamentais na América Latina e no Caribe para compreender em que nível os países que compõem estas duas regiões estão em relação à tecnologia supracitada. O levantamento da produção científica ocorreu através de buscas em quatro bases de dados e a coleta de informações sobre a publicação de Dados Abertos Governamentais se deu a partir dos Portais de Dados Abertos dos países latino-americanos e caribenhos. Encontramos, portanto, uma produção científica baixa que poderia também ser evidenciada pela ausência de iniciativas do governo para publicar seus dados abertos usando a tecnologia de Dados Conectados.

Palavras-chave

dados abertos conectados; produção científica; publicação de dados abertos conectados governamentais; América Latina; Caribe.

Abstract

Since 2006, when Tim Berner-Lee designed the concept of Linked Open Data, many researches and initiatives in this context began to be developed. Some countries stand out in scientific production and publication of government data using Linked Open Data technology, such as United States, England, Germany, etc. This research aimed to survey scientific production and governmental publication of Linked Open Data in Latin America

¹ Mestranda no Programa de Pós-graduação em Engenharia e Gestão do Conhecimento (PPGEGC)/Universidade Federal de Santa Catarina/freiburger.lm@gmail.com/Brasil.

² Doutorando no Programa de Pós-graduação em Engenharia e Gestão do Conhecimento (PPGEGC)/Universidade Federal de Santa Catarina/gbertonimachado@gmail.com/Brasil.

³ Doutor em Engenharia de Produção/Universidade Federal de Santa Catarina/jose.todesco@ufsc.br/Brasil.

and the Caribbean to understand at what level the countries that comprise these two regions are in relation to the aforementioned technology. The survey of scientific production occurred from search four databases and the collection of information about the publication of Linked Open Government Data was observed from the Open Data Portal of the countries in Latin America and Caribbean. We find, therefore, a low scientific production that could be also evidenced through the lack of government initiatives to publish their open data using Linked Data technology.

Keywords

linked open data; scientific production; linked open government data publish; Latin America; Caribe.

1 Introdução

Em 2006, Tim Berner-Lee modelou o conceito de Linked Data ou Dados Conectados, sendo esta uma tecnologia que dá suporte à consolidação da Web de Dados. Também conhecida como Web Semântica, esta é uma Web onde os dados não são apenas publicados isoladamente, mas podem ser conectados para que pessoas e máquinas pudessem melhor explorá-los, através de conexões semânticas entre os mesmos (Berners-Lee, 2006; Bizer, Heath & Berners-Lee, 2009; Bizer, 2009; Heath & Bizer, 2011).

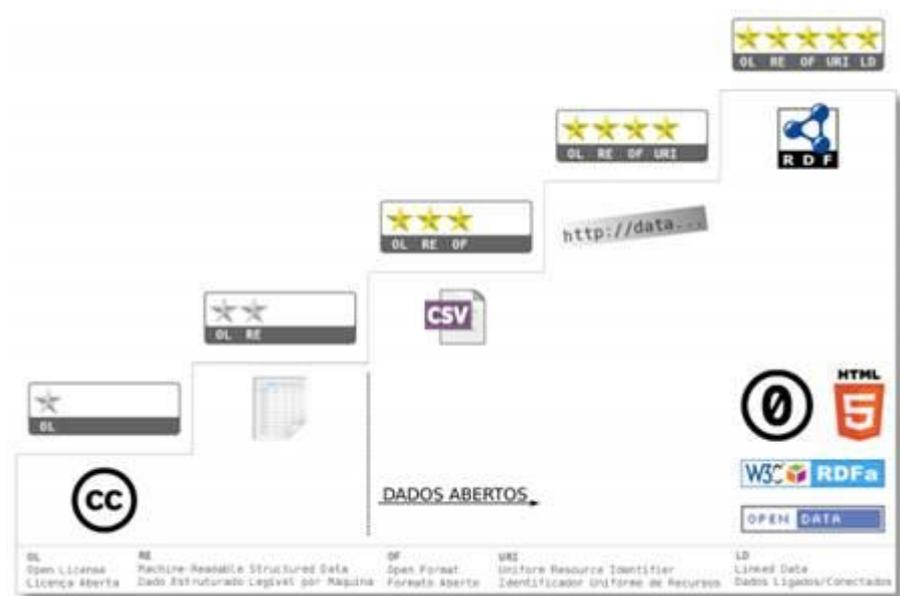
Mudando o conceito de publicação e consumo de dados na Web, Tim Berners-Lee traçou quatro princípios que gerem a utilização desta nova tecnologia (Berners-Lee, 2006):

1. Utilizar URIs (Uniform Resource Identifiers) para dar nome às coisas;
2. Utilizar HTTP URIs (HyperText Transfer Protocol) para que as pessoas possam pesquisar por esses nomes;
3. Quando alguém procurar um URI, fornecer informações úteis utilizando os padrões RDF (Resource Description Framework);
4. Incluir links para outros URIs, de modo que se possa descobrir mais coisas.

Com o contínuo crescimento da onda de dados abertos na Web, Berners-Lee (2006) propôs, ainda, uma classificação para dados abertos baseado em cinco estrelas (Figura 1).

1. Dados 1 estrela – Disponíveis na Web em qualquer formato, mas com licença aberta, para ser Dado Aberto; Dados 2 estrelas – Disponíveis em formato estruturado processável por máquina (por exemplo, excel ao invés de uma imagem de uma tabela);
2. Dados 3 estrelas – Como (2) + formato não-proprietário (por exemplo, CSV ao invés de excel);
3. Dados 4 estrelas – Dados 3 estrelas + uso de padrões estabelecidos pela W3C (RDF e SPARQL) para identificar coisas, de modo que as pessoas possam referenciá-los;
4. Dados 5 estrelas – Dados 4 estrelas + conexão de dados de diferentes datasets para prover contexto.

Figura 1 – Classificação 5 Estrelas para Dados Abertos



Fonte: Pereira, Machado & Todesco, 2015

Com o crescimento dos dados abertos que são publicados na Web e podendo estes dados serem publicados em um formato específico, padronizado pela W3C (RDF), e estando estes conectados entre si, observamos a construção de uma extensão da Web atual onde as relações semânticas estabelecidas entre diferentes datasets permitem a descoberta de conhecimento que, até então, não estava explícito, mas implícito entre os dados (Bizer, Heath, Idehen & Berners-Lee, 2008; Bizer, Cyganiak & Heath, 2007).

As iniciativas para publicação de dados utilizando a tecnologia de Dados Abertos Conectados vem crescendo desde então (Bizer, 2009) e, atualmente, o espaço de datasets conectados, denominado LOD Cloud ou Nuvem de Dados Abertos Conectados, conta com dados de variados tipos, como governamentais, geográficos, sobre ciências naturais, mídia, redes sociais, entre outros.

Em 2007, a LOD Cloud contava com 17 datasets publicados, já em 2014 este número estava em 570 (<http://lod-cloud.net>, recuperado em 13, junho, 2016). Este é um dado importante que reflete a importância que se tem dado à consolidação da WebSemântica, conforme apresentado na Tabela 1.

Tabela 1 – Crescimento da LOD Cloud

Data	Quantidade de Datasets
01/05/2007	12
01/03/2008	34
27/03/2009	93
22/09/2010	203
19/09/2011	295
30/08/2014	570

Fonte: Autoria Própria

Aqui ressaltamos a importância de iniciativas que tratem de novas abordagens para conversão de dados que se encontram em banco de dados relacionais para dados em formato RDF, tornando possível, assim, a construção de datasets que podem ser conectados uns com os outros (Auer, Dietzold & Lehmann, 2009). Essas iniciativas são de fundamental importância para a consolidação da Web de Dados, uma vez que, atualmente, a maior parte dos dados disponíveis da Web provém de banco de dados relacionais.

Tais iniciativas estão fortemente presentes especialmente em âmbitos governamentais, isso porque os governos, em geral, são grandes geradores de dados valiosos para a sociedade e, normalmente, não são utilizados e explorados com todo o potencial de informações que estes oferecem (Haslholfer & Isaac, 2011).

Com o objetivo de compreender como o tema “Dados Abertos Conectados” vem sendo tratado na América Latina e Caribe, fizemos uso, nesta pesquisa, de técnicas bibliométricas para fazer o levantamento da produção científica sobre o tema nessas regiões. Ainda, para atingirmos o objetivo inicial, observamos como os governos vêm utilizando esta nova tecnologia para publicarem seus dados, desta maneira, foi possível verificar em que escala, na prática, os dados governamentais vêm sendo publicados em formatos que propiciem a conexão destes com demais datasets.

O presente artigo está organizado da seguinte maneira: na seção 2 apresentamos o delineamento metodológico utilizado na condução desta pesquisa. Na seção 3 apresentamos a análise de dados realizada a partir da produção científica acerca de Dados Abertos Conectados na América Latina e Caribe. Na seção 4 apresentamos, igualmente, a análise realizada a partir da quantidade de datasets publicados em portais governamentais de países que compõem a América Latina e Caribe, disponibilizados em formatos padronizados pela W3C, que podem ser conectados com outros datasets. Na seção 5 apresentamos as principais considerações finais. Por fim, na seção 6, apresentamos o arcabouço teórico que nos auxiliou no desenvolvimento da pesquisa.

2 Procedimentos Metodológicos

Buscando quantificar pesquisas científicas e a própria produção de Dados Abertos Conectados na América Latina e Caribe, foi desenvolvida uma metodologia que será apresentada neste tópico, baseada em cinco etapas, conforme identificados na Figura 2.

Figura 2 – Delineamento Metodológico da Pesquisa



Fonte: Autoria Própria.

Inicialmente foram definidas as bases de dados em que seriam feitas as pesquisas dos artigos, a saber: Scopus, Web of Science e SciELO. Buscamos por artigos dos últimos dez anos (2006 – 2016) que continham em seus títulos, resumos ou palavras-chaves o termo “Linked Data” ou “Linked Open Data”, bem como o nome dos países que compõem a América Latina e a região do Caribe e os próprios termos “Latin America” e “Caribbean”.

Neste sentido, as String de busca utilizadas foram as seguintes:

1. “LATIN AMERICA” AND (“LINKED OPEN DATA” OR “LINKED DATA”)
2. “CARIBBEAN” AND (“LINKED OPEN DATA” OR “LINKED DATA”)
3. “NOME_DO_PAÍS” AND (“LINKED OPEN DATA” OR “LINKED DATA”)

Na Terceira String de busca substituímos o termo “NOME_DO_PAÍS” por cada um dos países que compõem a América Latina, bem como a região caribenha. Neste primeiro momento, obtivemos como resultado da pesquisa a quantidade de artigos apresentada na Tabela 2.

Tabela 2 – Resultado da Primeira Busca na Base de Dados

Base de Dados	Quantidade de Documentos
IEEE	4
Scopus	39
Web of Science	39
TOTAL	82

Fonte: Autoria Própria.

A partir dos resultados, seguimos para a segunda etapa. Nas próprias bases de dados geramos os arquivos para exportar os resultados para um software gerenciador de bibliografias, EndNote X7. Este software nos permitiu organizar os documentos para que pudéssemos seguir para a etapa seguinte. Nas bases de dados IEEE e Scopus, os arquivos foram gerados em formato .ris e na base de dados Web of Science os arquivos foram gerados no formato .ciw.

Na terceira etapa nós eliminamos todas as duplicidades dos resultados obtidos na busca nas bases de dados. A Universidade Federal de Santa Catarina dá aos seus pesquisadores acesso à diversas bases de dados, inclusive as três bases utilizadas nesta pesquisa. Apesar disso, alguns artigos não tinham disponíveis seu conteúdo completo nas bases de dados e estes também foram eliminados. Ainda nesta etapa fizemos a leitura dos títulos e resumos de todos os artigos cujo conteúdo estava disponível para acesso e eliminamos aqueles que não eram aderentes à proposta da pesquisa, com o objetivo de selecionarmos apenas artigos a serem analisados na etapa seguinte. Aqui ressaltamos que algumas áreas de pesquisa utilizam o termo “Linked Data” para caracterizar outros objetos de pesquisa, com a área da genética e biologia, por exemplo. Ao final deste processo, obtivemos o seguinte resultado, apresentado na Tabela 3:

Tabela 2 – Resultado do Processo de Seleção de Dados

Base de Dados	Quantidade de Documentos
IEEE	4
Scopus	10
Web of Science	2
TOTAL	16

Fonte: Autoria Própria.

PEREIRA, Larissa Mariany Freiberg; MACHADO, Guilherme Bertoni; TODESCO, José Leomar. Avaliando a produção científica e a publicação de dados conectados na América Latina e Caribe. **Revista Democracia Digital e Governo Eletrônico**, Florianópolis, v. 1, n. 17, p. 92-106, out. 2018.

Na quarta etapa fizemos a leitura criteriosa dos 16 documentos provenientes da etapa de seleção dos dados, com o intuito de analisa-los. Durante a análise, buscamos encontrar elementos que tratavam da produção e pesquisas acerca de Dados Abertos Conectados na América Latina e no Caribe, tais como pesquisas que utilizavam a tecnologia de Dados Abertos Conectados para construção de aplicações, análises da produção de mashups nestas regiões, revisões sistemáticas que tratassem do assunto em um nível teórico, avaliação da publicação de datasets em RDF (especialmente por governos), pesquisas que tratassem da conversão de datasets em formatos diversos para RDF em âmbito governamental, entre outros.

Consideramos que a quantidade de documentos que foram selecionados para a análise dos dados não seria suficiente para nos dar um panorama sobre o assunto. Diante disso, com o objetivo de quantificar melhor a publicação de Dados Conectados na América Latina e Caribe, decidimos visitar o portal governamental de cada um dos países que compõem as duas regiões citadas anteriormente e verificar a quantidade de datasets disponibilizados ao público, a partir disso, apurar quantos deles poderiam ser classificados como Dados Abertos Conectados.

Por fim, a partir da análise de todos os dados obtidos, trabalhamos no desenvolvimento do documento final para a apresentação dos resultados.

3 Produção científica acerca de dados abertos conectados

Após a etapa de coleta e seleção dos dados, foram obtidos 16 documentos para serem analisados. Trataram-se, portanto, de artigos publicados em congressos ou periódicos que discorressem sobre pesquisas desenvolvidas no contexto de dados abertos conectados em países da América Latina e Caribe (Tabela 4).

Tabela 4 – Pesquisas Desenvolvidas na América Latina e Caribe no contexto de Dados Abertos Conectados

Autores	Artigo	Instituições	Ano	Local de Publicação	País da Instituição
Base IEEE					
Ribeiro, C. E. Vivacqua, A. S.	A Framework for Composition and Reuse on the Linked Open Data Cloud	1. Universidade Federal do Rio de Janeiro	2013	7th Conference on Semantic Computing	1. Brasil
Piedra, N. Chicaiza, J. López, J. Caro, E. T.	Towards a Learning Analytics Approach for Supporting discovery and reuse of OER	1. Universidad Técnica Particular de Loja 2. Universidad Politécnica de Madrid	2015	Global Engineering Educations Conference (EDUCON)	1. Equador 2. Espanha

Parreiras, F. S. Gröner, G Schwabe, D. Silva, F. F.	Towards a Marketplace of Open Source Software Data	1. Universidade Fundação Mineira de Educação e Cultura 2. University of Duisburg-Essen 3. Pontifícia Universidade Católica do Rio de Janeiro	2015	48th Hawaii International Conference on System Sciences	1. Brasil 2. Alemanha 3. Brasil
Piedra, N. Chicaiza, J. López, J. Tovar, E.	Using linked open data to improve the search of open educational resources for engineering students	1. Universidad Técnica Particular de Loja 2. Universidad Politécnica de Madrid	2013	Frontiers in Education Conference (FIE)	1. Equador 2. Espanha
Base Scopus					
Ribeiro Junior, D. I.	Data repositories for e-Science: Open data, linked data and its technologies	1. Universidade Federal de Santa Catarina	2015	Ciência da Informação	1. Brasil
Santarem Segundo, J. E.	Information and Communication Technologies to provide Open Data in semantic formats	1. Universidade de São Paulo	2013	Ibersid	1. Brasil
Kauppinem, T Espindola, G. M. Jones, J. Sánchez, A. Gräler, B, Bartoschek, T.	Linked Brazilian Amazon Rainforest Data	1. University of Muenster 2. Instituto Nacional de Pesquisas Espaciais	2014	Semantic Web	1. Alemanha 2. Brasil
Rademaker, A. Oliveira, D. A. B. Paiva, V. Higuchi, S. Medeiros e Sá, A. Alvim, M.	A linked open data architecture for the historical archives of the Getulio Vargas Foundation	1. Fundação Getúlio Vargas 2. Nuance Communications	2015	International Journal on Digital Libraries	1. Brasil 2. Estados Unidos
Alves, M. V. C Bax, M. P.	Necessity and feasibility of adopting lod standard by national congress: A study in the context of the public budget	1. Fundação João Pinheiro 2. Universidade Federal de Minas Gerais	2014	Informação e Sociedade	1. Brasil 2. Brasil

Breitman, K. Viterbo, J. Salas, P. Saraiva, D. Magalhães, R. P. Gama, V. Casanova, M. A. Chaves, M. Franzosi, E.	Open government data in Brazil	1. Pontifícia Universidade Católica do Rio de Janeiro 2. Universidade Federal Fluminense 3. Universidade Federal do Ceará 4. Serviço Federal de Processamento de Dados 5. Ministério do Planejamento	2012	IEEE Intelligent Systems	1. Brasil 2. Brasil 3. Brasil 4. Brasil 5. Brasil
Speroni, R. M. Macedo, M. Gauthier, F. A. O.	Reference model for regional innovation indicators	1. Universidade Federal de Santa Catarina	2016	Revista Espacios	1. Brasil
Rodríguez-Enríquez, C. A. Alor-Hernández, G. Sánchez-Ramírez, C. Córtés-Robles, G.	Supply chain knowledge management: A linked data-based approach using skos	1. Instituto Tecnológico de Orizaba	2015	Revista DYNA	1. México
Cardoso, S. D. Amanqui, F. K. Serique, K. J. A. Santos, J. L. C. Moreira, D. A. Santos Neto, A. L.	SWI: A Semantic Web Interactive Gazetteer to support Linked Open Data	1. Universidade de São Paulo 2. Instituto Nacional de Pesquisas da Amazônia	2016	Future Generation Computer Systems	1. Brasil 2. Brasil
Marcondes, C. H. Pereira, D. V. Fonseca, E. R. Souza, I. V. P. Barbosa, N. Moraes, R. P. T. Martins, S. C.	Using open data technology to connect libraries, archives and museums: A machadian case	1. Universidade Federal Fluminense 2. Universidade Federal do Rio de Janeiro	2013	TransInformação	1. Brasil 2. Brasil
Base Web of Science					
Kauppinem, T. Espindola, G. M.	Linked Open Science-Communicating, Sharing and Evaluating Data, Methods and Results for Executable Papers	1. University of Muenster 2. Instituto Nacional de Pesquisas Espaciais	2011	International Conference on Computational Science (ICCS)	1. Alemanha 2. Brasil

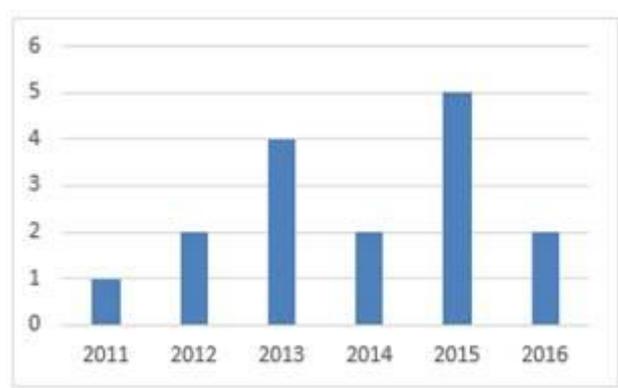
PEREIRA, Larissa Mariany Freiberger; MACHADO, Guilherme Bertoni; TODESCO, José Leomar. Avaliando a produção científica e a publicação de dados conectados na América Latina e Caribe. **Revista Democracia Digital e Governo Eletrônico**, Florianópolis, v. 1, n. 17, p. 92-106, out. 2018.

Tovar, E. Piedra, N. Chicaiza, J. López, J. Martinez- Bonastre, O	OER Development and Promotion. Outcomes of an International Research Project on the OpenCourseWare Model	1. Universidad Politécnica de Madrid 2. Universidad Técnica Particular de Loja 3. Miguel Hernandez University	2012	Journal of Universal Computer Science	1. Espanha 2. Equador 3. Espanha
--	--	--	------	--	--

Fonte: Autoria Própria.

Dos 16 artigos analisados, 12 tiveram a participação de pesquisadores vinculados a instituições brasileiras, 3 tiveram a participação de pesquisadores vinculados a instituições equatorianas e 1 foi desenvolvido por pesquisadores de instituição mexicanas. Ainda ressaltamos que 5 artigos foram apresentados em conferências e 11 artigos foram publicados em periódicos. Os 16 trabalhos foram produzidos por 57 autores, vinculados à 21 instituições distintas de 6 países.

Figura 3 – Produção Científica por Ano



Fonte: Autoria Própria.

Os trabalhos analisados foram publicados nos últimos 5 anos e, conforme apresentado na Figura 3, 2015 foi o ano com maior produção científica sobre o tema, contabilizando 5 artigos produzidos.

Destacamos aqui que a produção científica acerca da tecnologia de Dados Conectados nas regiões da América Latina e Caribe é oriunda de poucos países. Dos 32 países que compõem essas regiões, apenas 3 tiveram participação nos artigos analisados através de suas instituições. Os outros 3 países que tiveram participação na produção científica analisada são europeus (2) e norte-americanos (1).

O tema “Dados Conectados” é consideravelmente novo, tendo suas primeiras publicações em 2006 (Berners-Lee, 2006). Na região da América Latina e Caribe observamos que a primeira produção científica sobre o tema foi publicada em 2011, 6 anos mais tarde: *Linked Open Science-Communicating, Sharing and Evaluating Data, Methods and Results for Executable Papers* (Kauppinen and Espindola, 2011).

PEREIRA, Larissa Mariany Freiberg; MACHADO, Guilherme Bertoni; TODESCO, José Leomar. Avaliando a produção científica e a publicação de dados conectados na América Latina e Caribe. **Revista Democracia Digital e Governo Eletrônico**, Florianópolis, v. 1, n. 17, p. 92-106, out. 2018.

em âmbito governamental para RDF e este fato é refletido na baixa quantidade de datasets produzidos por governos das regiões da América Latina e Caribe (assunto tratado na seção seguinte).

Ressaltamos também a falta de iniciativas que trabalhem com mashups, desenvolvendo aplicações que impactem no dia-a-dia dos usuários a partir da conexão de datasets de diferentes fontes.

No campo da pesquisa, existe um gap considerável a ser explorado no que tange ao tema “Dados Abertos Conectados” nas regiões da América Latina e Caribe. Embora o tema tenha surgido há algum tempo, pouco se foi pesquisado à nível teórico e, menos ainda, à nível experimental. Quando pensamos no arcabouço de dados produzidos em nível governamental e na relevância que esses dados possuem para os cidadãos, empresas e sociedade em geral, pensamos também em quão rica seria a aproximação dos pesquisadores do tema “Dados Abertos Conectados” com os governos das regiões da América Latina e Caribe. Este fato possivelmente impactaria na quantidade de datasets disponibilizados aos cidadãos em formato RDF, bem como na quantidade de aplicações desenvolvidas para tornar mais simples o dia-a-dia do cidadão.

Ressaltamos também a necessidade da participação da iniciativa privada neste movimento, uma vez que esta também produz uma considerável quantidade de dados que interessam à sociedade em geral e, neste sentido, é possível imaginar a riqueza de informações descobertas a partir da conexão de dados governamentais e dados da iniciativa privada ou de outros tipos de instituições, entre outros.

4 Publicação de dados abertos conectados governamentais

A verificação da publicação de dados abertos governamentais foi realizada em portais de governo de todos os países que compõem a América Latina, bem como a região caribenha (Aqui ressaltamos a existência de muitos territórios que compõem a América Latina e Caribe, pertencentes a outros países. Esses territórios não foram incluídos na pesquisa, uma vez que fugiria do escopo da pesquisa).

Nestes países foi possível verificar que a produção de dados abertos conectados em âmbito governamental é consideravelmente baixa, conforme apresentado nas Tabelas 5 e 6.

Tabela 5 – Publicação de Dados Abertos Governamentais na América Latina

Nome	Classificação	Número de <i>Datasets</i>	<i>Datasets</i> em RDF	% RDF
Argentina	País	7 (http://www.datos.gob.ar/)	0	0
Bolívia	País	Portal de Dados Abertos Governamentais não encontrado		
Brasil	País	1.120 (http://dados.gov.br/)	2	0,17
Chile	País	2.456 (http://datos.gob.cl/)	0	0
Colômbia	País	2.233 (http://www.datos.gov.co/)	0	0
Costa Rica	País	2236 (http://www.datos.gov.co/)	0	0

PEREIRA, Larissa Mariany Freiberg; MACHADO, Guilherme Bertoni; TODESCO, José Leomar. Avaliando a produção científica e a publicação de dados conectados na América Latina e Caribe. **Revista Democracia Digital e Governo Eletrônico**, Florianópolis, v. 1, n. 17, p. 92-106, out. 2018.

Cuba	País	Portal de Dados Abertos Governamentais não encontrado		
República Dominicana	País	309 (http://datos.gob.do/)	0	0
El Salvador	País	Portal de Dados Abertos Governamentais não encontrado		
Equador	País	70 (http://sni.gob.ec/datosabiertos)	0	0
Guatemala	País	Portal de Dados Abertos Governamentais não encontrado		
Haiti	País	Portal de Dados Abertos Governamentais não encontrado		
Honduras	País	Portal de Dados Abertos Governamentais não encontrado		
México	País	2.934 (http://datos.gob.mx)	0	0
Nicarágua	País	Portal de Dados Abertos Governamentais não encontrado		
Panamá	País	Portal de Dados Abertos Governamentais não encontrado		
Paraguai	País	123 (https://www.datos.gov.py/)	0	0
Peru	País	Portal de Dados Abertos Governamentais não encontrado		
Uruguai	País	125 (http://datos.gub.uy/)	1	0,8
Venezuela	País	Portal de Dados Abertos Governamentais não encontrado		

Fonte: Autoria Própria.

Tabela 6 – Publicação de Dados Abertos Conectados Governamentais no Caribe

Nome	Classificação	Número de <i>Datasets</i>	<i>Datasets</i> em RDF	% RDF
Antigua e Barbuda	País	Portal de Dados Abertos Governamentais não encontrado		
Aruba	País	Portal de Dados Abertos Governamentais não encontrado		
Barbados	País	Portal de Dados Abertos Governamentais não encontrado		
Curaçao	País	Portal de Dados Abertos Governamentais não encontrado		
Dominica	País	Portal de Dados Abertos Governamentais não encontrado		
Grenada	País	Portal de Dados Abertos Governamentais não encontrado		
Jamaica	País	Portal de Dados Abertos Governamentais não encontrado		
São Cristóvão e Névis	País	Portal de Dados Abertos Governamentais não encontrado		
Santa Lúcia	País	Portal de Dados Abertos Governamentais não encontrado		
São Vicente e Granadinas	País	Portal de Dados Abertos Governamentais não encontrado		
Bahamas	País	Portal de Dados Abertos Governamentais não encontrado		
Trindade e Tobago	País	30 (http://data.tt/)	0	0

Fonte: Autoria Própria.

No contexto da América Latina é possível observar que, em relação à publicação de dados abertos governamentais, metade dos países não possui sequer o portal para publicação destes.

Quando se trata da publicação de dados abertos conectados, dos 20 portais governamentais analisados, apenas 2 produzem, em baixíssima escala, datasets em RDF, podendo conectar esses dados a outros conjuntos de dados.

Aqui ressaltamos que o país que tem publicada a maior quantidade de datasets em RDF (Brasil) é também o país com maior número de instituições pesquisadoras acerca do tema “Dados Conectados” identificados nos documentos analisados e apresentados na seção anterior. É possível observar, portanto, que o que vem sendo produzido na academia, em termos de conhecimento científico, vagarosamente vem ganhando espaço no âmbito governamental.

Em relação aos países caribenhos a situação é ainda mais grave. Dos 12 países verificados, apenas um possui portal governamental para publicação de dados abertos do governo e, ainda assim, há poucos datasets disponibilizados (apenas 30). Aqui percebemos que não existe nenhuma iniciativa governamental na região que fomente a produção de dados abertos por parte do governo.

Consideramos importante destacar aqui que, antes de tratarmos sobre Dados Conectados nesses países, é preciso que seja gerada uma cultura de publicação de dados entre os governos de países caribenhos. Sem esta cultura consolidada não é possível pensar que estes países venham a disponibilizarem seus dados em formatos que permitam a conexão destes com outros datasets.

5 Considerações finais

A pesquisa conduzida teve como objetivo levantar a produção científica, bem como a publicação em âmbito governamental de Dados Conectados na América Latina e Caribe para compreender em que nível os países que compõem essas duas regiões estão no que tange à tecnologia supracitada.

O processo de pesquisa em bases de dados não se mostrou suficiente, pois poucos artigos foram encontrados. Dessa maneira, foi necessário verificar a quantidade de datasets publicados em portais governamentais de países que compõem a América Latina e Caribe para que a pesquisa pudesse ser ampliada e gerar resultados mais consistentes. Aqui foi possível verificar que, de fato, pouco se foi pesquisado sobre o tema até o presente momento nas regiões da América Latina e Caribe, existindo um amplo nicho que pesquisa a ser explorado

A principal dificuldade foi que até mesmo os portais governamentais não têm informações de fácil acesso, quando as têm. Além disso, os dados publicados, quase em sua totalidade, ainda estão nos níveis baixos da classificação 5 estrelas proposta por Berners-Lee (2006), ou seja, não são classificadas como Dados Abertos Conectados e, tampouco estão disponibilizados ao público em formatos que possibilite que estes venham a ser conectados posteriormente.

A região estudada está consideravelmente atrás dos Estados Unidos e Europa no que tange à produção científica e publicação em âmbito governamental de Dados Abertos Conectados, o que pode ser observado através dos dados bibliométricos (Dal Pizzol, de Moura Speroni, Zancanaro, Gauthier & Todesco, 2015). Portanto, aqui identificamos um

nicho de pesquisa bastante relevante, uma vez que os estudos realizados até agora são poucos e, de fato, há muito trabalho a ser feito nessas regiões, principalmente no âmbito governamental, como foi verificado e citado anteriormente.

O artigo apresentou uma síntese que pode guiar os pesquisadores desta região nas futuras decisões acerca de pesquisas dados abertos conectados a partir de vários vieses, como publicação, visualização, consumo, entre outros. Da mesma maneira, mostramos a deficiência, em âmbito governamental, nas regiões da América Latina e Caribe de iniciativas que fomentem o uso da tecnologia de Dados Abertos Conectados.

É necessário, na verdade, antes disso, consolidar a cultura da publicação de dados abertos governamentais na região, visto que muitos países não chegaram nem ao menos na etapa de disponibilizar um espaço na Web onde seus dados possam ser publicados.

6 Referências

- AUER, S., DIETZOLD, S., LEHMANN, J., HELLMANN, S., & AUMUELLER, D. (2009, April). Triplify: light-weight linked data publication from relational databases. In **Proceedings of the 18th international conference on World wide web** (pp. 621-630). ACM.
- BERNERS-LEE, T. (2006). **Design issues: Linked data.**
- BIZER, C., CYGANIAK, R., & HEATH, T. (2007). **How to publish linked data on the web.**
- BIZER, C., HEATH, T., IDEHEN, K., & BERNERS-LEE, T. (2008, April). Linked data on the web (LDOW2008). In **Proceedings of the 17th international conference on World Wide Web** (pp. 1265-1266). ACM.
- BIZER, C. (2009). The emerging web of linked data. **IEEE intelligent systems**, 24(5), 87-92.
- BIZER, C., HEATH, T., & BERNERS-LEE, T. (2009). Linked data-the story so far. **Semantic Services, Interoperability and Web Applications: Emerging Concepts**, 205-227.
- DAL PIZZOL, L., DE MOURA SPERONI, R., ZANCANARO, A., GAUTHIER, F. O., & TODESCO, J. L. (2015). Análise bibliométrica da produção científica sobre Linked Data. **Informação & Informação**, 20(3), 77-112.
- HASLHOFER, B., & ISAAC, A. (2011, September). data. europeana. eu: The europeana linked open data pilot. In **International Conference on Dublin Core and Metadata Applications** (pp. 94-104).
- HEATH, T., & BIZER, C. (2011). Linked data: Evolving the web into a global data space. **Synthesis lectures on the semantic web: theory and technology**, 1(1), 1-136.
- PEREIRA, L. M. F., MACHADO, G. B., & TODESCO, J. L. (2015, November). Linked Open Data como fundamento da Web de Dados: uma revisão sistemática. In **Anais do Congresso Internacional de Conhecimento e Inovação** (pp. 1769-1788), Joinville, SC, Brasil, 5.