

QUALIDADE DA INFORMAÇÃO EM LINKED DATA: DADOS BIBLIOMÉTRICOS

Murilo Silveira Gomes¹, Lidiane Visintin², Larissa Freiburger Pereira³, Fernando A. O. Gauthier⁴, Marcelo Macedo⁵

Abstract. *There has been a significant increase in the volume of data and information on the web in recent years. All this accumulation of data and information can cause several problems, with a duplicate or an inconsistency. In this way, a quality of information is a way to soften problems in the context of the web, as well as in the context of semantic web. This study focuses This study has the objective of carrying out a quantitative survey of the literature through a bibliometry in order to identify the researches in the area of quality of information and Linked Data. With this research it was possible to identify that there is a large number of journals events that deals with quality of information and Linked Data, as well as the countries and universities that investigate the subject.*

Keywords: *Quality of Information, Linked data, Bibliometrics*

Resumo. *Houve um aumento significativo do volume de dados e informações contidas na web, nos últimos anos. Toda este acúmulo de dados e informações podem ocasionar vários problemas, como, dados duplicados ou inconsistentes. Desta forma a qualidade da informação é uma maneira para amenizar possíveis problemas, tanto no contexto da web como na web semântica. Este estudo por sua vez tem como foco realizar um levantamento quantitativo da literatura através de uma bibliometria a fim de identificar como estão as pesquisas no âmbito da qualidade da informação e Linked Data. Com a realização deste trabalho foi possível identificar que existe um grande número de journals e eventos sobre qualidade da informação que abordam Linked Data, assim como os países e universidades que pesquisam sobre o assunto.*

Palavras-Chave: *Qualidade da Informação, Linked data, Bibliometria*

¹ Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento– Universidade Federal de Santa Catarina (UFSC) Florianópolis – SC– Brazil. Email: lilo.flp@gmail.com

² Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento– Universidade Federal de Santa Catarina (UFSC) Florianópolis – SC– Brazil. Email: Lidiane.visintin@gmail.com

³ Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento– Universidade Federal de Santa Catarina (UFSC) Florianópolis – SC– Brazil. Email: larii.f@gmail.com

⁴ Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento– Universidade Federal de Santa Catarina (UFSC) Florianópolis – SC– Brazil. Email: Fernando.gauthier@gmail.com

⁵ Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento– Universidade Federal de Santa Catarina (UFSC) Florianópolis – SC– Brazil. Email: marcelo.macedo@ufsc.br

1 INTRODUÇÃO

A Web teve seu início potencializando apenas a exibição das informações de maneira estática e, em um segundo momento, a web já exibia informações de maneira dinâmica. Esse dinamismo possibilitou aos usuários uma melhor experiência, interatividade e tornou-a uma web mais atrativa. Toda esta evolução da web ocasionou também no crescimento do armazenamento de dados não estruturados, devido aos avanços das mídias digitais. Com estes avanços surgiram problemas de gerenciamento de informação e problemas referentes a padronização dos conteúdos que são disponibilizados (Group, 2014; Mendonça & Neto, 2015).

Percebe-se assim que a web continua em constante evolução, sendo que esta evolução vem sendo ampliada através da web semântica. A web semântica por sua vez, tem por objetivo dar significado aos dados transformando-os em informações e possibilitando a interpretação por homens e máquinas (Aghaei *et al.* 2012).

A web atual é composta por *hypertext*⁶, tal composição não é suficiente para dar suporte à web semântica. Deste modo o *Linked Data* é apresentado para auxiliar a web semântica a dar significado aos dados através das ligações entre dados e *datasets*. O *Linked Data* utiliza de links *Hyperdata*⁷ com informações expressas em *Resource Description Framework* - (RDF), ou seja, não são apenas os documentos conectados, mas sim os dados que estão conectados (Sheth, 2011; Berners-Lee, 2006).

Através do uso de *Linked Data* pode-se ter alguns benefícios, sendo estes: a uniformidade, a capacidade de ser referenciado, a coerência, a interoperabilidade e a pontualidade (Auer *et al.*, 2011). Ao fazer uso de *Linked Data* tem-se uma preocupação quanto a qualidade dos dados, devido ao grande volume de dados que podem ser manipulados (Acosta *et al.*, 2013).

Existem vários tipos de problemas relacionado a qualidade de dados e informação, como exemplo, informações com múltiplas fontes, utilização de julgamentos subjetivos, sistemáticos e erros na produção de informação (Calazans, 2008; Assis & Moura, 2011; Basharat *et al.* 2014).

Observando o contexto apresentado, o objetivo deste artigo é o de fornecer aos pesquisadores uma compreensão abrangente dos trabalhos presentes na literatura sobre Qualidade da informação e *Link Data*, estimulando assim a continuação da experimentação e desenvolvimento de novas abordagens centradas no sentido da qualidade de dados e informações, especificamente para *Linked Data*.

⁶ Hipertexto é um texto que contém ligações a outros textos.

⁷ Hyperdata indica objetos de dados ligados a outros objetos de dados em outros lugares.

A seguir, serão apresentadas as seções de trabalho: na seção 2 é abordada a revisão da literatura contendo os temas abordados pela pesquisa; na seção 3 é apresentada a metodologia utilizada para a realização deste estudo; na seção 4 são apresentados os resultados obtidos com a bibliometria; e por fim, são apresentados os encaminhamentos para futuros trabalhos e conclusões.

2 REVISÃO DE LITERATURA

O *World Wide Web*, ou apenas Web, evolui constantemente e é considerada um meio de encurtar distâncias, compartilhar informação e expressar opiniões. No ano de 1989 Tim Berners-Lee, propôs um ambiente onde se pudesse apresentar textos, imagens, vídeos e animações de modo que fossem integrados por meio da internet, conhecido como a Web (Berners-Lee, 1989).

A Web Semântica é uma extensão que busca materializar uma web onde não apenas os seres humanos sejam capazes de interpretar e distinguir conteúdos conforme determinado contexto, mas que as máquinas sejam capazes de compreender os significados (Alves *et al.*, 2005). A Web Semântica tem por objetivo tornar possível a interpretação dos conteúdos de forma automática. Para que isso ocorra, os documentos e dados precisam ser representados de maneira padronizada (Berners-Lee, 2001). Um dos conceitos sobre as práticas de conteúdo padronizados é o *Linked Data*.

1.1 LINKED DATA

O termo *Linked Data* foi cunhado por Tim Berners-Lee em 2006, sendo que no contexto de Web Semântica o *Linked Data* não visa apenas inserir dados na Web, também possibilita realizar as ligações de modo que uma máquina ou seres humanos possam explorar a Web. Deste modo *Linked Data* é definido como "um conjunto de boas práticas para conectar e publicar dados estruturados na Web de diferentes fontes" (Berners-Lee, 2006).

A publicação de dados na Web já vem sendo realizada há algum tempo, mas tais abordagens tradicionalmente são realizadas com a utilização de dados brutos no formato CSV, XML ou HTML, sem preocupação com a integração ou com sua semântica (Sheth, 2011).

Deste modo o uso de *Linked Data* proporciona um maior significado aos dados, a seguir são apresentados alguns benefícios do uso de *Linked Data* (Auer *et al.*, 2011), sendo eles:

- **Uniformidade:** Com a utilização do modelo de dados RDF toda a informação publicada é representada em fatos expressos por termos consistindo por sujeito, predicado e objeto. Deste modo todos os dados são baseados em um modelo uniforme.
- **Capacidade de ser referenciado:** Utiliza-se de URIs para identificar entidades e permite a localização, recuperação de recursos descritos e representados na web por estas entidades assim como uma URL.
- **Coerência:** As triplas RDF estabelecem uma ligação entre as entidades identificadas pelo sujeito e o objeto sendo interligadas por um predicado que representa a ligação entre eles.
- **Integrabilidade:** Todas as fontes de dados *Linked Data* são baseadas em um único mecanismo para representar a informação, sendo de fácil integração semântica por diferentes conjuntos de dados.
- **Pontualidade:** Todas as publicações e atualização dos dados tendem a reutilização, facilitando a disponibilização.

Tim Berners-Lee criou quatro recomendações básicas para que haja a possibilidade da publicação de dados *Linked Data* (Berners-Lee, 2006), sendo:

- I - Use URIs como nomes para as coisas;
- II - Use HTTP URIs para que as pessoas possam procurar esses nomes;
- III - Quando alguém procura um URI, fornecer informações úteis, usando de padrões (RDF ou SPARQL);
- IV - Incluir links para outros URIs. Para que se possa descobrir mais informações.

Com a adaptação das recomendações básicas para a publicação de dados proposta por Berners-Lee surgiram outros projetos como: o *Linked Open Data*, que visa à publicação de dados abertos; o *Linked Government Data* que visa a publicação dos dados abertos governamentais; e o *Enterprise Linked Data* visando a publicação de dados de empresas (Dal Pizzol *et al.*, 2015). No entanto, todos estes projetos preocupam-se com a qualidade dos dados, assim como a qualidade da informação de modo a obter aplicações confiáveis, ou seja, que façam uso de dados de qualidade.

1.2 QUALIDADE DA INFORMAÇÃO

A qualidade da informação – (QI) é um termo que começou a ser discutido em 1989 no seminário do Conselho Nórdico de Bibliotecas de Informação e Investigação Científicas, onde foi apresentado algumas propostas sobre a QI. Algumas dessas propostas abordaram a gestão da informação no âmbito do gerenciamento de negócios identificando algumas abordagens e

dimensões. Desde então a qualidade da informação vem sendo explorada em diversos domínios (Calazans, 2012).

A qualidade da informação é apresentada como um conceito multidimensional, de modo que um produto físico também possui dimensões associadas a qualidade, assim como uma informação ou um dado (Lin, Gao & Koronios, 2006; Wand & Wand, 1996).

Para a avaliação da qualidade da informação existem diferentes perspectivas com diferentes dimensões. Segundo dos Santos (2016) foram realizados alguns estudos e feito o agrupamento de quinze dimensões em quatro categorias, sendo estas:

- **Categoria intrínseca** (acurácia, credibilidade, objetividade, precisão e confiabilidade);
- **Categoria contextual** (relevância, oportunidade, completude e conveniência);
- **Categoria de representação** (compreensão, interpretabilidade, representação concisa e consistente);
- **Categoria de acessibilidade** (acessibilidade, segurança, disponibilidade do sistema, facilidade de operação e privilégios).

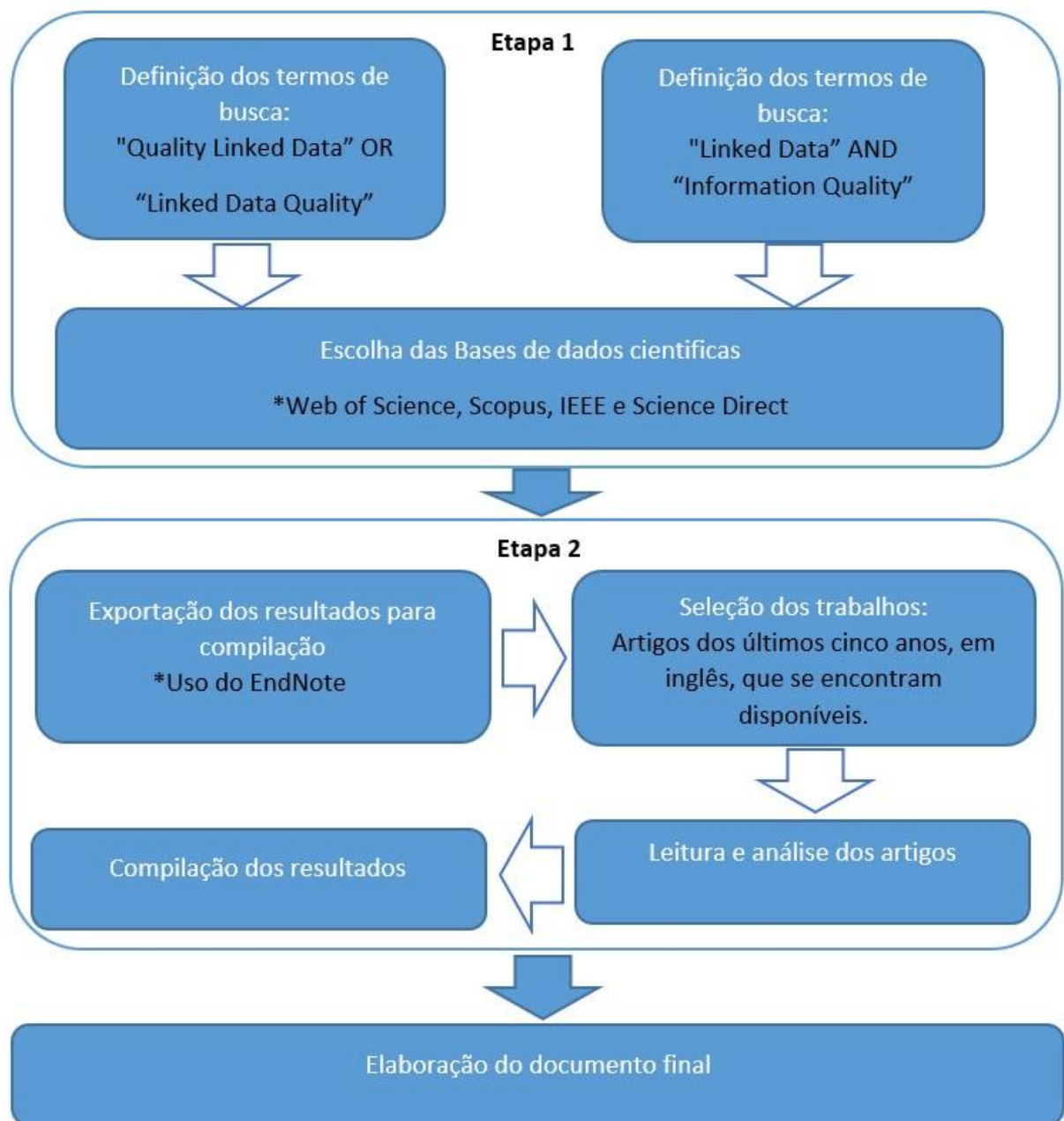
Apresentado o contexto de Qualidade da Informação e de *Linked Data*, percebe-se diversas áreas de pesquisa que permeiam os temas, no entanto o foco deste artigo é fornecer uma visão geral de como está a área de Qualidade da Informação para aqueles que desejam iniciar seus estudos neste contexto de *Linked Data*.

3 METODOLOGIA

Esta pesquisa utiliza a bibliometria, cujo objetivo é analisar de modo quantitativo as publicações e as atividades científicas de determinada área ou tema. Esta abordagem torna possível verificar, por exemplo, os anos que tiveram mais publicação, as revistas que publicam na área ou no tema pesquisado, quais os autores e as universidades que se destacam assim como a identificação de palavras chaves que mais são utilizadas (da Silva, Hayashi e Hayashi, 2011; Santos, 2003).

A Figura 1 apresenta a sumarização das etapas realizadas para a obtenção dos resultados que são abordados neste artigo, sendo que os mesmos são apresentados na sequência.

Figura 1. Fluxograma com a metodologia utilizada para a realização da pesquisa bibliográfica e compilação de resultados.



Fonte: Os autores (2017)

Na etapa 1 foram definidos dois conjuntos de busca diferentes: “*Quality Linked Data*”, “*Linked Data Quality*” e “*Linked Data*”, “*Information Quality*”. Foram utilizadas as palavras chaves entre aspas duplas a fim de buscar exatamente os termos contidos nos documentos. Utilizou-se também o operador booleano AND para buscar a expressão que contenha as duas palavras chaves utilizadas e o operador booleano OR para buscar a expressão que contenha uma das duas palavras chaves utilizadas, conforme especificado na Figura 1. Para a realização deste estudo foram utilizadas quatro bases de dados: SCOPUS, Web of Science, Science Direct e IEEE.

Foram selecionadas estas bases por serem bases multidisciplinares, interdisciplinares e tecnológicas possibilitando assim uma melhor aderência aos temas pesquisados.

Na etapa 2 os resultados obtidos através das bases foram exportados e importados no software *Endnote*, onde foi possível organizar todos os documentos e encontrar os documentos disponíveis. Após a organização foi possível realizar a análise e, posteriormente, a escrita desse documento. A seguir são apresentados alguns dos resultados obtidos.

4 RESULTADOS

Para a realização desta pesquisa, foi utilizado duas sintaxes, sendo elas: a primeira ("quality linked data" OR "linked data quality") e a segunda sintaxe ("linked data" AND "information quality"). A busca resultou em um total de 146 documentos, sendo na base Scopus recuperados 26, na base Word of Science 13, na Science Direct 57 e na IEEE 50 documentos. Com os documentos coletados, foram aplicados os seguintes filtros: dos últimos cinco anos, exclusão de documento duplicados e a utilização apenas dos documentos disponíveis para análise. Ao final da busca foram recuperados 61 documentos para análise.

Com este levantamento foi possível visualizar a quantidade de publicações por ano, os principais *journals* e eventos, as palavras chaves mais utilizadas e as universidades que mais se destacam nesta área de pesquisa.

4.1 JOURNALS E OU PERIÓDICOS

Selecionados os artigos, foi realizado um levantamento dos *journals* e/ou eventos que os documentos foram publicados. Abaixo é apresentado a Tabela 2, com informações sobre os periódicos como: os *journals* e o ISSN / ISBN.

Tabela 1 – Journals que abordaram a temática

| Journals / Eventos | ISSN / ISBN |
|---|-------------------|
| 26th International Workshop on Database and Expert Systems Applications | 978-1-4673-7582-5 |
| IEEE Transactions on Fuzzy Systems | 1063-6706 |
| Virtual Systems & Multimedia (VSMM) 2014 | 978-1-4799-7227-2 |
| Computer Science and Software Engineering (CSSE) 2015 | 978-1-4673-9181-8 |
| Biomedical and Health Informatics (BHI) 2014 | 2168-2194 |
| Data Engineering Workshops (ICDEW) 2014 | 978-1-4799-3481-2 |

| | |
|--|-------------------|
| IEEE 8th International Symposium on Service Oriented System Engineering | 978-1-4799-3616-8 |
| The Semantic Web: Trends and Challenges, | 978-3-319-07443-6 |
| Web Intelligence (WI), 2016 | 978-1-5090-4470-2 |
| Government Information Quarterly | 0740-624X |
| International Conference on Accounting Studies 2014, ICAS 2014 | 1877-0428 |
| International Workshop on Geospatial Big Data - Trends, Applications, and Challenges | 1877-0509 |
| Computer Communications and Networks (ICCCN) 2013 | 978-1-4673-5775-3 |
| The Semantic Web - ISWC 2015 | 978-3-319-25010-6 |
| Telecommunications Forum (TELFOR), 2016 | 978-1-5090-4086-5 |
| Big Data Computing Service and Applications (BigDataService) 2016 | 978-1-5090-2251-9 |
| eChallenges e-2014 Conference Proceedings | 978-1-905824-46-5 |
| IEEE Access Vol.4 | 2169-3536 |
| Methods of Information in Medicine | 0026-1270 |
| Engineering, Technology and Innovation/ International Technology Management Conference (ICE/ITMC) 2015 | 978-1-4673-7156-8 |
| 2014 Tenth International Conference on Signal-Image Technology & Internet-Based Systems | 978-1-4799-7978-3 |
| Semantic Computing (ICSC) 2014 | 978-1-4799-4003-5 |
| Computers in Human Behavior | 0747-5632 |
| OCEANS'15 MTS/IEEE Washington | 978-0-9339-5743-5 |
| System Sciences (HICSS) | 978-0-7695-5670-3 |
| Journal of Banking & Finance | 0378-4266 |
| 2014 Joint Conference of the International Workshop on Software Measurement and the International Conference on Software Process and Product Measurement | 978-1-4799-4174-2 |
| Web Semantics: Science, Services and Agents on the World Wide Web | 1570-8268 |
| Journal of Biomedical Informatics | 1532-0464 |
| The semantic web: ESWC 213 Satellite Events | 0302-9743 |
| Information and Software Technology | 0950-5849 |
| Global Information Infrastructure and Networking Symposium (GIIS), 2016 | 2150-329X |
| Geoscience and Remote Sensing Symposium (IGARSS), 2014 IEEE International | 978-1-4799-5775-0 |
| IEEE Transactions on Geoscience and Remote Sensing | 0196-2892 |
| International journal on Semantic Web and information systems · | 552-6283 |
| Procedia - Social and Behavioral Sciences | 1877-0428 |
| The Gray Jjournal | 1574-1796 |
| Computer | 0018-9162 |
| ISKO-Maghreb: Concepts and Tools for knowledge Management (ISKO-Maghreb) 2014 | 978-1-4799-7508-2 |
| Global Communications Conference (GLOBECOM), 2016 IEEE | 978-1-5090-1328-9 |

| | |
|---|-------------------|
| IEEE Transactions on Geoscience and Remote Sensing | 0196-2892 |
| eChallenges e-2015 Conference, 2015 | 978-1-9058-2453-3 |
| Internet of Things (WF-IoT), 2016 IEEE | 978-1-5090-4130-5 |
| 1st Workshop on Linked Data Quality co-located with 10th International Conference on Semantic Systems - CEUR Workshop Proceedings | 1613-0073 |
| IEEE Transactions on Geoscience and Remote Sensing | 0196-2892 |
| Chilean Computer Science Society (SCCC), 2014 33rd International Conference of the | 1522-4902 |
| CENTERIS 2014 - Conference on ENTERprise Information Systems ProjMAN 2014 - International Conference on Project MANagement / HCIST 2014 - International Conference on Health and Social Care Information Systems and Technologies | 2212-0173 |
| Web Semantics: Science, Services and Agents on the World Wide Web | 1570-8268 |
| IEEE Transactions on Knowledge and Data Engineering | 1041-4347 |
| Future Generation Computer Systems | 0167-739X |
| Journal of Financial Economics | 0304-405X |
| Management of Engineering & Technology (PICMET), 2014 | 978-1-890843-29-8 |
| Information and Communication Technology (ICoICT), 2015 | 978-1-4799-7752-9 |
| Applied Ergonomics | 0003-6870 |

Fonte: autores (2017)

A tabela mostra todos os *journals* e eventos em que os documentos recuperados foram publicados. A seguir será apresentada a descrição dos *journals* e eventos que possuem o maior número de documentos publicados e ordenados por maior quantidade de publicação:

O Jornal *Government Information Quarterly*⁸ foi o *journal* que mais teve documentos publicados, tendo um total de sete documentos. Este *journal* possui o foco na interseção política, tecnologia da informação e governo público. Este *journal* tem o objetivo de publicar pesquisas acadêmicas, revisões, ensaios e editoriais que informam aos pesquisadores e aos profissionais sobre a relação dos temas destacados anteriormente.

O *IEEE Transactions on Geoscience and Remote Sensing*⁹ é um *journal* que realiza publicações mensais e possui o foco em teorias, conceitos e técnicas de ciência e engenharia aplicadas na constatação da terra, dos oceanos, da atmosfera e do espaço, bem como o processamento, interpretação e divulgação destas informações. Este foi o segundo *journal* que teve mais documentos publicados.

⁸ *Government Information Quarterly* <<https://www.journals.elsevier.com/government-informationquarterly>>

⁹ *IEEE Transactions on Geoscience and Remote Sensing*
<<http://ieeexplore.ieee.org/xpl/aboutJournal.jsp?punumber=36>>

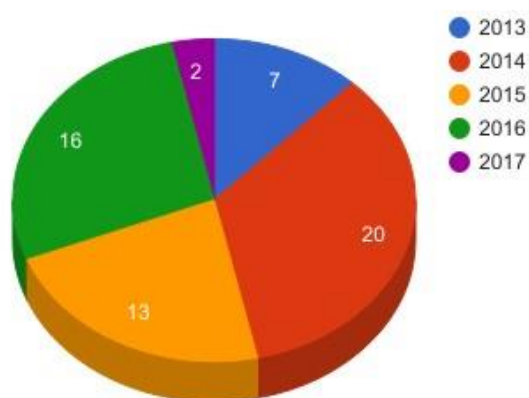
O *eChallenges*¹⁰ é um evento que ocorre todos os anos e está na sua vigésima quinta edição, possuindo o foco na promoção das TICs, empreendedorismo e inovação, facilitando a partilha de conhecimentos relacionados com a sociedade da informação, TICs aplicadas no governo e indústria. Nesta pesquisa foram encontradas publicações referentes aos anos de 2014 e 2015, possuindo um total de 2 documentos publicados.

O *The Semantic Web*¹¹ – ISWC é um evento que possui o foco de discussão sobre a comunidade da web semântica e *Linked Data*. Este evento vem acontecendo todos os anos. Nesta pesquisa foram encontrados documentos referente aos anos de 2013, 2014 e 2015. O *Web Semantics: Science, Services and Agents on the World Wide Web*¹² é um *journal* indexado na *Elsevier*. Este *journal* tem como foco pesquisas e aplicações que contribuem no desenvolvimento de conhecimento e serviços inteligentes na web, nas áreas de tecnologias do conhecimento, ontologia, agentes, bancos de dados e redes semânticas. Para este *journal* foram encontrados dois documentos.

3.2 QUANTIDADE X ANO DE PUBLICAÇÃO

A figura 6 apresenta o número de artigos por ano de publicação, em relação as duas sintaxes de busca utilizada.

Figura 2 – Publicação por ano



Fonte: autores (2017)

Observa-se que em 2013 há sete artigos publicados, 2014 há vinte artigos publicados (ano com maior incidência de publicação de artigos sobre o tema deste trabalho), 2015 treze artigos

¹⁰ eChallenges < <http://www.echallenges.org> >

¹¹ International Semantic Web Conference < iswc2017.semanticweb.org >

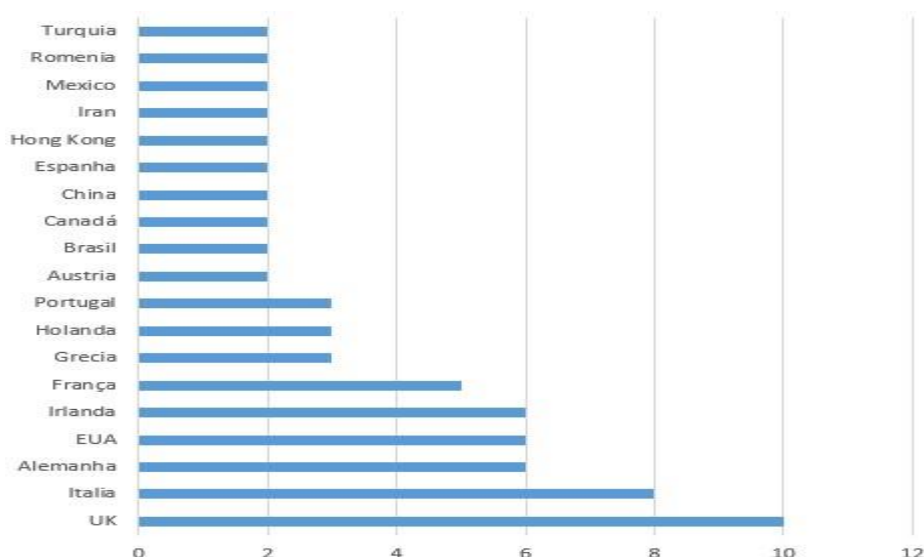
¹² *Web Semantics: Science, Services and Agents on the World Wide Web*
<<https://www.journals.elsevier.com/journal-of-web-semantics/>>

foram publicados, em 2016 foram dezesseis artigos e em 2017 foram até o momento publicados dois artigos. É possível observar que este tema é emergente.

3.3 PAISES QUE MAIS PUBLICAM

Os artigos encontrados nas bases de dados para a realização da pesquisa, apresentaram ao todo uma soma de 179 autores que estão afiliados à 106 institutos, universidades ou centros de pesquisa instalados em 35 países. O Gráfico 2 apresenta os países que mais publicaram nesses últimos cinco anos.

Figura 3 – Países que mais publicaram nos últimos cinco anos



Fonte: autores (2017)

Os países com maior concentração de autores dos 11 documentos escolhidos foram: Reino Unido com dez autores, Itália com oito autores, Alemanha, EUA, Irlanda com seis autores, França com cinco autores, Grécia, Irlanda e Portugal com três autores. Os demais países incluído o Brasil com dois autores. Como pode ser percebido a maior concentração de autores e, consequentemente, pesquisas na área de Qualidade da Informação e *Linked Data* está no Reino Unido.

3.4 PALAVRAS-CHAVES MAIS UTILIZADAS

É apresentado a seguir a nuvem das palavras-chaves criada a partir dos 62 documentos recuperados, tendo um total de 594 palavras. Para a construção desta nuvem, as palavras-chaves foram analisadas de maneira isolada, não sendo possível identificar termos compostos.

[illegible]

A nuvem de palavras-chaves tem como destaque as palavras “*Quality, information, semantics, ontologies, databases, web, framework*”. Sendo visto que são palavras que tendem aparecer naturalmente em virtude aos termos de busca utilizados.

As palavras “*open, government, health, analysis*” também se destacam, pois, são temas que possuem uma necessidade de ter qualidade de informação e dados, em aplicações nesse contexto.

5 CONCLUSÃO

Este trabalho apresentou uma perspectiva geral através de uma bibliometria sobre o tema de qualidade da informação focada em *Linked Data*. Foi possível identificar que existe um grande número de *journals* e eventos que abordam tanto temas isolados da qualidade da informação como temas da qualidade de dados em *Linked Data*. Também foi possível visualizar que nos anos de 2013, 2014 e 2016 foram os anos que mais tiveram publicação e até junho de 2017 há poucos documentos que abordam esses temas nas bases de dados analisadas. Em relação aos países que mais publicam sobre os temas está o Reino Unido – UK possuindo dez publicações.

Por fim, foi possível identificar na nuvem de palavras-chave que as palavras que mais se destacam são “*Quality, information, semantics, ontologies, databases, web, framework*” por estarem ligadas diretamente aos temas pesquisados.

Como trabalhos futuros é sugerida a realização de uma pesquisa sistemática a fim de identificar como é abordada a qualidade de dados junto ao *Linked Data*. Outro ponto que pode ser explorado é como as ontologias influenciam na qualidade da informação.

REFERÊNCIAS

- Aghaei, S., Nematbakhsh, M. A., & Farsani, H. K. (2012). Evolution of the world wide web: From WEB 1.0 TO WEB 4.0. *International Journal of Web & Semantic Technology*, 3(1), 1.
- Alves, R. C. V. (2005). Web Semântica: uma análise focada no uso de metadados.
- Auer, S., Lehmann, J., & Ngomo, A. C. N. (2011). Introduction to linked data and its lifecycle on the web. In *Reasoning Web. Semantic Technologies for the Web of Data* (pp. 1-75). Springer Berlin Heidelberg.
- Berners-Lee, T. (1989). Information management: A proposal.
- Berners-Lee, T. (2011). Linked data-design issues (2006). URL <http://www.w3.org/DesignIssues/LinkedData.html>, 10, 11.
- Berners-Lee, T., Hendler, J., & Lassila, O. (2001). The Semantic Web. A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities. *Scientific American*, 284(5), 1-5.
- Calazans, A. T. S. (2012). Qualidade da informação: conceitos e aplicações. *Transinformação* ISSN 2318-0889, 20(1).
- Calazans, A. T. S. (2012). Qualidade da informação: conceitos e aplicações. *Transinformação* ISSN 2318-0889, 20(1).
- da Silva, M. R., Hayashi, C. R. M., & Hayashi, M. C. P. I. (2011). Análise bibliométrica e cientométrica: desafios para especialistas que atuam no campo. *InCID: Revista de Ciência da Informação e Documentação*, 2(1), 110-129.
- 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.
- dos Santos, F. (2016). *Qualidade da informação estratégica organizacional utilizando a casa da qualidade* (Dissertação). Universidade Federal de Santa Catarina, <http://btd.egc.ufsc.br/wp-content/uploads/2016/05/Fernanda-dos-Santos.pdf>.
- Gomes, F. C. D. R. (2013). Arquitetura de repositório semântico de dados para organização de pesquisa agropecuária.
- Group, D. R. (2014). *A world that counts mobilising the data revolution for sustainable development*. Retrieved from <http://www.undatarevolution.org/wp-content/uploads/2014/12/A-World-That-Counts2.pdf>
- Lee, S. H., & Haider, A. (2013). *Identifying relationships of information quality dimensions* (pp. 1217-1228). IEEE.
- Lin, S., Gao, J., & Koronios, A. (2006, June). The need for a data quality framework in asset management. In *Australian Workshop on Information Quality* (Vol. 1).

- Santos, R. N. M. D. (2003). Produção científica: por que medir? o que medir?. *Revista digital de Biblioteconomia e Ciência da Informação*, 1(1).
- Sheth, A. (Ed.). (2011). *Semantic Services, Interoperability and Web Applications: Emerging Concepts: Emerging Concepts*. IGI Global.
- Wand, Y., & Wang, R. Y. (1996). Anchoring data quality dimensions in ontological foundations. *Communications of the ACM*, 39(11), 86-95.
- Acosta, M., Zaveri, A., Simperl, E., Kontokostas, D., Auer, S., & Lehmann, J. (2013, October). Crowdsourcing linked data quality assessment. In *International Semantic Web Conference* (pp. 260-276). Springer Berlin Heidelberg.
- de Assis, J., & Moura, M. A. (2011). A Qualidade da Informação na Web: uma abordagem semiótica. *Informação & Informação*, 16(3), 96-117.
- Basharat, A., Arpinar, I. B., Dastgheib, S., Kursuncu, U., Kochut, K., & Dogdu, E. (2014, June). CrowdLink: Crowdsourcing for large-scale linked data management. In *Semantic Computing (ICSC), 2014 IEEE International Conference on* (pp. 227-234). IEEE.
- Mendonça, A. P. B., & Neto, A. P. (2015). Critérios de avaliação da qualidade da informação em sites de saúde: uma proposta. *Revista Eletrônica de Comunicação, Informação & Inovação em Saúde*, 9(1).