

UMA REVISÃO DE LITERATURA ACERCA DE LINKED ENTERPRISE DATA

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

Abstract. *The technologies used in the semantic web have been gaining prominence in the researches in recent years, but little has been researched on the business context. Firms are beginning to be interested in semantic technologies, but there is still a fear of get in this area because of the small amount of research in this area. Linked Enterprise Data and Enterprise Linked Data are concepts that incorporate Linked Data into an enterprise environment. In this sense, the objective of this work is to present definitions found in the literature for Linked Enterprise Data and Enterprise Linked Data through a preliminary review. With this research, some conceptual differences between the terms were identified, and opportunities were presented for future work.*

Keywords: *Linked Enterprise Data, Enterprise Linked Data, Linked Data, Semantic Technology*

Resumo. *As tecnologias presentes na web semântica vêm ganhando destaque nas pesquisas nos últimos anos, mas pouco é pesquisado sobre um viés empresarial. As empresas estão começando a se interessar pelas tecnologias semânticas, mas ainda há receio em se aproximar dessa área por conta de poucas pesquisas neste âmbito. O Linked Enterprise Data e o Enterprise Linked Data são conceitos que tratam da incorporação do Linked Data em um ambiente corporativo. Nesse sentido, o objetivo deste trabalho é apresentar as definições encontradas, através de uma revisão preliminar da literatura para Linked Enterprise Data e Enterprise Linked Data. Com a realização deste estudo, foram identificadas algumas diferenças conceituais entre os termos, e apresentadas oportunidades para trabalhos futuros.*

Palavras-Chave: *Linked Enterprise Data, Enterprise Linked Data, Linked Data, Tecnologia Semântica*

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 semântica é apresentada como uma extensão da web atual, sendo que há um número crescente de pesquisas que buscam potencializar o uso da mesma (Berners-Lee, Hendler & Lassila, 2001). O conceito de Linked Data (LD) relaciona-se com a web semântica por utilizar de duas tecnologias presentes na arquitetura da web semântica, sendo estes: o RDF (*Resource Description Framework*) que é um modelo de dados simples para descrever recursos e seus relacionamento. Estes recursos são representados como triplas (recurso-propriedade-valor), sendo que o valor pode ser outro recurso. Já o URIs (*Uniform Resource Identifier*) que é utilizado como identificador único de qualquer recurso, como: pessoas, lugares, documentos, conceitos, etc. Deste modo, os conjuntos de dados RDF podem ser associados a outros conjuntos de dados para formar uma grande base de dados conectados. Esse movimento está chamando atenção do meio empresarial, pois possibilita às empresas fazer uso da web semântica com intuito de padronizar seus dados, minimizando assim os problemas relacionados à gestão dos mesmos.

As empresas possuem diversos problemas de gestão de dados, como por exemplo, conjunto de dados oriundos de sistemas de informações independentes, dados desatualizados, incompletos e/ou inconsistentes. Estes problemas podem impactar diretamente na receita das empresas (Inria, 2017).

No mercado existem soluções que minimizam o problema da gestão de dados, por exemplo, *Business Intelligence* (BI), *Master Data Management* (MDM), *Service Oriented Architecture* (SOA) e *Big Data*, no entanto essas soluções possuem algumas limitações vinculadas, principalmente, a grande quantidade de conjunto de dados que exigem abordagens tecnológicas complexas e de alto custo (Inria, 2017).

Por sua vez o *Linked Data* pode minimizar tais limitações e dar semântica aos dados (Berners-Lee et al. 2001). Segundo pesquisa do Gartner Group (2017), há uma tendência de que as máquinas em breve tomarão decisões que, anteriormente, só podiam ser tomadas por humanos. Percebe-se, portanto, que o *Linked Enterprise Data* tem grande valia neste âmbito por contribuir agregando semântica aos dados empresariais.

O conceito de *Linked Enterprise Data* (LED) é uma extensão do conceito de *Linked Data*, com foco empresarial. As empresas podem obter benefícios com o uso deste conceito, por meio de: interoperabilidade, homogeneidade, proveniência, agilidade, coerência, acesso, identificação e governança (Galkin, Auer & Scerri, 2016). O LED também possibilita que as empresas possam consumir os dados presentes na nuvem de *Linked Open Data* (LOD), ou seja,

podem fazer uso de *datasets* disponibilizados com licença aberta, com o intuito de obter uma nova percepção de análise de dados (Glachs, Schaffert & Bauer, 2012).

O objetivo deste estudo é apresentar uma revisão preliminar da literatura a fim de identificar qual dos termos é mais adequado para referenciar o *Linked Data* no contexto empresarial, sendo que na literatura são encontrados os termos “*Linked Enterprise Data*” e “*Enterprise Linked Data*”.

Para realizar este estudo se utilizou dos termos de busca “*Linked Enterprise Data*” e “*Enterprise Linked Data*”. A busca foi realizada com os termos entre aspas duplas, a fim de buscar exatamente os termos contidos nos documentos, dando a preferência aos artigos completos dos últimos cinco anos e em língua inglesa. Foram utilizadas sete bases de dados para a realização desta pesquisa (SCOPUS, WEB OF SCIENCE, SCINCE DIRECT, ACM, SPRINGER, IEEE E EBESCO) a fim de identificar quais são as definições apresentadas para *Linked Enterprise Data* e *Enterprise Linked Data*, e responder a seguinte pergunta de pesquisa “Qual a diferença entre *Linked Enterprise Data* e *Enterprise Linked Data*? ”.

2 REVISÃO DE LITERATURA

O *Linked Enterprise Data* é um conceito emergente e que está chamando atenção das empresas (Galkin, Auer & Scerri, 2016; Blumauer, 2014). O *Linked Enterprise Data* é definido por Galkin, Auer e Scerri (2016, [informar página](#)) como “uma estrutura base para incorporar tecnologias semânticas em ambientes empresariais de TI”. O uso do LED está focado no consumo de dados, no entanto, o baixo índice de estudos relacionados ao tema traz insegurança as empresas que desejam fazer uso destas tecnologias. Esse é um dos motivos que faz com que o LED não seja tão difundido (Graube, Carnerero, Uriarte & Urbas, 2013; Li & Zhai, 2016; Glachs, Schaffert & Bauer, 2012). Contudo, esse tema é apresentado como uma alternativa para que as empresas possam se libertar das tecnologias tradicionais usando de *datasets* presentes na nuvem de *Linked Open Data* e agregando-os com seus bancos de dados a fim de obter-se benefícios. Outro ponto destacado na literatura é que com o uso desta tecnologia as empresas podem obter um alto desempenho sobre os sistemas de informação em geral (Li & Zhai, 2016).

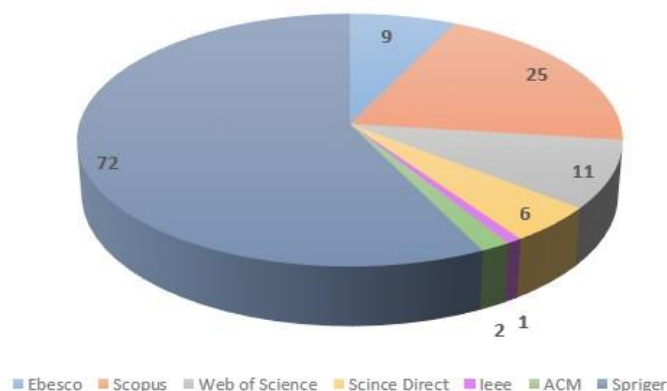
O *Enterprise Linked Data* é um conceito próximo ao *Linked Enterprise Data*, no entanto, para *Enterprise Linked Data* não foram encontradas definições que deixassem claro seu papel. A literatura apresenta que a abordagem do LD no meio empresarial vem se mostrando interessante as empresas, principalmente nos quesitos de gerenciamento de informação e interoperabilidade dos dados (Blumauer, 2014). É possível observar na literatura que o

Enterprise Linked Data é utilizado em cinco diferentes cenários: (i) na interoperabilidade através do uso de web services semânticos, (ii) na apresentação do mesmo junto com conceitos da web semânticas ou no conceito de otimização, (iii) na integração com as coisas, ou seja, com o uso da Internet das Coisas utilizando-se de sensores, (iv) em sistemas de gestão de conhecimento e (v) na integração de mídias, por exemplo, stream e músicas (Pinto & Parreiras, 2014).

3 DISCUSSÃO

No estudo realizado foi observado que o termo de busca “Enterprise Linked Data” resultou em um número menor de resultados e tais resultados estão inclusos na busca realizada com o termo “Linked Enterprise Data”, deste modo optou-se por apresentar apenas dos resultados encontrados com o termo de busca “Linked Enterprise Data”, por ter um maior número de documentos recuperados. A seguir é apresentado no Gráfico 1 com o resultado de documentos recuperados das sete bases de dados nos últimos cinco anos.

Gráfico 1 – Pesquisa nas Bases de Dados



Fonte: Autor, 2016

Para a coleta desses documentos foram aplicados dois filtros, o primeiro filtro dos últimos cinco anos e o segundo filtro somente documentos em inglês, resultando um total de 126 documentos. Desse total foram excluídos os documentos duplicados e indisponíveis, resultando em um total de 37 documentos disponíveis, sendo que foram analisados para a realização deste estudo apenas 22 documentos devido a pesquisa estar em andamento.

Com a realização da análise inicial foi encontrado um único documento Pinto e Parreiras (2014) que aborda o estado da arte sobre o *Enterprise Linked Data* sendo que o mesmo

apresenta uma revisão sistemática. Porém, o autor realizou a busca nas bases de dados utilizando três sintaxes e três termos de busca diferentes a fim de verificar o estado da arte de *Linked Data* em um contexto empresarial, sendo que o estudo apresentou uma busca desde o ano de 2010 até 2014.

Observou-se que além do trabalho de Pinto e Parreiras (2014), que existem diferentes ramificação para os estudos que abordam *Linked Enterprise Data*, sendo estes: (i) a segurança em que as empresas devem ter ao utilizar o LED, (ii) a aplicabilidade em empresas através de estudos de casos, (iii) estudos visando o uso de LD para a criação de grafos de conhecimento através de bases de conhecimento interligadas no contexto empresas, (iv) estudos sobre a Internet das Coisas (IoT) usufruindo do LD no meio empresas, (v) o uso de LD para a gestão do conhecimento e auxílio na tomada de decisão, assim como a possibilidade da criação de uma nuvem de dados privada, a fim de que as empresas possam armazenar seus *datasets* e possam usufruir de todos os benefícios propostos pelo LD.

No que se remete a pergunta de pesquisa “Qual a diferença entre *Linked Enterprise Data* e *Enterprise Linked Data*?”, foi possível identificar na literatura que não há diferença em sua aplicabilidade, na sua essência e nos seus benefícios as empresas que utilizam deste meio. No trabalho de Pinto e Parreiras (2014), os autores identificaram os dois termos, no entanto os autores relatam que não há diferenças entre os termos. Deste modo, recomenda-se a utilização da nomenclatura *Linked Enterprise Data (LED)*, pois grande parte dos documentos analisados abordam o tema utilizando-se desta nomenclatura e por ser o único termo que foi encontrado uma definição.

4 CONCLUSÃO

Neste trabalho foi possível concluir que o tema *Linked Enterprise Data* é emergente e possibilita vários caminhos para realização de pesquisas nesta área. É possível identificar que o *Linked Data* é um tema conhecido a algum tempo, mas no que se diz respeito ao contexto empresarial é novo, sendo observada a viabilidade de pesquisas no âmbito de *Linked Enterprise Data*, possibilitando assim o surgimento de outras pesquisas proporcionando uma maior segurança às empresas que desejam fazer uso deste conceito e suas tecnologias semânticas.

Para trabalhos futuros sugere-se outra revisão sistemática e uma revisão bibliométrica, atualizando o estado da arte a fim de verificar outras aplicações do *Linked Data* no contexto empresarial, assim como apresentar quais os países e universidades que estão pesquisando o tema, outra possibilidade de trabalho é a realização de uma pesquisa sobre a construção de uma

nuvem particular para as empresas. Sugere-se ainda realizar um estudo abordando o benefício do LED no aspecto da gestão da informação e na tomada de decisão, assim como realizar uma pesquisa abordando quais *datasets* contidos na nuvem LOD podem ser utilizados em nível empresarial e em quais segmentos, outro ponto que pode ser explorado é uma pesquisa sobre as soluções existentes no mercado e compara-las com o *Linked Enterprise Data*.

REFERÊNCIAS

- Berners-Lee, T., Hendler, J., and 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.
- Blumauer, A. (2014). SKOS as a Key Element in Enterprise Linked Data Strategies. In *International Semantic Web Conference (Industry Track)*.
- Pinto, V. A., & Parreiras, F. S. (2014, October). Enterprise linked data: A systematic mapping study. In *International Conference on Conceptual Modeling* (pp. 253-262). Springer International Publishing.
- Li, H., & Zhai, J. (2016, June). Constructing Investment Open Data of Chinese Listed Companies Based on Linked Data. In *Proceedings of the 17th International Digital Government Research Conference on Digital Government Research* (pp. 475-480). ACM.
- Galkin, M., Auer, S., & Scerri, S. (2016, October). Enterprise Knowledge Graphs: A Backbone of Linked Enterprise Data. In *Web Intelligence (WI), 2016 IEEE/WIC/ACM International Conference on* (pp. 497-502). IEEE.
- Graube, M., Ortiz, P., Carnerero, M., Lázaro, O., Uriarte, M., & Urbas, L. (2013, December). Flexibility vs. security in linked enterprise data access control graphs. In *Information Assurance and Security (IAS), 2013 9th International Conference on* (pp. 13-18). IEEE.
- Glachs, D., Schaffert, S., & Bauer, C. (2012). Interlinking Media Archives with the Web of Data. In *I-SEMANTICS (Posters & Demos)* (pp. 17-21).
- Bizer, C. (2009). The emerging web of linked data. *IEEE intelligent systems*, 24(5).
- 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.
- Inria, Linked Enterprise Data Principles, Uses and Benefits. Recuperado em 19 de abril de 2017, de <<https://www.inria.fr/content/download/18171/514403/>>
- Gartner Group, Gartner Reveals Top Predictions for IT Organizations and Users for 2016 and Beyond. Recuperado em 19 de abril de 2017, de <<http://www.gartner.com/newsroom/id/3143718/>>