

UMA REVISÃO PRELIMINAR SOBRE A DIFUSÃO DE DADOS CONECTADOS NO ÂMBITO EMPRESARIAL

A PRELIMINARY REVIEW ON THE DIFFUSION OF LINKED DATA IN THE ENTERPRISE

Murilo Silveira Gomes⁽¹⁾, Lidiane Visintin⁽¹⁾, Fernando Álvaro Ostuni Gauthier⁽²⁾

(1) Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento, Universidade Federal de Santa Catarina, {lilo.flp lidiane.visintin}@gmail.com.

(2) Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento, Universidade Federal de Santa Catarina, fernando.gauthier@ufsc.br.

Resumo: As empresas necessitam estar atentas as mudanças no cenário mundial, para que possam responder rapidamente as mudanças de mercado. Para isso, a Web contribuiu possibilitando dinamismo, impactando com isso em um aumento significativo no volume de dados, que podem ser explorados, com o intuito de se obter benefícios. O Linked Enterprise Data é apresentado como um formato de dados que pode auxiliar os gestores na tomada de decisão, oportunizando explorar dados internos e externos a empresa, porém este conceito não é muito explorado no âmbito empresarial. Por este motivo o presente estudo objetiva-se em discutir os fatores que impedem a difusão de Linked Data neste contexto. Deste modo, observa-se que a cultura organizacional e alinhamento estratégico, entre outros, são alguns dos fatores encontrado para que não ocorra a difusão do Linked Enterprise Data. O que permite concluir que Linked Enterprise Data está começando a chamar atenção das empresas, mas se faz necessário mais trabalhos que abordem as reais contribuições ao meio empresarial, para difundir o Linked Data neste contexto.

Palavras-chave: Empresa; Dados Conectados; Fatores de Impacto.

Abstract: Companies need to be aware of changes in the global landscape, so they can respond quickly to market changes. For this purpose, the Web has contributed to the dynamism, thus impacting on a significant increase in the volume of data that can be exploited in order to obtain benefits. Linked Enterprise Data is presented as a concept that can help managers in decision making, allowing the exploration of internal and external data to the enterprise, but this concept is not much explored in the business scope. For this reason, the present study aims at discussing the factors that prevent the diffusion of Linked Data in this context. In this way, it is observed that the organizational culture and strategic alignment, among others, are some of the factors found so that the dissemination of Linked Enterprise Data does not occur. This allows us to conclude that Linked Enterprise Data is beginning to attract attention from companies, but more work is needed to address the real contributions to the business environment, to disseminate Linked Data in this context.

Keywords: Enterprise; Linked Data; Impact Factors.

1 Introdução

As empresas têm a necessidade de se tornarem dinâmicas, devido as mudanças que ocorrem nos mais diversos segmentos de mercado (HU; SVENSSON, 2010). Para isso otimizar o uso de seus dados e explorá-los, pode possibilitar reagir as rápidas mudanças, gerando benefícios para a tomada de decisão.

Há algumas soluções que buscam auxiliar na integração de dados e informações, bem como no processo de tomada de decisão, sendo algumas destas soluções: MDM (Master Data Management), SOA (Service-Oriented

Architecture) e BI (Business Intelligence), mas o grande desafio ainda se encontra em como integrar grandes quantidades de dados heterogêneos (ANTIDOT, 2017). Por sua vez, os dados conectados (Linked Data) podem minimizar tal desafio e dar semântica aos dados (BERNERS-LEE, 2001).

O uso dos princípios de dados conectados no contexto empresarial é interessante, pois está área de pesquisa promete integrar e consolidar fontes de dados heterogêneos, possibilitando analisar dados empresariais, com dados disponibilizados na Web, reutilizando de vocabulários conhecidos, como: *Friend-*

of-a-Friend (FOAF), Dublin Core (DC) e Simple Knowledge Organization System (SKOS) (WEICHSELBRAUN; STREIFF; SCHARL, 2015).

Analisado o contexto apresentado e verificando que o conceito de dados abertos conectados tem ganhado grande ênfase no setor público, este trabalho tem por objetivo discutir quais são os principais fatores que impedem a difusão de dados conectados no âmbito empresarial.

Na sequência, são apresentadas as seções deste trabalho: na seção 2 aborda-se o contexto de *Linked Enterprise Data*; na seção 3 é apresentada uma breve descrição dos procedimentos metodológicos; na seção 4 apresenta-se uma discussão preliminar dos resultados obtidos com a análise realizada; e por fim, na seção 5 são apresentados os encaminhamentos para trabalhos futuros e considerações finais.

2 Linked Enterprise Data

O Linked Enterprise Data (LED) é uma extensão do conceito de dados conectados, com foco empresarial. Que é definido como: “uma estrutura base para incorporar tecnologias semânticas em ambientes empresariais de TI” (GALKIN; AUER; SCERRI, 2016). Contudo, esse tema é apresentado como uma alternativa para que as empresas possam se libertar das tecnologias tradicionais. Assim como podem obter

benefícios como: interoperabilidade e 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 dados abertos conectados (Linked Open Data), 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). Outro ponto destacado na literatura é que com o uso deste conceito as empresas podem obter um alto desempenho sobre os sistemas de informação em geral (LI; ZHAI, 2016).

3 Procedimentos Metodológicos

Para a realização desta revisão literária preliminar, utilizou-se de 6 artigos que abordam o tema *Linked Enterprise Data*, sendo que os mesmos são apresentados entre os anos de 2010 a 2016. Os artigos foram selecionados com base no número de citações e priorizado alguns dos últimos anos, devido a fornecer uma visão mais atual das pesquisas na área. No Quadro 1 apresenta-se os artigos selecionados, destacando o título, ano de publicação, a quantidade de citações e os autores, apresentados na ordem de maior número de citações.

Quadro 1 – Artigos Selecionados para análise

Autores	Título	Ano	Qtd. Citação	Objetivo
HU, SVENSSON	A Case Study Of Linked Enterprise Data	2010	10	Apresenta um estudo piloto em uma organização internacional, tem por objetivo criar um espaço compartilhado que proporcione a integração de dados inter-organizacionais em Linked Data.
GRAUBE et al.	Flexibility Vs. Security In Linked Enterprise Data Access Control Graphs	2013	3	Apresentar o projeto ConVantage, focado na segurança dos dados em Linked Data em domínio empresarial.
BIANCHINI ANTONELLIS MELCHIORI	A Linked Data Perspective For Collaboration In Mashup Development	2013	3	Apresenta a abordagem e o protótipo LINKSMAN(Linked Data Supported Mashup Collaboration) baseando-se na integração de dados internos e externos a empresa utilizando de

				Linked Data, a fim de localizar novos colaboradores.
WEICHSELBRAUN STREIFF SCHARL	Consolidating Heterogeneous Enterprise Data For Named Entity Linking And Web Intelligence	2015	2	Apresenta o componente Recognize, que tem por objetivo identificar entidades nomeadas com uso de bases de dados conectados, a fim de promover a interoperabilidade de conjunto de dados públicos abertos.
TANEJA et al.	Linked Enterprise Data Model And Its Use In Real Time Analytics And Context-Driven Data Discovery	2015	2	Aborda as deficiências para o gerenciamento de IoT e apresentam a abordagem desenvolvida pelos autores denominada de LEDM (Linked Enterprise Data Model) desenvolvida para atender o domínio de IoT com foco em Big Data. O LEDM é baseado nos princípios de Linked Data com foco na interoperabilidade entre sistemas e subsistemas.
GALKIN AUER SCERRI	Enterprise Knowledge Graphs - A Backbone Of Linked Enterprise Data	2016	0	Apresenta um estudo focado em EKG (Enterprise Knowledge Graphs) utilizando de Linked Data para a integração de diferentes bases de conhecimento, utilizando de três abordagens (Indefinido, Transição e Federado).

Fonte: Autores, (2017)

Na próxima seção serão discutidos os artigos selecionados, apresentando os fatores identificados que evitam a difusão do LED.

4 Resultados e Discussões Preliminares

Com a análise realizada, foi possível identificar que todos os autores deixam claro que a web semântica e suas tecnologias podem beneficiar o meio empresarial, perceptivelmente através da interoperabilidade.

Com a realização desta análise também é possível perceber que o foco de discussão dos artigos analisados de dados conectados no âmbito empresarial, englobam áreas como: IoT, reconhecimento de entidades nomeadas, grafos de conhecimento, integração de dados entre outros, como pode ser observado no Quadro 1.

A literatura apresenta que o uso do LED está focado no consumo de dados, no entanto, o baixo índice de estudos

relacionados ao tema geram uma certa “insegurança” por parte das empresas que desejam fazer uso deste conceito. Esse é um dos motivos que faz com que o LED não seja tão difundido (GRAUBE et al., 2013, LI; ZHAI, 2016, GLACHS, SCHAFFERT; BAUER, 2012). Porém, há esforços em termos de pesquisa para difundir o LED, mas é possível perceber outros fatores que influenciam negativamente o aproveitamento deste contexto.

Com base nos artigos analisados, destaca-se e discute-se os seguintes fatores identificados: (1) cultura organizacional, (2) alinhamento estratégico, (3) trabalhos apenas com foco em domínio público e no meio acadêmico e (4) confusão quanto aos conceitos de dados abertos e dados conectados.

- Cultura Organizacional: Percebe-se que de nada adianta um ou dois indivíduos quererem implantar dados conectados em uma organização, caso não se tem uma conscientização e o

interesse das pessoas envolvidas no processo de implantação de dados conectados (principalmente os gestores da empresa) e dos benefícios que poderão ser obtidos com o uso deste conceito.

Nota-se ainda que a maioria dos usuários corporativos não tem uma “mente semântica” (HU; SVENSSON, 2010), fazendo com que o medo da mudança crie barreiras para aderir ao conceito de LED. Também se evidencia que em sua maioria os gestores das empresas estão preocupados apenas com que os dados sejam fornecidos de forma oportuna e precisa, independente das tecnologias utilizadas.

- Alinhamento estratégico: Constata-se que o gerenciamento de dados possui por finalidade o melhoramento da eficiência do *Core Business* da empresa (HU; SVENSSON, 2010), para isso se faz necessário conhecer os processos da empresa para que haja um alinhamento entre a estratégia empresarial e o uso dos dados. Da mesma forma, que os links entre dados não devem ser aleatoriamente criados, independentes dos processos de negócios (HU; SVENSSON, 2010).

Verifica-se ainda que alinhar dados a estratégia da empresa, possibilita o relacionamento dos dados empresariais internos ou externos, proporcionando assim novas perspectivas para a tomada de decisões.

- Trabalhos apenas com foco em domínio público e no meio acadêmico: Percebe-se que há pesquisas em meio acadêmico, que muitas vezes não chegam a ser testadas no âmbito empresarial, ou seja, aparenta haver um *gap* entre o que a academia desenvolve e o que as empresas necessitam para fazer uso (GOMES, 2017). Também se percebe que as pesquisas aplicadas em sua maioria estão voltadas para o domínio público.

- Confusão quanto aos conceitos de dados abertos e dados conectados: Para se fazer uso de dados conectados é preciso ter claro o seu conceito. Onde percebe-se que devido ao movimento de Governo Aberto (MEIJER; CURTIN; HILLEBRANDT, 2012), o conceito de dados abertos ganhou grande ênfase. No entanto quando se menciona dados conectados percebe-se que há uma confusão quanto aos conceitos apresentados, pois dados abertos podem ser disponibilizados a todos sem que estejam em um formato RDF. Ao mesmo tempo, que os dados podem ser conectados sem estarem livremente disponíveis para serem utilizados ou distribuídos, este é o caso de dados conectados. Assim, como se tem dados abertos conectados, que necessitam ter uma licença aberta bem como fazer uso do formato RDF, conforme pode ser observado no Quadro 2.

Quadro 2 – Justaposição de dados abertos e fechados

Representação/Grau de abertura	Fechado	Aberto
Modelo de dados estruturados Ex.: XML, CSV, SQL, etc...	Dados	Dados Abertos
Modelo de Dados RDF	Dados Conectados	Dados Abertos Conectados

Fonte: AUER et al., (2014)

Analisando ainda os artigos, observa-se que são apresentadas possibilidades de pesquisas futuras, sendo que a maioria dos autores

apresentam possibilidades de pesquisas focando na continuidade de seu próprio trabalho. Conforme segue no Quadro 3.

Quadro 3 – Relação de trabalhos futuros

Título	Ano	Trabalhos Futuros
A Case Study Of Linked Enterprise Data	2010	Realizar uma avaliação em grande escala com usuários convidados de diferentes departamentos e regiões

		geográfica. Identificar outros cenários de uso que possa demonstrar valor a nova diversidade de usuários.
Flexibility Vs. Security In Linked Enterprise Data Access Control Graphs	2013	Criar regras políticas de acesso a fim de minimizar as barreiras impostas pelas empresas para começar a obter informações de ambientes de dados conectados.
A Linked Data Perspective For Collaboration In Mashup Development	2013	Identificar novos padrões de colaboração adicionais usando de novas fontes externas, estender a abordagem para considerar a experiência externa dos desenvolvedores e realizar a avaliação com base no protótipo.
Consolidating Heterogeneous Enterprise Data For Named Entity Linking And Web Intelligence	2015	Melhorar a performance de desambiguação da Recognize, considerando bases estruturadas mais complexas no processo de vinculação. Otimizar e avaliar perfis de desambiguação que funcionem com fontes de dados públicas, explorar as opções de combinação de dados empresariais conectados com dados abertos conectados, fornece avaliações a outros tipos de entidades e ampliar a abordagem para outros idiomas.
Linked Enterprise Data Model And Its Use In Real Time Analytics And Context-Driven Data Discovery	2015	Os autores não evidenciam as oportunidades de trabalhos futuros.
Enterprise Knowledge Graphs - A Backbone Of Linked Enterprise Data	2016	Melhorar o conceito de EKG em domínios técnicos e empresariais e especificar melhor as características do EKG, a fim de aprofundar o estudo de modo que as empresas possam fazer uso.

Fonte: Autores, (2017)

Verifica-se que para a difusão do LED, há a necessidade de trabalhos que abordem mais claramente os benefícios e os ganhos reais que são obtidos ao utilizar de dados conectados. Deste modo, constata-se também que o fator primordial para a difusão de dados conectados no âmbito empresarial são as pessoas (HU; SVENSSON, 2010), pois sem elas não há a compreensão dos conceitos, bem como o interesse em utilizá-lo.

5 Considerações Finais

Com a realização deste trabalho observou-se que dados conectados agregam um valor notável no domínio público (BIZER; HEATH; BERNERS-LEE, (2009). No entanto quando se fala do domínio empresarial nota-se que os seguintes fatores impedem sua difusão: cultura organizacional, alinhamento estratégico, trabalhos apenas com foco em domínio público e no meio acadêmico e uma confusão quanto aos conceitos de dados abertos e dados conectados.

Identificou-se que os primeiros trabalhos que abordam LED possuem

foco em publicar dados, com o intuito de favorecer o compartilhamento e a interoperabilidade dos dados, em um espaço inter-organizacional, porém o foco dos trabalhos mais recentes está na interoperabilidade de sistemas, para o consumo de dados.

Notou-se também que o artigo mais citado encontrado na literatura apresenta uma abordagem mais abrangente sobre as possibilidades no uso do LED, pois ele engloba além questões técnicas, questões estratégicas, que são cruciais para as empresas.

Outros pontos que foram observados é que todos os trabalhos relatam a possibilidade de fazer uso de dados internos e externo em um mesmo ambiente e também há relatos de uma estrutura de nuvem de dados privada. No entanto não são relatadas propostas para materializar estas possibilidades e percebe-se que estas propostas são viáveis de serem desenvolvidas em pesquisas futuras.

Referências

- ANTIDOT (França). **Enterprise Data Principles, uses and benefits**. 2012. Disponível em: <<http://www.antidot.net/wp-content/uploads/2012/11/LinkedEnterpriseData-WP-en-v2.2.pdf>>. Acesso em: 03 ago. 2017
- AUER, Sören; LEHMANN, Jens; NGOMO, Axel-Cyrille Ngonga. Introduction to linked data and its lifecycle on the web. In: **Reasoning Web. Semantic Technologies for the Web of Data**. Springer Berlin Heidelberg, 2011. p. 1-75.
- BERNEERS-LEE, Tim. The Semantic Web: A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities. **Scientific American**, v. 284, n. 5, p. 34-43, 2001.
- BIANCHINI, Devis; DE ANTONELLIS, Valeria; MELCHIORI, Michele. A linked data perspective for collaboration in mashup development. In: **Database and Expert Systems Applications (DEXA), 2013 24th International Workshop on**. IEEE, 2013. p. 128-132.
- BIZER, Christian; HEATH, Tom; BERNERS-LEE, Tim. Linked data-the story so far. **Semantic services, interoperability and web applications: emerging concepts**, p. 205-227, 2009.
- GALKIN, Mikhail; AUER, Sören; SCERRI, Simon. Enterprise Knowledge Graphs: A Backbone of Linked Enterprise Data. In: **Web Intelligence (WI), 2016 IEEE/WIC/ACM International Conference on**. IEEE, 2016. p. 497-502.
- GLACHS, Dietmar; SCHAFFERT, Sebastian; BAUER, Christoph. Interlinking Media Archives with the Web of Data. In: **I-SEMANTICS (Posters & Demos)**. 2012. p. 17-21.
- GOMES, Murilo Silveira. **PROPOSTA DE ARQUITETURA PARA ECOSSISTEMA DE INOVAÇÃO EM DADOS ABERTOS**. 2017. 104 f. Dissertação (Mestrado) - Curso de Engenharia e Gestão do Conhecimento, Centro Tecnológico, Universidade Federal de Santa Catarina, Florianópolis, 2017. Disponível em: <<http://btd.egc.ufsc.br/wp-content/uploads/2017/04/Murilo-Gomes.pdf>>. Acesso em: 05 ago. 2017.
- GRAUBE, Markus et al. Flexibility vs. security in linked enterprise data access control graphs. In: **Information Assurance and Security (IAS), 2013 9th International Conference on**. IEEE, 2013. p. 13-18.
- HU, Bo; SVENSSON, Glenn. A case study of linked enterprise data. **The Semantic Web-ISWC 2010**, p. 129-144, 2010.
- LI, Hongqin; ZHAI, Jun. 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**. ACM, 2016. p. 475-480.
- MEIJER, Albert J.; CURTIN, Deirdre; HILLEBRANDT, Maarten. Open government: connecting vision and voice. **International Review of Administrative Sciences**, v. 78, n. 1, p. 10-29, 2012.
- TANEJA, Kunal et al. Linked enterprise data model and its use in real time analytics and context-driven data discovery. In: **Mobile Services (MS), 2015 IEEE International Conference on**. IEEE, 2015. p. 277-283.
- WEICHSELBRAUN, Albert; STREIFF, Daniel; SCHARL, Arno. Consolidating heterogeneous enterprise data for named entity linking and Web Intelligence. **International Journal on Artificial Intelligence Tools**, v. 24, n. 2, p. 1540008, 2015.