

XVIII ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2017

GT-4 – Gestão da Informação e do Conhecimento

FLUXO INFORMACIONAL PARA REDUÇÃO DE DADOS DUPLICADOS NA PLATAFORMA LATTES

Marcos de Souza (Universidade Federal de Minas Gerais – UFMG)

Renato Rocha Souza (Universidade Federal de Minas Gerais – UFMG / Fundação Getúlio
Vargas – FGV)

INFORMATION FLOW FOR REDUCING DATA DUPLICATED ON THE LATTES PLATFORM

Modalidade da Apresentação: Pôster

Resumo: A Arquitetura da Informação é a arte e a ciência de estruturar e organizar sistemas de informação com o intuito de ajudar as pessoas a atingirem o seu objetivo. A plataforma de currículo Lattes permite aos estudantes, professores e pesquisadores realizarem o registro de publicações científicas como artigos completos publicados em periódicos, artigos aceitos para publicação, livros e capítulos; trabalhos publicados em anais de eventos nas modalidades de artigo completo, resumo ou resumo expandido. O que se tem observado junto ao ambiente virtual de currículo são dados duplicados de registros de publicações científicas mediante aos diferentes comportamentos de seus próprios usuários que acabam por realizar tais registros sem conhecimento. Partindo desse princípio, questiona-se: de que forma seria possível minimizar o quantitativo de registro de dados duplicados de publicações científicas junto a plataforma de currículo Lattes? Essa pesquisa possui como objetivo geral propor uma mudança no fluxo informacional na plataforma de currículo Lattes, tais como as partes envolvidas no processo de diferentes tipos de publicações, seja agências regulamentadoras, instituições de ensino, editoras e periódicos. Busca refletir até quando o usuário da plataforma poderia controlar suas informações científicas sem gerar duplicidades de dados e os impactos que tal mudança de fluxo poderia ocasionar. Para elaboração da pesquisa, foram utilizados dissertação, livros e artigos científicos de bases de dados como SCIELO e Google Acadêmico que serviram como suporte teórico e embasamento para os resultados e discussões da pesquisa. O fluxo informacional proposto para cadastro das informações das publicações científicas junto a plataforma Lattes segue quatro etapas: 1 - editoras, periódicos ou instituições de ensino superior organizam as publicações científicas; 2 - solicitarem os respectivos números indexadores aos órgãos responsáveis; 3 – realizarem o cadastro das produções científicas junto a plataforma Lattes; 4 – atualização de forma automatizada dos currículos dos autores envolvidos em uma determinada obra.

Palavras-Chave: Arquitetura da Informação; Fluxo Informacional; Dados.

Abstract: Information Architecture is the art and science of structuring and organizing information systems in order to help people achieve their goal. The Lattes curriculum platform allows students, professors and researchers to register scientific publications as complete articles published in

periodicals, articles accepted for publication, books and chapters; papers published in annals of events in the modalities of full article, abstract or expanded abstract. What has been observed next to the virtual environment of curriculum are duplicated data of records of scientific publications through the different behaviors of its own users that end up making such records without knowledge. Based on this principle, it is questioned: in what way could it be possible to minimize the quantitative recording of duplicate scientific publications with the Lattes curriculum platform? This research aims at proposing a change in the information flow in the Lattes curriculum platform, such as the parties involved in the process of different types of publications, be it regulatory agencies, educational institutions, publishers and periodicals. It seeks to reflect how far the platform user could control their scientific information without generating duplicate data and the impacts that such a change of flow could cause. For the elaboration of the research, we used dissertation, books and scientific articles of databases such as SCIELO and Google Scholar that served as theoretical support and basis for the results and discussions of the research. The proposed information flow for the registration of the information of the scientific publications along the Lattes platform follows four stages: 1 - publishers, periodicals or institutions of higher education organize the scientific publications; 2 - request the respective index numbers from the responsible bodies; 3 - to register the scientific productions along the Lattes platform; 4 - automated updating of the curricula of the authors involved in a particular work.

Keywords: Information Architecture; Information Science; Metadata; Information Retrieval.

1 INTRODUÇÃO

Na década de 1990 os profissionais cientistas da informação, buscando tratar e preservar documentos como elementos construídos socialmente desde sua origem, começaram a dar importância aos processos de conhecer e reconhecer a natureza dos objetos manipulados, perpassando assim, pelos interesses das instituições que produzem, utilizam, preservam ou disponibilizam informações para consulta (TRACE, 2002).

A plataforma de currículo *Lattes*¹, desenvolvida pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico² (CNPq) é uma base de dados pública de inserção de dados e recuperação da informação do progresso da vida acadêmica de estudantes, professores e pesquisadores. Destaca-se como registros de dados de publicações científicas junto a plataforma *Lattes*: artigos completos publicados em periódicos; artigos aceitos para publicação; livros e capítulos; trabalhos publicados em anais de eventos nas modalidades de artigo completo, resumo ou resumo expandido.

Davenport define dados como “observações sobre o estado do mundo” (DAVENPORT, 1998, p.19) e tais dados brutos ou entidades quantificáveis podem ser trabalhados por pessoas

1 Plataforma Lattes – Sobre a plataforma Lattes. Disponível em: <<http://lattes.cnpq.br/>>; Acesso em: 06/07/2017.

2 Conselho Nacional de Desenvolvimento Científico e Tecnológico. Disponível em: <<http://cnpq.br/>>; Acesso em: 06/07/2017.

ou pelas tecnologias da informação. A partir desse princípio, o que se tem observado é um volume de publicações científicas contidas junto aos currículos da plataforma *Lattes* com dados duplicados. Levando em consideração esse fator, questiona-se: de que forma seria possível minimizar o quantitativo de registro de dados duplicados de publicações científicas junto a plataforma de currículo *Lattes*?

Essa pesquisa possui como objetivo geral propor uma mudança no fluxo informacional na plataforma de currículo *Lattes*, tais como as partes envolvidas no processo que envolvem os diferentes tipos de publicações, seja agências regulamentadoras, instituições de ensino, editoras e periódicos. Além disso, busca refletir até quando o usuário da plataforma poderia controlar suas informações científicas no ambiente de currículos sem gerar duplicidades de dados e os impactos que tal mudança de fluxo poderia ocasionar.

Justifica-se a importância dessa pesquisa uma vez que dados duplicados em quaisquer sistemas de informação não apresentam veracidade aos números, interferindo assim, em tomadas de decisão por exemplo. Além disso, um novo fluxo informacional poderia minimizar consideravelmente a possibilidade de dados duplicados junto a plataforma *Lattes*, servindo assim, como parâmetro real para instituições de fomento ou mesmo órgãos governamentais realizarem a gestão da informação e a gestão do conhecimento.

2 DESENVOLVIMENTO

De acordo com Rezende (2013) o advento das mídias interativas e da internet potencializou o volume de dados e informações gerados em todo o mundo. De acordo com Stair (1998), O'Brian (2003) os dados são definidos como fatos em uma forma primária e não necessariamente são compostos por fatos alfanuméricos. Os dados podem possuir características diferentes, dentre elas, os dados alfanuméricos tradicionais, compostos por números e caracteres alfanuméricos e os dados audiovisuais, compostos por parágrafos, imagem, áudio, cifras, sons e a própria voz humana.

Para Felix (2003), O'Brian (2003), Oliveira (2012) um dado isolado, de qualquer elemento em uma forma bruta, apresenta um significado restrito e pouco informativo. Porém, atribuído a um contexto, um conjunto de dados pode gerar informações. Tais conceitos de dado e informação são empregados de forma intercambiável, onde dado são recursos de matéria-prima e conseqüentemente, convertidos em um contexto de significado e sendo de serventia para determinado usuário, são considerados informação (O'BRIAN; 2003).

Felix (2003) destaca que o conceito de conhecimento perpassa por uma sequência de etapas desde os dados brutos até as informações pré-selecionadas, destacando assim, regras e diretrizes dessa transformação. Ainda segundo o autor, um conjunto de conhecimento permite gerar uma base de conhecimento, que pode ser utilizada a qualquer momento.

Davenport (1998) destaca que a constituição de um determinado conhecimento está na informação. Apesar de ser algo de muita valia, torna-se difícil de ser gerenciada. O valor da informação é importante pois, inserida em um contexto e interpretada por alguém, acrescenta sabedoria ao indivíduo.

A Arquiteta da Informação está relacionada ao fluxo efetivo da informação representada por desenhos informacionais conforme aponta (MACEDO, 2005). Lima (2016) salienta que o termo Arquitetura da Informação associada a Arquitetura tradicional segue os processos do conhecimento, desde o planejamento e sua estruturação, visando assim, proporcionar espaços funcionais e eficientes. Na Arquitetura da Informação estão relacionados quatro pilares fundamentais para ajudarem as pessoas e atenderem as necessidades da informação, sendo elas: 1 – investigação; 2 – análise; 3 – desenho e; 4 – implementação (LIMA-MARQUES; MACEDO, 2006).

A Arquitetura da informação é responsável por auxiliar na organização das informações projetadas em sistemas de informação de forma que seus usuários possam realizar os processos traçados. “[...] é a arte e a ciência de estruturar e organizar sistemas de informação com o intuito de ajudar as pessoas a atingirem o seu objetivo” (BAILEY, 2003, p.9).

Estão relacionadas também disciplinas correlatas a Arquitetura da Informação, dentre elas, destacam-se: design gráfico e design de interação, biblioteconomia e ciência da informação, jornalismo, engenharia de usabilidade, marketing, ciência da computação, redação técnica, arquitetura e gerenciamento de produtos conforme aponta (MORVILLE; ROSENFELD, 2006).

3 METODOLOGIA

Quanto aos procedimentos metodológicos, a pesquisa se classifica de acordo com Gil (2010) como: do ponto de vista de sua natureza, como básica; quanto a abordagem do problema, como qualitativa; do ponto de vista dos objetivos, como exploratória; do ponto de vista dos procedimentos técnicos, como pesquisa bibliográfica. Para elaboração da pesquisa, foram utilizados dissertação, livros e artigos científicos de bases de dados como SCIELO e Google

Acadêmico que serviram como suporte teórico e embasamento para os resultados e discussões da pesquisa.

4 RESULTADOS E DISCUSSÕES

No mundo acadêmico, instituições de ensino ou editoras utilizam o *International Standard Book Number* (ISBN) e o *International Standard Serial Number* (ISSN) como números vinculados as diferentes obras científicas como anais de eventos, livros e periódicos. O *Digital Object Identifier*³ (DOI) é representado por um sistema de identificação especificamente para conteúdo como livros, artigos, documentos, ambos em formato digital. A obra, imagens ou textos, recebe um número único que é atribuído a um link de acesso junto com o preenchimento dos metadados. Posteriormente, o autor da obra receberá o número único onde consequentemente fará o cadastro da obra junto a plataforma *Lattes* e aparecerá todos os metadados cadastrados pelo periódico ou editora ao qual o autor tenha realizado a publicação.

Um dos grandes problemas da plataforma *Lattes* está em permitir que o usuário final realize o cadastro de um determinado tipo de pesquisa científica. A duplicidade de dados em um mesmo currículo pode acontecer em algumas situações, como por exemplo: 1 – um autor cadastra um artigo científico publicado em um periódico no qual ainda não foi gerado um DOI; em um segundo momento, o usuário administrador do periódico responsável pelo registro do DOI realiza o cadastro dos metadados gerando uma sequência numérica do DOI; em um terceiro momento, um co-autor desse mesmo artigo atualiza o currículo *Lattes* utilizando o DOI; por fim, os metadados acabam por ser inseridos no currículo do autor principal. Caso ele não observe, na próxima atualização de seu currículo, apresentará duas publicações, sendo uma cadastrada de forma manualmente e outra mediante ao cadastro do co-autor.

Uma questão a ser contextualizada não está relacionada a lógica de programação, que certamente, poderia resolver algumas dessas problemáticas, porém, o que se é discutido nessa pesquisa vai além de uma questão técnica, pois fatores como dados cadastrados de forma errônea por usuários junto a plataforma *Lattes* também contribuem para potenciais conflitos de informações, por exemplo, estatísticos. O que se pode refletir está associado as contribuições da Arquitetura da Informação para tratamento dos dados das publicações científicas que envolvem um todo, desde as agências regulamentadoras dos ISSN e ISBN, perpassando por instituições de ensino, periódicos e editoras.

³ DOI. Disponível em: < <https://www.doi.org> >; Acesso em: 06 jul. 2017.

O atual fluxo da plataforma *Lattes* possibilita ao usuário do ambiente realizar o preenchimento de dados de produção acadêmica e de educação e popularização de ciência e tecnologia livremente ou por meio de importação de metadados externos como o DOI. O fluxo proposto para cadastro das informações das publicações científicas junto a plataforma *Lattes* seguiria algumas etapas como conforme apresentado na figura 1: 1 – as editoras, periódicos ou instituições de ensino superior, por meio de seus setores de publicações (quando existentes) ao realizarem a organização de anais de eventos, livros ou publicações de revistas científicas; 2 – responsáveis solicitarem os respectivos números indexadores aos órgãos responsáveis; 3 – os organizadores de uma determinada obra científica tornam-se responsáveis por realizarem o cadastro das produções científicas junto a plataforma *Lattes*, vinculado por uma chave primária, como o CPF por exemplo ou um link único da página de cada autor e co-autor(es) se existente; 4 – atualização de forma automatizada dos currículos dos autores envolvidos em uma determinada obra.

Figura 1: Proposta de fluxo de atualização junto a plataforma *Lattes*.



Fonte: O autor.

Embora tal proposta possa minimizar a quantidade de dados publicados na plataforma *Lattes*, uma questão para se refletir está na liberdade do autor em poder realizar os seus próprios registros científicos. Torna-se importante ressaltar que a proposta desse cenário se aplicaria somente a publicações científicas realizadas no Brasil e para estudantes, professores ou pesquisadores que utilizam a plataforma de currículo *Lattes*. Uma questão também a ser discutida está na forma como os pesquisadores aceitariam possíveis mudanças.

5 CONSIDERAÇÕES FINAIS

A contribuição da Arquitetura da Informação está em auxiliar na organização das informações projetadas em sistemas de informação por meio de diversos profissionais qualificados e em diversas áreas do conhecimento, passando assim por fases de estudos como investigação, análise, desenho e implementação.

A problemática da pesquisa foi respondida apresentando uma proposta no fluxo informacional do registro das pesquisas científicas realizadas por estudantes, professores e pesquisadores que utilizam a plataforma *Lattes*. Essa mudança reduziria consideravelmente o número de dados duplicados encontrado na base de currículos, porém, trata-se de um estudo que envolve não só os aspectos tecnológicos, mas principalmente dois aspectos: 1 – políticos: que relaciona as agências reguladoras dos ISSN e ISBN, DOI, plataforma *Lattes*, editoras, instituições de ensino e; 2 – culturais: relacionado aos aspectos de aceitação por parte dos usuários do ambiente de currículo.

Uma questão refletida está em saber se existiria um bom aceite por parte da comunidade de pesquisadores. Em contrapartida, um dos benefícios está nos dados consolidados, que podem ser utilizados para tomadas de decisão de forma estratégica por instituições de fomento ou pelos próprios órgãos governamentais como forma de gerir as informações e o conhecimento.

Sugere-se para pesquisas futuras um aprofundamento na temática, tais como estudos de viabilidade envolvendo os órgãos responsáveis e também uma pesquisa quantitativa com estudantes, professores e pesquisadores que utilizam a plataforma de currículo *Lattes* como forma de identificar o nível de aceitação com relação a possíveis mudanças no fluxo informacional, apresentando assim os prós e contras.

REFERÊNCIAS

- BAILEY, Samantha. Information architecture: a brief introduction. **The Information Architecture Institute**, Beverly, Massachusetts, v. 12, mar: 2003. Disponível em: <<http://www.iainstitute.org/sites/default/files/bailey-iaintro.pdf>>; Acesso em: 07 jul. 2017.
- DAVENPORT, Thomas Hayes. **Ecologia da informação**: por que só a tecnologia não basta para o sucesso na era da informação. São Paulo - SP: Futura, 1998. 318p.
- FELIX, Wellington. **Introdução à Gestão da Informação**. Campinas - SP: Alínea, 2003. 97p.
- GIL, Antonio Carlos. **Métodos e Técnicas de Pesquisa Social**. 6. ed. São Paulo - SP: Atlas, 2008. 200p.

XVIII ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2017
23 a 27 de outubro de 2017 – Marília – SP

LIMA, Gercina Ângela Borém de Oliveira. Arquitetura da informação. In: MIRANDA, Roberto Campos da Rocha. **Arquitetura da Informação na Câmara dos Deputados**. Brasília - DF: Edições Câmara, 2016. 282p.; p.45-62. (Série relatos de pesquisa nº 7). Disponível em: <http://bd.camara.gov.br/bd/bitstream/handle/bdcamara/30199/arquitetura_informaca_miranda.pdf?sequence=1>. Acesso em: 08 jul. 2017.

LIMA-MARQUES, Mamede; MACEDO, Flávia Lacerda Oliveira de. Arquitetura da informação: base para a gestão do conhecimento. In: TARAPANOFF, Kira. **Inteligência, Informação e Conhecimento**. Brasília - DF: IBICT, 2006. 453p.; p.241-255. Disponível em: <<http://unesdoc.unesco.org/images/0014/001469/146980por.pdf>>. Acesso em: 08 jul. 2017.

MACEDO, Flávia. Lacerda. **Arquitetura da Informação: aspectos epistemológicos, científicos e práticos**. 186f. Dissertação (Tese de Mestrado) – Universidade de Brasília, Brasília - DF, 2005. Disponível em: <<https://goo.gl/2wAmVL>>; Acesso em: 08 jul. 2017.

MORVILLE, Peter; ROSENFELD, Louis. **Information architecture for the world wide web**. 3. ed. Cambridge: O'Reilly, 2006. 504p.

O'BRIEN, James A.. **Sistemas de Informação: e as decisões gerenciais na era da internet**. 9. ed. São Paulo - SP: Saraiva, 2003.

OLIVEIRA, Djalma de Pinho Rebouças de. **Sistemas de Informações Gerenciais: Estratégica - Táticas - Operacionais**. 15. ed. São Paulo - SP: Atlas, 2012. 299p.

REZENDE, Denis Alcides. **Sistemas de Informações Organizacionais: Guia prático para projetos em cursos de Administração Contabilidade Informática**. 5.ed. São Paulo - SP: Atlas, 2013. 143p.

STAIR, Ralph M.. **Princípios de Sistemas de Informação: uma abordagem gerencial**. 1. ed. Rio de Janeiro - RJ Ltc, 1998. 452p.

TRACE, Ciaran B. What is recorded is never simply 'what happened': record keeping in modern organizational culture. **Archival Science**, v.2, n.1, p.137-159, 2002. Disponível em: <<https://link.springer.com/content/pdf/10.1007%2F02435634.pdf>>; Acesso em: 10 jul. 2017.