

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

### GT-2 – Organização e Representação do Conhecimento

#### A ARQUITETURA E A RECUPERAÇÃO DA INFORMAÇÃO PARA A MELHORIA DO ÍNDICE H NA PLATAFORMA LATTES

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

Renato Rocha Souza - Universidade Federal de Minas Gerais (UFMG) e Fundação Getúlio Vargas (FGV)

#### THE ARCHITECTURE AND RECOVERY OF INFORMATION FOR THE IMPROVEMENT OF INDEX H IN THE LATTES PLATFORM

#### Modalidade da Apresentação: Pôster

**Resumo:** A plataforma de Currículo *Lattes* é adotada pela maioria das instituições de pesquisa, fomento e universidades do Brasil. Um dos problemas que interfere no índice h de um pesquisador é a quantidade de nomes de citação que um autor pode possuir. Esses novos nomes de citação podem ser gerados mediante o cadastro de metadados de um *Digital Object Identifier* (DOI) realizado por usuários administradores de periódicos. Partindo desse princípio, de que forma seria possível minimizar a inserção de novas citações de um mesmo autor junto a plataforma *Lattes* mediante a erros cometidos por terceiros ao se realizar o preenchimento dos metadados durante o cadastramento de um DOI? Busca como objetivo geral identificar possíveis soluções para uma melhor arquitetura da informação da plataforma *Lattes* como forma de minimizar a possibilidade de inserção de novas formas de citação para um determinado autor mediante a erros de cadastros de metadados de um DOI realizado por administradores de periódicos. Foram utilizados livros, dissertações e artigos científicos disponibilizados em bases de dados para construção da pesquisa. Em uma das discussões, sugere-se como solução a criação por parte da plataforma *Lattes* de um endereço único contendo os metadados básicos como nome, sobrenome e nome de citação que possa ser recuperado em um ambiente de cadastro do DOI, minimizando assim, a possibilidade de erros e transferindo a responsabilidade para o autor.

**Palavras-Chave:** Arquitetura da Informação; Ciência da Informação; Metadados; Recuperação da Informação.

**Abstract:** The Curriculum *Lattes* platform is adopted by most Brazilian research, development and universities institutions. One of the problems that interferes with a researcher's index is the number of citation names an author can possess. These new citation names can be generated through the metadata register of a *Digital Object Identifier* (DOI) performed by journal administrator users. Based on this principle, in what way could it be possible to minimize the insertion of new quotations from the same author on the *Lattes* platform by means of errors made by third parties when completing the

metadata during the registration of a DOI? It seeks as a general objective to identify possible solutions for a better information architecture of the Lattes platform as a way of minimizing the possibility of insertion of new citation forms for a certain author through the errors of metadata registrations of a DOI performed by journal administrators. We used books, dissertations and scientific articles available in databases for the construction of the research. In one of the discussions, it is suggested as a solution the creation by the Lattes platform of a unique address containing the basic metadata such as name, surname and citation name that can be retrieved in a DOI registration environment, thus minimizing the possibility and transferring responsibility to the author.

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

## 1 INTRODUÇÃO

A plataforma de Currículo *Lattes*<sup>1</sup> é o resultado de um estudo do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e adotada pela maioria das instituições de pesquisa, fomento e universidades do Brasil. Permite o registro da vida progressiva e atual de pesquisadores e estudantes do país em um único ambiente de Sistema de Informação.

O *Digital Object Identifier*<sup>2</sup> (DOI) foi desenvolvido pela Associação de Publicadores Americanos com o objetivo de localizar e acessar materiais disponibilizados na web como as publicações de periódicos, livros e documentos em geral. O DOI representa um sistema numérico para conteúdos digitais no qual pode ser inserido na plataforma *Lattes* para recuperação da informação e preenchimento automatizado dos dados publicados.

O abastecimento do DOI de um periódico é realizado por um usuário administrador que preenche de forma manual os metadados referente ao objeto. A estrutura de um metadado apresenta os elementos necessários para a descrição de um conteúdo de dados como por exemplo, autor, título, ano de publicação, volume, número, página inicial e final, *link* de hospedagem e quaisquer informações necessárias para a recuperação da informação (SOUZA; CATARINO; SANTOS, 1997).

Problemas de interpretação por parte dos usuários administradores de periódicos podem ocorrer durante o preenchimento dos metadados de um DOI. Um exemplo está nas preposições contidas em nomes de autores, que se forem cadastrados de forma errada, acabará

---

<sup>1</sup> Plataforma Lattes – Sobre a plataforma Lattes. Disponível em: <<http://lattes.cnpq.br/>>; Acesso em: 27/06/2016.

<sup>2</sup> Digital Object Identifier. Disponível em: <<https://www.doi.org/>>; Acesso em: 27/06/2017.

por gerar uma nova forma de citação para esse determinado autor ao inserir os dados na plataforma *Lattes*.

Partindo desse princípio, questiona-se: de que forma seria possível minimizar a inserção de novas citações de um mesmo autor junto a plataforma *Lattes* mediante a erros cometidos por terceiros ao se realizar o preenchimento dos metadados durante o cadastramento de um determinado DOI?

Essa pesquisa tem como objetivo geral identificar possíveis soluções para uma melhor arquitetura da informação da plataforma *Lattes* como forma de minimizar a possibilidade de inserção de novas formas de citação para um determinado autor mediante há erros de cadastros de metadados de um DOI realizado por administradores de periódicos. Além disso, possui como objetivos específicos realizar uma abordagem conceitual sobre os termos de Ciência da Informação, Arquitetura da Informação, Metadados e Recuperação da Informação como aporte teórico para a realização da proposta da pesquisa destacada na seção de desenvolvimento.

Justifica-se a importância dessa pesquisa uma vez que o número de citações de autores gerados erroneamente por usuários administradores de periódicos durante o preenchimento dos metadados interfere diretamente no índice h, que é uma proposta de quantificar a produtividade e o fator de impacto das publicações científicas dos pesquisadores.

## **2 DESENVOLVIMENTO**

O termo Ciência da Informação está embasado em teorias, experimentos e combinações inter-relacionadas como efetividade, comunicação humana, conhecimento, registros do conhecimento, informação, necessidades de informação, uso da informação, contexto social, contexto institucional, contexto individual e tecnologia da informação (SARACEVIC, 1996). O autor conceitua Ciência da Informação como:

[...] é um campo dedicado às questões científicas e à prática profissional voltadas para os problemas da efetiva comunicação do conhecimento e de seus registros entre os seres humanos, no contexto social, institucional ou individual do uso e das necessidades de informação. No tratamento. destas questões são consideradas de particular interesse as vantagens das modernas tecnologias informacionais (SARACEVIC, 1996, p. 47).

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

Le Coadic (1996) destaca que a ciência da informação é um estudo interdisciplinar e explora os campos de conhecimentos das áreas da psicologia, filosofia, sociologia, linguística, matemática, lógica, informática, telecomunicações, economia, direito e política.

Correlacionado aos conceitos de Ciência da Informação, a Arquitetura da Informação pode ser definida como a viabilização do fluxo efetivo da informação que são estabelecidos por meio de desenhos informacionais (MACEDO, 2005). As autoras LLarena, Duarte e Lira (2016) destacam que o conceito relaciona-se a uma nova área do conhecimento:

A Arquitetura da Informação (AI) como uma nova área do conhecimento, que enfatiza no ambiente web a satisfação de necessidades de informação e envolve os processos de investigação, análise, desenho e implementação, requer habilidades e competências cognitivas sistematizadas, a exemplo da combinação de áreas do conhecimento e linguagens diversificadas (LLARENA; DUARTE; LIRA, 2016, p.37).

O conceito da Arquitetura da Informação pode ser classificado destacando três pontos fundamentais, sendo eles: 1 – A organização de itens que compõe um sistema de informação; 2- O desenho organizacional da distribuição da informação de forma que os conteúdos, tais como as funcionalidades, sejam intuitivos para o usuário final; 3 – A classificação de serviços web, como exemplo, sites e intranets. Esse processo de classificação beneficia diretamente na localização e processos gerenciais da informação (MORVILLE; ROSENFELD, 2006).

Mediante aos processos envolvidos na Arquitetura da Informação, o termo metadados significa dados sobre dados conforme aponta Souza, Vendrúsculo e Melo (2000); Takahashi (2000). Os autores ressaltam que trata-se de dados estruturados que descrevem as características de outros dados no formato digital. Os metadados apresentam vantagens utilizadas para auxiliar na identificação, descrição e localização de informações como um “conjunto de elementos que descrevem as informações contidas em um recurso, com o objetivo de possibilitar sua busca e recuperação” (GRÁCIO, 2002, p. 23).

Alves (2005) apresenta três tipos diferentes de padrões de metadados, sendo eles: 1 – formato simples – não estruturados e extraídos de forma automática por robôs; 2 – formato estruturado – baseados em normas e apresenta a quantidade mínima de dados utilizada para identificação, localização e recuperação. Faz necessário um especialista da informação para preencher as descrições dos campos e; 3 – formatos ricos – altamente estruturados, contemplam metadados complexos, formal e detalhada. Também são baseados em normas.

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

Com base na Arquitetura da Informação e o preenchimento dos metadados, em um mundo imerso nas tecnologias, a recuperação da informação, que necessita de técnicas de representação, organização e indexação se apresenta em novos suportes tecnológicos e sistemas de informação, no qual, podem ser associadas diretamente ao domínio e ao contexto (BARACHO, 2016). O termo recuperação da informação foi definido por Calvin (1951, p2:20-32) “[...] é o nome do processo onde um possível usuário de informação pode converter a sua necessidade de informação em uma lista real de citações de documentos armazenados que contenham informações úteis a ele...”.

Um sistema de recuperação de informação possui componentes e funções que resultam a partir da interação entre o usuário e o sistema. Dentre eles, destaca-se os componentes básicos: 1 - aquisição e representação da informação; 2 – identificação e representação do conteúdo do documento e; 3 – especificação da função de comparação baseada em representação que seleciona os documentos com maior importância. Conseqüentemente, a função de um sistema de recuperação de informação deve comparar e retornar os resultados de uma busca, de forma ordenada, levando em consideração o grau de relevância dos documentos (BAEZA-YATES; RIBEIRO-NETO, 1999).

### **3 METODOLOGIA**

A pesquisa se classifica do ponto de vista de sua natureza como básica; do ponto de vista da abordagem do problema como qualitativa; do ponto de vista dos objetivos como explicativa e; do ponto de vista dos procedimentos técnicos como bibliográfica (GIL, 2010).

Para realização da pesquisa, foram utilizadas dissertações, artigos científicos disponibilizados nas bases de dados do Google Acadêmico, SCIELO, CAPES e livros que abordaram os conceitos de Ciência da Informação, Arquitetura da Informação, Metadados e Recuperação da Informação para construção do aporte teórico.

### **4 RESULTADOS E DISCUSSÕES**

Partindo da perspectiva em que os registros das atividades da vida acadêmica de um pesquisador junto a Plataforma *Lattes* podem servir como progressão em sua vida profissional, o DOI tem facilitado no preenchimento dos dados dos currículos, uma vez que, o pesquisador insere uma sequência numérica referente a uma publicação e a plataforma *Lattes* realiza a

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

recuperação das informações contidas nos metadados, tais como: autor, título, periódico, página inicial e final, *link* referentes ao objeto único cadastrado. Posteriormente, o pesquisador pode inserir no currículo o conjunto de dados clicando em confirmar, porém, sem a possibilidade de alterar quaisquer informações caso esteja errada.

Os metadados de um DOI são registrados por usuário(s) administrador(es) de periódico(s) que o faz(em) manualmente por meio de um ambiente autenticado ou por importação de arquivo *eXtensible Markup Language* (XML). Caso os metadados de nome, nome do meio e sobrenome referente ao DOI sejam cadastrados de forma errada, o número de registro que será utilizado para ser inserido junto a Plataforma *Lattes* acompanhará os dados com erros, conseqüentemente, acarretará em um novo nome de citação para o pesquisador. O ambiente de cadastro de um DOI requer muita atenção por parte do usuário administrador, além disso, trata-se de um ambiente pouco intuitivo, sem tratamento de erros e de baixa usabilidade.

O grande problema de vários nomes de citação para um mesmo autor junto a plataforma *Lattes* está relacionado diretamente ao índice h. Com a inserção de novos nomes de citação para um mesmo pesquisador mediante há erros gerados por usuários durante o cadastro do DOI, acaba por interferir diretamente no índice de relevância de um determinado autor. Fatores como a mudança no estado civil também podem contribuir, já que o pesquisador passa ter mais de um nome de citação. Torna-se necessário evidenciar que o processo de alteração de um DOI também é possível, já os trâmites, depende de uma relação interpessoal entre o autor e os editores. Destaca-se também situações de mudança no estado civil que acompanham inserções ou exclusões de sobrenomes.

Uma possibilidade de minimizar a inserção de novas formas de citação de um mesmo autor a partir de erros cometidos por usuários administradores de periódicos que são responsáveis pelos cadastros do DOI está associada diretamente na Arquitetura da Informação e Recuperação da Informação entre a Plataforma *Lattes* e a *CrossRef*<sup>3</sup>, que é uma associação de editores e instituições que publicam documentos na Internet e que necessitam registrar seu conteúdo através de identificadores únicos. Existem algumas discussões que podem ser

---

<sup>3</sup> Crossref - Metadata enables connections. Disponível em: < <https://www.crossref.org/>>; Acesso em: 02 jul. 2017.

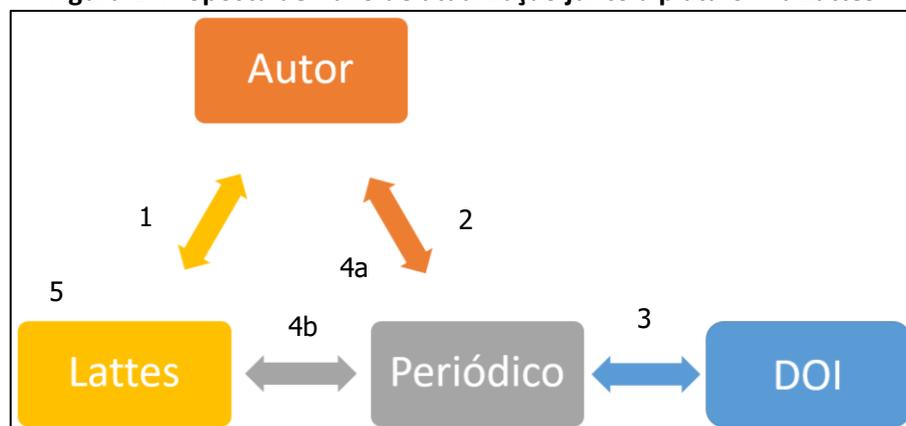
**XVIII ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2017**  
**23 a 27 de outubro de 2017 – Marília – SP**

estudadas para minimizar a quantidade de novas citações que podem ser geradas por erros de usuários, dentre elas:

1. Um cadastro com os dados básicos do pesquisador junto a *Crossref*, incluindo nome de citação. A busca do autor poderia ser realizada por número de documento ou de cadastro;
2. A plataforma *Lattes* criar um endereço único, visível ou não visível para cada autor, no qual, também contemplaria dados básicos incluindo nome de citação; tais dados seriam recuperados por meio de uma *Uniform Resource Locator* (URL) única de cada autor informada durante o preenchimento do DOI, onde apresentariam os metadados básicos.

Pode-se clarificar a segunda discussão conforme apresentado na figura 1, onde: 1 - o autor realiza o preenchimento dos seus metadados como nome, nome do meio, sobrenome e nome de citação que são armazenados junto a plataforma *Lattes* e conseqüentemente é gerado um endereço único com os dados; 2 – o periódico solicita o número de endereço único ao autor; 3 – o periódico realiza a recuperação da informação do autor por meio do endereço único e posteriormente o preenchimento dos metadados do DOI; 4a – o administrador do periódico remete ao usuário o número do DOI com os dados de autoria fornecido pelo próprio autor para posterior atualização junto a plataforma *Lattes*; 4b – o periódico por meio de uma proposta de reformulação da Arquitetura da Informação da plataforma *Lattes* ser tornaria responsável por realizar a atualização da publicação junto a plataforma de currículos sem intermediação do autor; 5 – os dados são atualizados.

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



Fonte: O autor - 2017.

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

A seção discutiu sobre a possibilidade de minimizar erros de preenchimento dos metadados de um documento no que diz respeito aos dados dos autores, como por exemplo nome de citação que interfere diretamente no índice h de cada autor.

## **5 CONSIDERAÇÕES FINAIS**

O uso do DOI para registro de identificação de objeto digital tem aumentado por parte de periódicos de publicação, porém, o processo de cadastro, tem se tornado doloroso no que diz respeito a usabilidade. Trata-se de um ambiente pouco intuitivo e cognitivo, possibilitando assim, erros de cadastramento dos metadados por parte de usuários administradores de periódicos. Um dos erros está no preenchimento do nome dos autores, principalmente, os que possuem preposições. Ao se utilizar o DOI com erros de cadastro, os metadados inseridos junto a plataforma *Lattes* acabará por recuperar informações divergentes e gerar um novo nome de citação para determinado autor, interferindo assim diretamente no índice h do pesquisador.

A problemática e o objetivo geral da pesquisa foram respondidas ao discutir duas soluções que possa minimizar a inserção de novas citações de um mesmo autor junto a plataforma *Lattes*, dentre elas, destaca-se o uso de um *link* único contendo as informações de citação do autor para que seja utilizado por usuários administradores de periódicos ao se realizar o cadastro de um DOI. Dessa forma, além de transportar a responsabilidade dos dados para o autor, evita-se erros de cadastro, já que os metadados seriam carregados mediante a inserção do endereço único no ambiente de cadastro do DOI. O referencial teórico serviu como base para a sugestão de alteração no fluxo, perpassando pelos conceitos de arquitetura da informação, passando pelos metadados e recuperação da informação como forma de contribuição da Ciência da Informação para tratamento dos dados do pesquisador. Sugere-se como pesquisas futuras um aprofundamento na temática, tais como um estudo aprofundado na Arquitetura da Informação da plataforma *Lattes* levando em consideração que a quantidade de formas de citação de um mesmo autor interfere diretamente no Índice h.

## **REFERÊNCIAS**

ALVES, Rachel Cristina Vesú. **Web Semântica: uma análise focada no uso de metadados**. 2005, 180f. Dissertação (Mestrado) Programa de Pós-Graduação em Ciência da Informação, Universidade Estadual Paulista (Unesp), Marília - SP, 2005. Disponível em: <[https://repositorio.unesp.br/bitstream/handle/11449/93690/alves\\_rcv\\_me\\_mar.pdf?sequence=1&isAllowed=y](https://repositorio.unesp.br/bitstream/handle/11449/93690/alves_rcv_me_mar.pdf?sequence=1&isAllowed=y)>; Acesso em: 02 jul.2017.

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

BAEZA-YATES, Ricardo; RIBEIRO-NETO, Berthier. **Modern information retrieval**. New York: ACM press, 1999.

BARACHO, Renata Abrantes. Organização e recuperação da informação pilares da arquitetura da informação. **Tendências da Pesquisa Brasileira em Ciência da Informação**, Brasil, v. 9, n. 1, p. 126–146, 2016. Disponível em: <<http://inseer.ibict.br/ancib/index.php/tpbci/article/view/231/314>>; Acesso em: 1 jul. 2017.

Calvin N. Mooers. **Zatoeodmg applied to mechanical organization of knowledge**. American Documentation, 2:20–32, 1951

GIL, Antonio Carlos. **Como elaborar projetos de pesquisa**. 5. ed. São Paulo - SP: Atlas, 2010.

GRÁCIO, José Carlos Abbud. **Metadados para a Descrição de Recursos da Internet: o padrão Dublin Core, aplicações e a questão da interoperabilidade**. 2002. 104 f. 2002. Dissertação (Mestrado) – Programa de Pós-Graduação em Ciência da Informação, Universidade Estadual Paulista (Unesp), Marília – SP, 2002. Disponível em: <[https://www.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/gracio\\_jca\\_dr\\_mar.pdf](https://www.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/gracio_jca_dr_mar.pdf)>; Acesso em: 03 jul. 2017.

LE COADIC, Yves-François. **A ciência da informação**. Tradução Maria Yêda F. S. de Figueiras Gomes. Brasília - DF: Briquet de Lemos, 1996.

LLARENA, Rosilene Agapito; DUARTE, Emeide Nóbrega; LIRA, Suzana Lucena. A arquitetura da informação à luz da teoria de Piaget: uma possibilidade epistemológica para a gestão do conhecimento. **Perspectivas em Gestão & Conhecimento**, João Pessoa - PB, v. 6, p.36-52, 2016. Semestral. Disponível em: <<http://periodicos.ufpb.br/ojs/index.php/pgc/article/view/23029/15472>>. Acesso em: 29 jun. 2017.

MACEDO, Flávia. Lacerda. **Arquitetura da Informação: aspectos epistemológicos, científicos e práticos**. 186f. Dissertação (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.

SOUZA, Marcia Isabel Fugisawa; VENDRUSCULO, Laurimar Gonçalves; MELO, Geane Cristina. Metadados para a descrição de recursos de informação eletrônica: utilização do padrão Dublin Core. **Ciência da Informação**, Brasília - DF, v. 29, n. 1, 2000. Disponível em: <<http://revista.ibict.br/ciinf/article/view/903/940>>; Acesso em: 02 jul. 2017.

SARACEVIC, Tefko. Ciência da Informação: origem, evolução e relações. **Perspectivas em Ciência da Informação**, Belo Horizonte - MG, v. 1, n. 1, p. 41-62, jan./jun. 1996. Semestral. Disponível em: <<http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/235/22>>. Acesso em: 07 nov. 2016.

SOUZA, Terezinha Batista De; CATARINO, Maria Elizabete; SANTOS, Paulo Cesar Dos. Metadados: catalogando dados na Internet. **Transinformação**, Campinas, v. 9, n. 2, p. 93–105, 1997. Disponível em: <<http://periodicos.puc-campinas.edu.br/seer/index.php/transinfo/article/view/1586/1558>>; Acesso em: 27 jun. 2017.

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

TAKAHASHI, T. (Org.). **Sociedade da Informação no Brasil: o livro verde**. Brasília - DF:

Ministério da Ciência e Tecnologia, 2000. 153p, Disponível em:

<<http://livroaberto.ibict.br/bitstream/1/434/1/Livro%20Verde.pdf> >; Acesso em: 27 jun. 2017.