

**A RELAÇÃO DA NORMA GERAL INTERNACIONAL DE DESCRIÇÃO
ARQUIVÍSTICA [ISAD (G)] COM AS RECOMENDAÇÕES DO WORLD WIDE
WEB CONSORTIUM (W3C): O MODELO RDF**

Diana Vilas Boas Souto Aleixo – UEL
Maria Elisabete Catarino – UEL

Resumo

Explora a aplicação do modelo *Resource Description Framework* (RDF) na descrição arquivística, tendo como foco a Norma Geral Internacional de Descrição de Arquivos [ISAD (G)]. Tem como objetivo investigar a relação da Norma ISAD (G) com as recomendações do *World Wide Web Consortium* (W3C) para a aplicação do modelo RDF na descrição arquivística. Numa abordagem qualitativa está sendo realizada uma pesquisa exploratória a qual terá como resultado um diagrama em *Unified Modelling Language* (UML) que representará a estrutura conceitual da aplicação do RDF, a ser validada pelos especialistas da Câmara Técnica de Normalização da Descrição arquivística. Este trabalho permitirá visualizar a utilização do padrão RDF para a organização das informações conforme as recomendações da norma ISAD (G). Portanto, investigar a relação da Norma ISAD (G) com as recomendações do W3C que orientam a aplicação do modelo RDF se faz necessário, a fim de corroborar com a organização e recuperação das informações disponíveis hoje na *web*, e de modo particular, que orientam o campo de pesquisa voltado a Ciência da Informação e a Arquivística.

Palavras-chave: ISAD (G). Normas de Descrição Arquivística. *Web Semântica*. RDF.

Abstract

Explores the application of the *Resource Description Framework* (RDF) model in archival description, focusing on the *General International Standard File Description* [ISAD (G)]. Aims to investigate the relationship of the standard ISAD (G) with the *World Wide Web Consortium* (W3C) recommendations for implementing the RDF model in archival description. A qualitative approach is being performed an exploratory research which will result in a *Unified Modelling Language* (UML) diagram that represents the conceptual structure of the application of RDF to be validated by experts of the *Technical Standardization of archival description*. This paper will view the use of the standard RDF for the organization of information according to the recommendations of ISAD (G). Therefore, to investigate the relationship of the standard ISAD (G) with the W3C recommendations that guide the implementation of the RDF model is needed in order to corroborate the organization and retrieval of information available today on the web, and in particular, that guide the research field aimed *Information Science* and the *Archival*.

Keywords: ISAD (G). *Standard Archival Description*. *Semantic Web*. RDF.

1 INTRODUÇÃO

As tecnologias de informação no âmbito da Arquivística tem ampliado as possibilidades de novas aplicações, principalmente no que se relaciona aos processos de organização dos documentos voltados na *web*.

Neste sentido, é imprescindível a padronização no processo de descrição documental, tal como previsto na Norma Geral Internacional de Descrição Arquivística¹ [ISAD (G)]. E neste sentido, remete-se a Ramalho (2013), o qual relata que a adoção da *Web Semântica* permite criar meios alternativos para localizar os arquivos na *web*, por meio da criação de índices e da adoção de padrões é possível relacionar os *links* em *Resource Description Framework* (RDF) presentes em diferentes bancos de dados.

Para se entender o contexto, descreve-se que em meados de 2001, Tim Berners-Lee trouxe à tona uma nova visão da *web*, o que ele nominou de “*Web Semântica*”, que é uma extensão da *web* atual, porém constituída por documentos auto-descritíveis, onde o conteúdo pode ser compreendido por programas, o que por sua vez possibilitaria aos agentes de *software* raciocinar e realizar inferências relativas ao conteúdo dos documentos, auxiliando o usuário na recuperação das informações. (MARCONDES; CAMPOS, 2008).

Para tornar essa *web* uma realidade, os dados contidos na rede devem ser representados por um formato padrão, acessível e gerenciável por tecnologias apropriadas. Além do acesso aos dados, é imprescindível, criar relações entre os dados. Neste sentido, é preciso adotar o (RDF) recomendado pelo *World Wide Web Consortium* (W3C), o qual é um modelo padrão para intercâmbio de dados na *web* (RDF WORKING GROUP, 2004).

Atentos a estes avanços, os conselhos que orientam e normalizam a área de arquivos iniciam um caminho aberto a adoção de metadados a partir das recomendações estabelecidas na norma ISAD (G), como *Encoded Archival Description Document Type Definition*² (DTD EAD) e o *Europeana Data Model*³ (EDM), os quais fazem uso do *eXtensible Markup Language* (XML).

O objetivo central deste trabalho constitui em investigar a relação da Norma ISAD (G) com as recomendações do W3C para a aplicação do modelo RDF. Para atingir o objetivo

¹ Disponível em: <http://www.conarq.arquivonacional.gov.br/Media/publicacoes/isad_g_2001.pdf>. Acesso em: 20 jun. 2013.

² Desenvolvido pela Biblioteca da Universidade da Califórnia e amparado pela Biblioteca do Congresso Americano e pela *Society of American Archivists* (SAA). Disponível em: <<http://www.loc.gov/ead/>>. Acesso em: 20 jun. 2013.

³ Criado pela *The Foundation Europeana*. Disponível em: <<http://pro.europeana.eu/edm-documentation>>. Acesso em: 20 jun. 2013.

proposto pretende-se: contextualizar as recomendações elaboradas pelo W3C e as Normas de descrição arquivística; identificar e analisar as correlações entre os campos de atributos e os campos de descrição existentes na norma ISAD (G) com o modelo RDF de forma mútua; apresentar por meio de um diagrama em *Unified Modelling Language* (UML) o desenho conceitual de relacionamento entre documentos descritos em RDF; e caracterizar as opiniões dos integrantes da Câmara Técnica de Normalização da Descrição Arquivística (CTNDA) do CONARQ frente a aplicação do modelo RDF junto a norma ISAD (G).

Acredita-se que a maior contribuição deste trabalho está em levar a Arquivística e a Ciência da Informação (CI) para este círculo de debate, propiciando subsídios e estabelecendo pontes com a *Web Semântica*, além de possibilitar o início de novas pesquisas referentes as temáticas abordadas e corroborar com a organização e recuperação das informações disponíveis hoje na *web*.

2 A DESCRIÇÃO ARQUIVÍSTICA

Para garantir o acesso às informações contidas nos documentos, a Arquivística estabelece seus próprios critérios elaborados a partir de uma lógica e de metodologias específicas a cada tipo documental, a fim de auxiliar à pesquisa e ao conhecimento. (CARBONE, 1993).

O Conselho Internacional de Arquivos (CIA, 2000) define a descrição arquivística como a representação de uma unidade⁴ de descrição e de suas partes componentes, por meio da extração, análise, organização e registro de informação, a fim de auxiliar na identificação e localização dos documentos de arquivo, além de elucidar o contexto de sua produção e acumulação.

Desde a publicação do Manual Holandês (ASSOCIAÇÃO DOS ARQUIVISTAS HOLANDESES, 1960, uma grande mudança aconteceu, surgiram várias normas e publicações a fim de auxiliar os profissionais e os usuários de arquivo.

A ISAD(G), norma eleita como foco neste trabalho, determina regras para descrição arquivística que podem ser aplicadas em qualquer documento independente de forma ou suporte em que se encontre o documento. Essas regras consistem em assegurar a criação de descrições consistentes, facilitarem a recuperação e a troca de informação sobre documentos arquivísticos, possibilitar o compartilhamento de dados de autoridades e tornar possível a

⁴ “Documento ou conjunto de documentos, sob qualquer forma física, tratado como uma unidade, e que, como tal, serve de base a uma descrição particularizada [...]”. (CIA, 2000, p. 16).

integração de descrições de diferentes arquivos em um sistema unificado de informação (CIA, 2000).

O intuito é desenvolver uma discussão em torno da especificidade da descrição, e sua relação com a *Web Semântica*. Para isso, será discorrido de forma breve, conceitos relativos a temática.

3 A WEB SEMÂNTICA

O termo '*Web Semântica*' ou como Tim Berners-Lee denomina *web* de dados, é um nome dado ao projeto do W3C a qual refere-se à visão da *web* dos 'Dados Lincados'⁵. Conforme o consórcio W3C (2013) a *Web Semântica* dá às pessoas a capacidade de criarem repositórios de dados na Web, de forma a facilitar a troca de informação. Conforme Catarino (2009) o termo foi dito pelos autores Tim Berners-Lee, James Hendler e Ora Lassila em 2001 quando divulgaram na revista *Scientific American* o estudo que relacionava o desenvolvimento de tecnologias que possibilitassem a informação legível pelas máquinas.

Esta proposta, na visão de Breitman (2005) parte da necessidade de organizar a informação na internet. A ideia de categorizar a informação de forma padronizada como a autora relata, pode ser comparada a classificação dos seres vivos. Conforme Berners-Lee, Hendler e Lassila (2001) a *Web Semântica* não é uma *web* separada, na verdade ela é uma extensão da *web* atual, onde a informação possui significados definidos, o que permite uma cooperação entre computadores e pessoas.

As autoras Catarino e Baptista (2008) descrevem que o princípio que norteia a *Web Semântica* é simples, e requer a utilização de linguagens e formatos comuns para a descrição. Nesse sentido, a *web* semântica fornece um enquadramento comum que permite o compartilhamento dos dados. O compartilhamento e armazenamento são realizados através da manipulação de dados, dados estes habilitados por tecnologias como o *Resource Description Framework* (RDF), a *Web Ontology Language* (OWL) e *Simple Knowledge Organization System* (SKOS). (W3C, 2013).

A seguir será explanada a metodologia elencada e em andamento.

⁵ "Dados lincados é o nome que se dá às coleções de dados relacionados na *Web*." (CATARINO; SOUZA, 2012, p. 79).

4 PROCEDIMENTOS METODOLÓGICOS

O presente trabalho está sendo realizado numa abordagem qualitativa por meio de pesquisa exploratória. Amparada na análise dos documentos, os quais fazem parte as normas de descrição arquivística e as recomendações para aplicação do modelo RDF elencadas pelo consórcio W3C, segue a elaboração de um diagrama conceitual em *Unified Modelling Language* (UML), o qual representará a aplicabilidade do padrão RDF na norma ISAD (G).

De posse do diagrama, o trabalho partirá para a fase de validação, onde será apresentado o resultado da análise aos especialistas da área de descrição arquivística, aqui nomeados os integrantes da Câmara Técnica de Normalização da Descrição Arquivística (CTNDA) do Conselho Nacional de Arquivos (CONARQ).

A consulta será realizada por meio da técnica Delphos, onde o estudo será conduzido por meio de questões direcionadas a temática (amparadas pelo diagrama) ao grupo de especialistas em rodadas, as quais permitirão que a cada devolução dos questionários seja possível visualizar um quadro de posicionamentos, até a convergência de todas as opiniões do grupo quanto a aplicação do padrão RDF na norma ISAD (G).

5 CONTRIBUIÇÕES ESPERADAS

O resultado obtido após a análise e interpretação dos dados, permitirá visualizar a utilização do padrão RDF para a organização das informações conforme as recomendações da norma ISAD (G).

O presente trabalho está em andamento, porém é possível ressaltar a ausência de estudos teóricos, o que elucida a necessidade de um maior aprofundamento teórico e metodológico das temáticas abordadas.

Atualmente, em andamento o levantamento bibliográfico e documental, relacionado a *Web Semântica* e as normas de descrição arquivística, e a elaboração do diagrama em UML.

REFERÊNCIAS

- ASSOCIAÇÃO DOS ARQUIVISTAS HOLANDESES. **Manual de arranjo e descrição de arquivos**. Tradução de Manuel Adolfo Wanderley. Rio de Janeiro: Arquivo Nacional. 1960.
- BERNERS-LEE, T.; HENDER, J.; LASSILA, O. The semantic web: a new form of Web content that is meaningful to computers will unleash a revolution of new possibilities. **Scientific American**, New York, may. 2001. Disponível em: <<http://www.med.nyu.edu/research/pdf/mainim01-1484312.pdf>>. Acesso em: 14 maio de 2013.
- BREITMAN, K. K. **Web semântica: a internet do futuro**. Rio de Janeiro: LTC, 2005.

CARBONE, S. História e arquivística. **Revista de Biblioteconomia**. Brasília, DF. v.11, n.1, p. 45-53, jan-jun.1993. Disponível em: <<http://www.brapci.ufpr.br/documento.php?dd0=0000001594&dd1=adef2>>. Acesso em: 17 jan. 2013.

CATARINO, M. E. **Integração das folksonomias nos metadados**: identificação de novos elementos como contributo para a descrição de recursos em repositórios. 2009. Tese (Doutorado em Tecnologias e Sistemas de Informação) – Departamento de Sistemas de Informação, Universidade do Minho, 2009.

CATARINO, M. E.; BAPTISTA, A. A. *Web semântica e a qualidade no intercâmbio da informação*. In: TOMAÉL, M. I. **Fontes de informação na internet**. Londrina: EDUEL, 2008, p. 31-51.

CATARINO, M. E.; SOUZA, T. B. A representação descritiva no contexto da *web* semântica. **TransInformação**, Campinas, v. 24, n. 2, p. 77-90, maio/ago., 2012. Disponível em: <<http://revistas.puc-campinas.edu.br/transinfo/viewissue.php?id=33>>. Acesso em: 01 jul. 2013.

CONSELHO INTERNACIONAL DE ARQUIVOS (CIA). ISAD (G): **Norma Geral Internacional de Descrição Arquivística**. Rio de Janeiro: Arquivo Nacional, 2000.

MARCONDES, C. H.; CAMPOS, M. L. A. Ontologia e *Web Semântica*: o espaço da pesquisa em ciência da informação. **Ponto de Acesso**, Salvador, v. 2, n. 1, p. 107 – 136, 2008. Disponível em: <<http://www.portalseer.ufba.br/index.php/revistaici/article/view/2669/1885>>. Acesso em: 01 jul. 2013.

RAMALHO, J.C. Arquivos digitais na 3ª geração da *Web*: uma ideia com mais de 20 anos. Conferência apresentada no **VII Encontro CTDI**, Vila do Conde, 2013. Disponível em: <<http://hdl.handle.net/1822/23979>>. Acesso em: 28 jul. 2013.

RDF WORKING GROUP. **Resource description framework**. Massachusetts: W3C, 2004. Disponível em: <<http://www.w3.org/RDF/>>. Acesso em: 01 jul. 2013.

WORLD WIDE WEB CONSORTIUM (W3C). **Web Semântica**. Disponível em: <<http://www.w3c.br/Padroes/WebSemantica>>. Acesso em: 17 mar. 2013.