

Análise dos padrões XML e RDF para a representação na web sob a perspectiva da Ciência da Informação: um estudo preliminar

Janailton Lopes Sousa

Universidade Federal de São Carlos – UFSCar, Email: janailtonlopes20@gmail.com

Paulo George Miranda Martins

Universidade Federal de São Carlos – UFSCar, Email: pgeorgemm22@gmail.com

Rogério Aparecido Sá Ramalho

Universidade Federal de São Carlos – UFSCar, Email: rogerio.ramalho@gmail.com

Resumo

A partir do desenvolvimento das tecnologias de informação novas possibilidades de representação foram incorporadas ao ambiente Web, entre os principais padrões desenvolvidos nas últimas décadas e que têm influenciado as práticas profissionais na área de Ciência da Informação pode-se destacar a linguagem *eXtensible Markup Language (XML)* e o *Resource Description Framework (RDF)*. Ao longo dos últimos anos, tais tecnologias têm revolucionado a forma como manipulamos recursos informacionais, contudo, é possível observar que muitos profissionais da área de Ciência da Informação ainda possuem dificuldades em relação ao entendimento e utilização destes padrões. Neste sentido, este artigo apresenta uma pesquisa em andamento com o objetivo de descrever as principais características dos padrões XML e RDF, analisando os aspectos conceituais que fundamentam tais tecnologias sob a perspectiva da área de Ciência da Informação. Trata-se de um estudo teórico e exploratório, realizado mediante levantamento bibliográfico sobre a temática em questão, revisitando a literatura da área de Ciência da Informação e documentos técnicos publicados pelo W3C relacionados aos padrões analisados. Como resultados são apresentados sucintamente algumas das relações existentes entre os fundamentos que norteiam o desenvolvimento de tais padrões e conceitos da área de Ciência da Informação, no intuito de favorecer uma maior discussão desta temática a contribuir para o aprimoramento profissional em relação às novas tecnologias de representação.

Palavras-chave: Representação da Informação. Linguagens de marcação. Tecnologias Semânticas. XML. RDF.

1 INTRODUÇÃO

A web é composta por vários documentos que muitas vezes não apresentam formalmente estruturação lógica que favoreçam conexões semânticas. Os desafios para organização da informação na web não refletem apenas anseios de um campo científico como a Computação, mas o esforço de vários campos que podem colaborar na sistematização dessa ampla gama de informações. O *Word Wide Web Consortium (W3C)* é o órgão regulamentador de boas práticas para publicação de informações e dados no ambiente digital. A adoção de

padrões reconhecidos pela W3C implica na acessibilidade e interoperabilidade de diversos formatos na web.

Ao efetuar buscas na Web é possível identificar grandes volumes de informações e dados produzidos por seres humanos e processados por máquinas, tal sobrecarga impede a leitura de muitos desses textos. *Browsers* indexam páginas web diariamente e para isso necessitam identificar descritores, palavras-chave, *tags*, termos correlatos ou sentenças, para gerar um resultado de busca. Isto parece muito comum no campo da Biblioteconomia e Ciência da Informação, que lida com indexação e Sistemas de Recuperação de Informações (SRI). Seja de forma automática ou manual criando formas de representação de documentos por meio referências, descrição bibliográfica, indexação, resumos, cabeçalhos de assunto, vocabulários controlados, sumários, índices catálogos entre outros, formas estas, que sempre remetem ao documento original.

O objetivo deste trabalho, em andamento, pauta-se na discussão conceitual e alguns aspectos usuais dos formatos XML e RDF. Apresentação de suas características principais, além de fomentar as discussões sobre padrões de representação de informações utilizados na web sob a ótica da Ciência da Informação. Trata-se de um estudo teórico e exploratório feito por meio de pesquisa bibliográfica e da literatura existente sobre a temática em questão revisitando alguns conceitos da Ciência da Informação e os documentos técnicos publicados pelo W3C sobre a usabilidade dos respectivos padrões.

Tim Berners-Lee criador da *World Wide Web* ao publicar o artigo *The Semantic Web* apresenta uma proposta de extensão da web atual. Dando um significado bem definido à informação, habilitando computadores e pessoas a trabalharem em cooperação (BERNERS-LEE; HENDLER; LASSILA, 2001). O grande desafio permeia na recuperação de informações relevantes em meio a milhões de bytes em conteúdo produzido por humanos e para humanos. Exige o processamento semântico desse conteúdo, para retornar respostas compatíveis com o tipo de busca efetuada, segundo Goñi; Fernandes; Lucena (2002, p.4):

Na Web Semântica, um conceito importante estará associado à uma URI, (*Uniform Resource Identifier*), cuja principal função é assegurar que conceitos não são apenas palavras num documento, mas estão amarrados a um único significado, que qualquer pessoa poderá encontrar na Web.

A web semântica, apesar de amplamente discutida na área tecnológica não sofreu avanços tão imediatos como se espera deste setor. A grande dificuldade sempre pairou sobre a conversão do que se havia produzido até então, para formatos semânticos, pois simplesmente criar uma web semântica descartando milhares de *exabytes* da web sintática não parece uma proposta favorável para um mercado que se consolidou nesse ambiente. Depois de algum tempo a proposta *Linked Data* surge como uma grande alternativa para otimizar esse volume de dados, o grande segredo estava em descrever para máquinas do que se trata cada documento na web.

Essa análise, que aparentemente parece uma regressão, foi necessária para avaliar o *status quo* da área e propor alternativas com tecnologias existentes, pesquisando os fundamentos da própria web semântica percebeu-se que ao descrever um documento é possível atribuir uma URI que apontaria um significado único para um documento, como um conceito, conforme supracitado em (GOÑI; FERNANDES; LUCENA, 2002).

Neste sentido uma das tecnologias amplamente utilizadas para descrever coisas no ambiente virtual é o XML - *eXtensible Markup Language*, uma linguagem de marcação que permite criar outras linguagens com as propriedades do XML, como o formato OWL, RDF - *Resource Description Framework* e SPARQL. Estes formatos tornam a semântica possível à medida que estruturam em classes, subclasses, axiomas e os conceitos que compõem um determinado documento.

Pickler (2007) faz uma distinção entre web semântica e web sintática a primeira busca significados e relações entre si para compreensão homem/máquina já a segunda foi elaborada propriamente para humanos compreenderem, pois os significados do que é apresentado fica a cargo de quem efetuou a busca para saber se o resultado é relevante ou não.

Os *links* que conectam uma página a outra na web sintática não são estruturados semanticamente ficando a cargo do ser humano consultá-los para saber se atendem a necessidade de uma busca. O que normalmente ocasionam a abertura de diversas abas durante a navegação, implicando numa análise demorada do conteúdo de cada página. Apesar de existir páginas colaborativas como a Wikipédia que por meio de *hyperlinks* concatenam vários assuntos marcados ao longo de um artigo publicado. Isto envolve a participação de muitos usuários o que implicaria numa ampla ação colaborativa.

A Web atual consiste de recursos e links, porém estes links são criados apenas para o uso de seres humanos, de modo que é relativamente simples para um ser humano identificar se um link, contido em um determinado recurso, referência uma fatura, um romance ou um trabalho científico, porém tais informações não estão acessíveis para as máquinas, pois os links na Web atual não indicam quais são os tipos de relações existentes entre os recursos referenciados. (RAMALHO, 2006, p. 39).

O W3C conceitua a web semântica como uma web de dados formalmente interligados, composta por uma coleção de tecnologias web, que fornecem um ambiente onde a aplicação pode consultar esses dados e realizar inferências usando vocabulários (W3C, 2015.). Este estudo exploratório e discursivo centra-se apenas em dois formatos de representação, neste caso o XML e o RDF que são elementos chave a serem estudados.

2 REPRESENTAÇÃO NA CIÊNCIA DA INFORMAÇÃO

Tornar eficiente a prática do profissional da informação de forma reflexiva e relevante tanto para quem disponibiliza quanto para quem consome informação é um desafio a ser enfrentado. Explorar tecnologias é um passo simples e muitas vezes intuitivo. A questão centra-se na avaliação do que está sendo produzido e como isto altera as formas de processamento da informação, tanto por humanos, quanto por máquinas. Existem diversas formas de avaliação da subjetividade humana, quicá de máquinas e softwares programáveis. Portanto torna-se necessário a avaliação destes insumos digitais e aprofundamento das discussões sobre as novas formas de se recuperar informações no ambiente web, seja no modo sintático ou semântico.

O contexto que situa as discussões na Ciência da Informação sempre foi marcado pela influência tecnológica. Todavia, o uso de técnicas que subsidiam as práticas do profissional da informação é destacado pela consistência e perenidade, que os modelos tradicionais de representação da informação fixaram nas unidades de informação. O desafio que desencadeia o cerne da prática profissional é impulsionado pela dinâmica das transformações sociais influenciadas pelo uso de tecnologias digitais.

Habilitar as reformulações que caracterizam o avanço da ciência não indica o menosprezo aos fundamentos básicos, entretanto um aprofundamento epistemológico, que expõe tal ciência a um posicionamento em relação aos desafios emergentes de sua área. A

Ciência da Informação, como área que nasce de forma interdisciplinar em uma reunião do *Georgia Institute of Technology* em 1962 (BRAGA, 1995). Tem contato com diversas áreas, como Biblioteconomia, Ciência da Computação, Comunicação, Linguística, etc.

Para Pinheiro; Loureiro (1995) na década de 60 são elaborados os primeiros conceitos, definições e se inicia o debate sobre as origens e os fundamentos teóricos na nova área, no qual se identifica uma tentativa de melhor demarcá-la e estabelecer relações interdisciplinares com outros campos. No entanto sua sedimentação teórica conseguiu se efetivar rapidamente, não abrindo espaço para conflitos de identidades que definiriam esta área. E sim um aproveitamento de disciplinas, cujo objeto de estudo é a informação. De acordo com Braga (1995, p. 4):

Os relatos da literatura de língua inglesa da década de 60 e mesmo boa parte da década de 70 mostram uma ciência da informação atrelada à teoria matemática da comunicação; empolgada com a nascente automação de sistemas de recuperação da informação e de bases de dados; voltada para os problemas de semântica, visando à representação da informação[...]

Problemas semânticos envolvendo a representação da informação não é algo recente para esta ciência. Pois a informação comporta um elemento de sentido (Le Coadic, 1996). Portanto as relações linguísticas que orientam os arranjos representativos da informação são estabelecidos por um sentido lógico e coeso da língua. A utilização de propriedades linguísticas formais é uma ação relativamente comum para concepção de termos representativos da informação.

De acordo com Novellino (1996) a principal característica do processo de representação da informação é a substituição de uma entidade linguística longa e complexa - o texto do documento - por sua descrição abreviada. Esta compressão textual, objetiva a extração ou atribuição de termos representativos da informação é uma das ações principais desta área, com o uso de linguagens documentárias para o controle terminológico de termos, que tornou-se um instrumento basilar do profissional da informação, para a construção de tesouros, vocabulários controlados, cabeçalhos de assuntos, entre outros.

Neste sentido, o ambiente digital não exige apenas um código ou linguagem específica, mas padrões interoperáveis que possam descrever, representar e criar relações entre si. Além disso, formatos como o XML proporcionam a criação de outras linguagens, como

RDF. A Ciência da Informação oferece os preceitos teóricos e práticos para o delineamento das dinâmicas e fluxos de informação, o estudo destes modelos tecnológicos não é uma tarefa fácil, no entanto é possível gerar aproximações teóricas que permitam um melhor aproveitamento deste campo de estudo.

3 XML COMO FORMATO DE REPRESENTAÇÃO

Representar conceitos e coisas é um dos principais desafios para construir uma semântica legível para máquinas diante de inúmeros documentos, dados e informações contidas na web. O *Extensible Markup Language* (XML) é um importante formato de representação que possibilita a utilização de tags personalizadas, pois se trata de uma linguagem de marcação extremamente versátil e amplamente utilizada em softwares e aplicações web.

O W3C gerencia pacotes de recomendações para padronizar formatos e práticas na disponibilização de recursos e aplicações da web. Isto facilita sobretudo a interoperabilidade entre softwares e aplicações web para a produção e consumo de informações. Dentre essas diversas recomendações, o XML é um formato de representação de qualquer tipo de documento, como o nome propõe é uma linguagem de marcação extensível. O XML é derivado de uma linguagem desenvolvida pela IBM na década de 1980, a *General Markup Language* (GML), (Moreno; Brascher, 2007).

O XML é uma versão simplificada da linguagem GML, menos complexa, no entanto extremamente exigente na formatação e validação dos documentos, uma linguagem que não tolera erros em seu escopo. Por ser um formato simples, mas de alto nível qualitativo é amplamente utilizado em soluções web, banco de dados, software de processamento de informações, bibliotecas virtuais e quaisquer mecanismos que necessitam de troca de informações. A disponibilização de informações bibliográficas no formato XML é uma realidade em bibliotecas como a *Library of Congress* (LC), criadora do formato *Machine Readable Cataloging* (MARC) desenvolvido na década de 1960 com o objetivo de viabilizar a comunicação de descrições bibliográficas em formato legível por computador (MORENO, 2006).

Com o avanço intermitente da tecnologia, fazer uma máquina apenas ler arquivos de descrição bibliográfica não era mais viável, pois os softwares de gerenciamento de

bibliotecas utilizavam cada vez mais o ambiente web para viabilização de seus produtos e serviços. E o MARC não era apropriado para este tipo de ambiente, devido a sua complexidade. Assim, a incorporação do XML em padrões de descrição bibliográfica surgiu como importante solução, tendo em vista que este serviria para descrever qualquer tipo de artefato. O MARCXML é uma iniciativa de adaptar o formato MARC para um formato nativo da web, pois uma linguagem independente e própria possui um alto nível de resistência em soluções alçadas em interoperabilidade de formatos e sistemas.

A LC reconheceu que a incorporação do XML no formato MARC foi necessária para possibilitar a interoperabilidade que o ambiente web precisa para a composição de rede e fluxo de informações. O XML é recomendado pela W3C desde 1998 e como síntese da SGML é reconhecido pela *International Organization for Standardization* (ISO) desde 1986 (BRAY et al, 2006). Apesar de ser uma versão simplificada, o XML possui o mesmo rigor do SGML identificadas pela ISO, o que garante a confiabilidade de um recurso que usa este tipo de tecnologia.

De acordo com Lima; Carvalho (2005) o XML é uma metalinguagem, ou seja, ela oferece recursos para a definição de gramáticas que caracterizam linguagens para classes de documentos específicos, com conjunto de elementos, atributos e regras de composição bem determinados. Isto possibilita descrever qualquer tipo de artefato material ou virtual, diferente do *HyperText Markup Language* (HTML) ele não possui marcações pré-definidas, o que aumenta sua possibilidade de uso. Para Moreno (2006) um dos problemas relacionados a HTML é a falta de padronização, que prejudica a interoperabilidade, diferente do XML, que é compatível com qualquer navegador e não admite erros em sua estruturação.

O XML permite adicionar qualquer tipo de marcação, ou seja, os campos de descrição podem ser expandidos para, ano, editora, local, localização, etc. Outra característica do XML são um conjunto de regras expressas em *Document Type Definition* (DTD) e *XML Schema*. Entendidas como parâmetros que validam ou confirma se a estrutura do arquivo XML está correta. Para Siqueira (2003) um arquivo DTD é usado para especificar quais elementos, *tags* ou atributos são permitidos no arquivo XML. Segundo Lima; Carvalho (2005) um documento criado de acordo com uma sintaxe (DTD ou *XML Schema*) é um documento XML

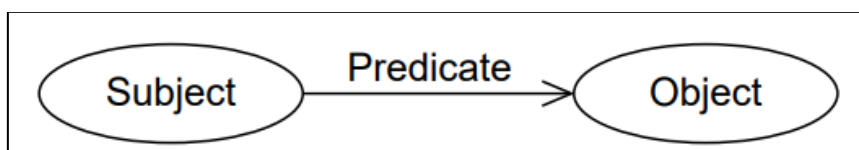
válido. Isto não é uma obrigatoriedade para arquivo.xml, todavia ressalta a qualidade ou expansão dos dados que o compõem.

O uso do XML em diversas instituições públicas e privadas já se tornou um realidade. As notas fiscais eletrônicas são um grande exemplo utilizado pelo poder público e setor privado, beneficiando ambos em compras, relatórios, transações financeiras, entre outras atividades. Sistemas de Gerenciamento de Banco de Dados (SGBD) precisam se comunicar com outros sistemas de informações, que podem operar apenas no ambiente web, como é o caso das bibliotecas virtuais necessitam de mecanismos interoperáveis entre esses dois sistemas, para importar, exportar, gerar ou ler arquivos, com o XML isto possível e ampliável dependendo da necessidade dos envolvidos, pois é possível adaptá-lo sem perder a qualidade que o XML proporciona.

4 O RDF COMO PADRÃO DE REPRESENTAÇÃO DE RECURSOS

Dentre os diversos padrões e modelos possíveis a partir do XML, o *Resource Description Framework* – RDF se destaca como forma de representação de informações no ambiente web. De acordo com Furgeri (2006) o RDF tem como objetivo principal prover intercâmbio de informações entre aplicações sem a perda do significado. Portanto preocupa-se com a definição de um vocabulário padrão de representação que possa permitir a interoperabilidade entre aplicações.

Figura 1 – Gráfico RDF simples



Fonte: CYGANIAK; WOOD; LANTHALER (2014)

O RDF funciona por meio de declarativas que expressam a relação tripla entre recurso, que representa o sujeito de uma declaração, propriedade, que é o predicado de uma declaração e valor, o objeto de uma declaração (FURGERI, (2006)). Esta este modelo relacional fundamenta-se nas propriedades linguísticas que expressão, sujeito, predicado e objeto nas orações. Isto expõe as relações semânticas entre as declarativas que serão interpretadas por uma

máquina. O RDF, amplamente discutida sob a ótica da web semântica compõe a camada dura expressa por Tim Berners-Lee, logo, a web semântica se beneficia das ontologias à medida que efetua buscas específicas em determinados domínios, os quais possuem relações formais entre os conceitos.

Apesar desta discussão que envolve Web semântica, ontologias, *Linked Data*, *Big Data*, etc, estarem no auge das conferências, palestras, treinamentos e atividades voltadas ao profissional da informação. Este trabalho volta-se para o elemento comum dessas nuances, que é o RDF, diante dos avanços e usos que o envolvem, a perspectiva da Ciência da Informação precisa atentar para esta forma de representação da informação na web, que atualmente é explorada por diversas instituições, como a própria LC.

De acordo com Ferreira; Santos (2013) o RDF pode utilizar-se do XML como sintaxe comum para o intercâmbio e o processamento de metadados, portanto a qualidade desse modelo para interoperabilidade de recursos no ambiente web está em consonância com a acuidade do SGML. Por isso é um modelo reconhecido pela W3C, o RDF estabelece na verdade um padrão de metadados para ser embutido na codificação XML (SOUZA; ALVARENGA, 2004). Neste sentido o RDF oferece subsídios para a construção semântica da web por meio de dados interligados que podem ser implementados em sistemas de recuperação de informações. De acordo com Martins; Ramalho (2017) atualmente a principal tecnologia que fundamenta a representação semântica na Web é denominada como RDF.

O uso do RDF permite utilizar as propriedades do XML e atribuir relacionamentos entre elas. Tudo isso por meio da sintaxe RDF e validações de esquemas que comprovem a correta utilização dos prefixos de marcação utilizados. Uma das principais características do RDF é a possibilidades de relacionamentos entre recursos. Utilizando não somente a tripla sujeito predicado e objeto, mas o uso expansivo de classes e propriedades que tornam efetivo esse relacionamento conceitual.

As classes e propriedades do RDF podem ser comparadas a conjuntos e suas características, no qual é possível relacionar com outros conjuntos, expressando relações hierárquicas ou axiomáticas. Como toda linguagem entender a sintaxe é saber as regras que condicionam o uso da língua. No ambiente virtual isso não é diferente, pois trata-se de uma preditiva para leitura dos arquivos em RDF e correta apropriação do assunto. A figura 2

Análise dos padrões XML e RDF para a representação na web sob a perspectiva da Ciência da Informação: um estudo preliminar

apresenta um exemplo elaborado a partir dos modelos disponíveis da LC, um arquivo de texto simples expresso em RDF, no qual é apresentado um termo “Designer de som” e uma nota de descrição.

Figura 2 – Arquivo com descrições em RDF

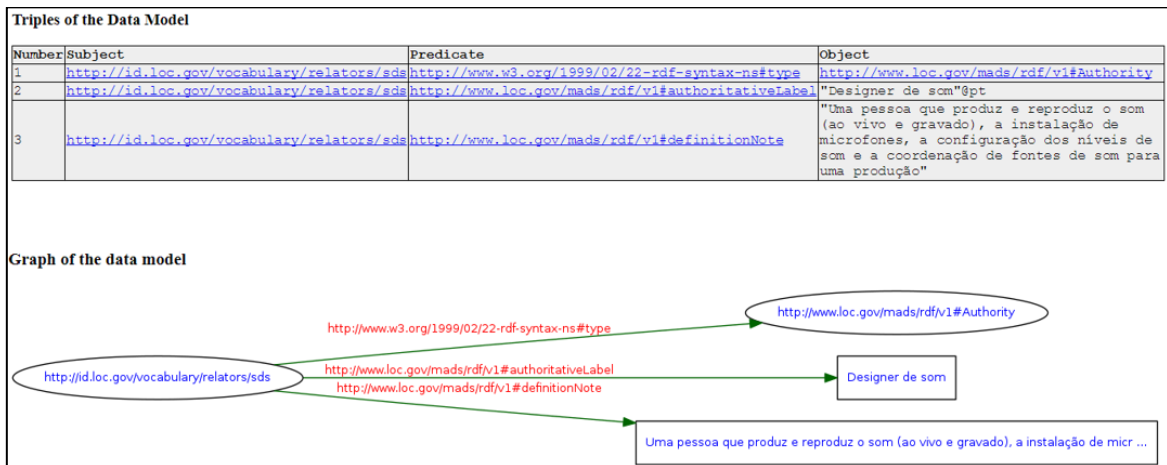
```

<?xml version="1.0" encoding="utf-8"?>
2: <rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
3:   <rdf:Description rdf:about="http://id.loc.gov/vocabulary/relators/sds"
4:     <rdf:type rdf:resource="http://www.loc.gov/mads/rdf/v1#Authority"/>
5:   </rdf:Description>
6:   <rdf:Description rdf:about="http://id.loc.gov/vocabulary/relators/sds"
7:     <ns0:authoritativeLabel xmlns:ns0="http://www.loc.gov/mads/rdf/v1#" xml:lang="pt">Designer de som</ns0:authoritativeLabel>
8:   </rdf:Description>
9:   <rdf:Description rdf:about="http://id.loc.gov/vocabulary/relators/sds"
10:    <ns0:definitionNote xmlns:ns0="http://www.loc.gov/mads/rdf/v1#">Uma pessoa que produz e reproduz o som (ao vivo e gravado),
11:    a instalação de microfones, a configuração dos níveis de som e a coordenação de fontes de som para uma produção</ns0:definitionNote>
12: </rdf:RDF>
  
```

Fonte: Os autores

Com o RDF não só é possível descrever, mas apresentar um conteúdo processável por máquina onde os relacionamentos que expressam o que é um termo e uma definição são legíveis por máquinas. E a partir dos esquemas de referência, os arquivos são validados, o XML como linguagem padrão transfere suas propriedades para o RDF. A figura 3 mostra o resultado deste processamento, utilizando o script da figura 2, apresentado por meio das triplas e de um grafo.

Figura 3 – Triplas e grafo RDF



Fonte: os autores

Nota-se que a relação sujeito, predicado e objeto é grande premissa do RDF, associada descrição de algo, pois não se estabelece limite para o que pode ser descrito, é possível estabelecer *n* relacionamentos entre arquivos e recursos na web, valendo-se como regra

precípua, a descrição, e a partir disso gerar esquemas de representação, como o grafo apresentado. Envolver a ótica da Ciência da Informação é um processo natural, diante da natureza desta ciência, todavia a apropriação destas tecnologias é um grande desafio, diante das transformações que ocorrem no curto espaço/tempo.

5 CONSIDERAÇÕES FINAIS

Neste estudo é possível identificar alguns conceitos chave para entender um pouco mais sobre a linguagem XML e o modelo RDF. Apesar de ser muito conhecido no ambiente tecnológico, por programadores e desenvolvedores no caso do XML, esse conhecimento muitas vezes passa a ser técnico e aplicado. Discutir metalinguagens e modelos de descrição no ambiente digital tem sido uma forma de aproximar estes aspectos técnicos para as discussões conceituais do fluxo da informação na web.

As implicações que envolvem o RDF está justamente na compreensão técnica de alguns termos, que tradicionalmente advém da Ciência da Computação, o que não restringe seu uso. Para estudiosos da Ciência da Informação, compreender as relações teóricas e práticas envolvendo tais tecnologias representam um avanço nas discussões que timidamente tem permeado a área. Apresenta-se sucintamente algumas das relações existentes entre os fundamentos que norteiam o desenvolvimento de tais padrões e conceitos da área de Ciência da Informação, no intuito de favorecer uma maior discussão desta temática a contribuir para o aprimoramento profissional em relação às novas tecnologias de representação

Este estudo preliminar apresenta um correlação entre os aspectos teóricos e práticos envolvendo XML/RDF, com caráter exploratório, para indicar os possíveis usos e aproximações da Ciência da Informação no manuseio e discussão dessas tecnologias. Espera-se que este trabalho possa contribuir com futuras pesquisas da área e possibilite um sólido aprendizado. Considera-se um pequeno passo para posteriores avanços da área e enriquecimentos dos profissionais da área.

Analysis of standards XML and RDF for Web representation under the perspective of information science: a preliminary study

ABSTRACT

From the development of information technologies new possibilities of representation were incorporated into the Web environment, among the main standards developed in the last decades and that have influenced the professional practices in the area of Information Science can highlight the language eXtensible Markup Language (XML) and the Resource Description Framework (RDF). Over the last few years, such technologies have revolutionized the way we manipulate information resources, however, it is possible to observe that many professionals in the area of Information Science still have difficulties regarding the understanding and use of these standards. In this sense, this article presents an ongoing research with the objective of describing the main characteristics of the XML and RDF standards, analyzing the conceptual aspects that underpin these technologies from the perspective of the Information Science area. This is a theoretical and exploratory study, carried out by means of a bibliographical survey on the subject matter, revisiting the literature of the area of Information Science and technical documents published by the W3C related to the analyzed standards. The results are briefly presented some of the relationships between the fundamentals that guide the development of such standards and concepts in the area of Information Science, in order to favor a greater discussion of this theme to contribute to the professional improvement in relation to the new technologies of representation

Keywords: Information Representation. Marking languages. Semantic Technologies. XML. RDF.

REFERÊNCIAS

BERNERS-LEE, T.; HENDLER, 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. 2001. Disponível em: < https://www-sop.inria.fr/acacia/cours/essi2006/Scientific%20American_%20Feature%20Article_%20The%20Semantic%20Web_%20May%202001.pdf>. Acesso em: jun. 2017.

BRAGA, Gilda Maria. Informação, ciência da informação: breves reflexões em três tempos. **Ciência da Informação**, Vol 24, número 1, 1995.

BRAY, Tim et al. **Extensible Markup Language (XML)**. W3C, 2006. Disponível em: < <https://www.w3.org/TR/xml11/>> Acesso em 27 out 2017.

BRICKLEY, Dan; Guha, RV. **RDF Schema 1.1**. W3C, 2014. Disponível em: < <https://www.w3.org/TR/rdf-schema/>> Acesso em 27 out 2017.

CYGANIAK, Richard; WOOD, David, LANTHALER, Markus. **RDF 1.1 Concepts and Abstract Syntax**. W3C, 2014. Disponível em: < <http://www.w3.org/TR/rdf11-concepts/>> Acesso em 27 out 2017.

Análise dos padrões XML e RDF para a representação na web sob a perspectiva da Ciência da Informação: um estudo preliminar

FERREIRA, Jaider Andrade; SANTOS, Plácida Leopoldina Ventura Amorim da Costa. O modelo de dados Resource Description Framework (RDF) e o seu papel na descrição de recursos. *Inf. & Soc.:Est.*, João Pessoa, v.23, n.2, p. 13-23, maio/ago. 2013.

FURGERI, Sérgio. O papel das linguagens de marcação para a Ciência da Informação. **TransInformação**, Campinas, 18(3):225-239, set./dez., 2006.

GOÑI, J. L.; FERNANDES, M. C P.; LUCENA, C. J. P. de. E-Learning e a Web Semântica. **Inf.MCC**, PUC-RIO, v. 12, n. 02, jun., 2002.

LE COADIC, Yves – François. **A Ciência da Informação**. Brasília: Brinquet de Lemos, 1996.

LIMA Júnio César de; CARVALHO, Cedric Luiz de. **Extensible Markup Language (XML)**. Relatório Técnico, Instituto de Informática da Universidade Federal de Goiás, 2005. Disponível em: <www.inf.ufg.br/sites/default/files/uploads/relatorios-tecnicos/RT-INF_002-05.pdf> Acesso em 27 out 2017.

MARTINS, Paulo; RAMALHO, Rogério Aparecido Sá. Tecnologias semânticas: fundamentos para o entendimento do conceito de Linked Data. In: SEMINÁRIO EM CIÊNCIA DA INFORMAÇÃO, 7., 2017, Londrina...**Anais...**Londrina: UEL, 2017. p.1040-1051. Disponível em: <<http://www.uel.br/eventos/cinf/index.php/secin2017/secin2107/paper/view/483/329>> Acesso em 27 out 2017.

MORENO, Fernanda Passini. **Requisitos Funcionais para Registros Bibliográficos – FRBR**: um estudo no catálogo da Rede Bibliodata. Dissertação (mestrado em Ciência da Informação), Brasília: UNB, 2006.

MORENO, Fernanda Passini; BRASCHER, Marisa. MARC, MARCXML E FRBR: relações encontradas na literatura. **Inf. & Soc.:Est.**, João Pessoa, v.17, n.3, p.13-25, set./dez. 2007.

NOVELLINO, Maria Salet Ferreira. Instrumentos e metodologias de representação da informação. **Inf.Inf.**, Londrina, v.1, n.2, p.37-45, jul./dez. 1996.

PALETTA, Francisco Carlos; MUCHERONI, Marcos Luiz. O desenvolvimento da WEB 3.0: Linked Data e DBPEDIA. **PRISMA.COM** (25) 2014, p. 73-90.

PICKLER, M. E. C. Web semântica: ontologias como ferramentas de representação do conhecimento. **Perspectivas em Ciência da Informação**, v. 12. n. 1, p. 63-83, jan/abr. 2007.

PINHEIRO, Lena Vania Ribeiro; LOUREIRO, José Mauro Matheus. Traçados e limites da ciência da informação. **Ciência da Informação**, Vol 24, número 1, 1995.

Análise dos padrões XML e RDF para a representação na web sob a perspectiva da Ciência da Informação: um estudo preliminar

RAMALHO, R. A. S. **Web Semântica**: aspectos interdisciplinares da gestão de recursos informacionais no âmbito da Ciência da Informação. Dissertação (mestrado em Ciência da Informação)- Faculdade de Filosofia e Ciências da Universidade Estadual Paulista – UNESP, Marília, 2006.

SIQUEIRA, Marcos Antonio. **XML na Ciência da Informação**: uma análise do MARC 21. Dissertação (mestrado em Ciência da Informação)- Faculdade de Filosofia e Ciências da Universidade Estadual Paulista – UNESP, Marília, 2003.

SOUZA, Renato Rocha; ALVARENGA, Lídia A Web Semântica e suas contribuições para a ciência da informação. **Ci. Inf., Brasília**, v. 33, n. 1, p. 132-141, jan./abr. 2004.
W3C. World Wide Web Consortium. **Web semântica**: dados vinculados. 2015. Disponível em:< <https://www.w3.org/standards/semanticweb/data.html> >Acesso em 20 jun 2017.