

# O MODELO DE DADOS RESOURCE DESCRIPTION FRAMEWORK (RDF) E O SEU PAPEL NA DESCRIÇÃO DE RECURSOS<sup>1</sup>

Jaider Andrade Ferreira\*

Plácida Leopoldina Ventura Amorim da Costa Santos\*\*

## RESUMO

Considerando o modelo de dados Resource Description Framework - RDF como um modo de representação e de descrição de recursos que visa à interoperabilidade e o acesso à informação em ambientes digitais, objetiva-se apresentar tal modelo, sua origem e suas principais características. Tratando-se de uma pesquisa teórica e exploratória com finalidade descritiva, foi utilizada a revisão de literatura como meio de investigação bibliográfica. Pôde-se perceber a importância do RDF para a descrição de quaisquer recursos, sejam eles digitais ou não, bibliográficos ou não. O RDF oferece para as comunidades de descrição de recursos a possibilidade de definirem a semântica de seus metadados de maneira formal, ou seja, de determinarem o significado dos elementos de metadados conforme as suas necessidades específicas de descrição e de modo processável por máquinas. Utilizando-se da Extensible Markup Language- XML como sintaxe para o intercâmbio e o processamento de metadados, o RDF colabora positivamente para a interoperabilidade entre os vários sistemas de informação e de descrição existentes, contribuindo, desse modo, para a construção de mecanismos de busca mais integrados que permitirão a oferta de serviços mais especializados aos seus usuários.

**Palavras-chave:** Resource Description Framework- RDF. Descrição de recursos. Web Semântica. Informação e tecnologia.

---

\* Bacharel em Biblioteconomia pela Universidade Estadual Paulista "Júlio de Mesquita Filho" – Campus Marília, Brasil. Mestrando no Programa de Pós-Graduação em Ciência da Informação da Universidade Estadual Paulista "Júlio de Mesquita Filho" – Campus Marília, Brasil.  
E-mail: jaideraf@gmail.com.

\*\* Livre-docente pela Faculdade de Filosofia e Ciências da Universidade Estadual Paulista "Júlio de Mesquita Filho" – Campus Marília, Brasil. Docente permanente do Programa de Pós-Graduação em Ciência da Informação da Universidade Estadual Paulista "Júlio de Mesquita Filho" – Campus Marília, Brasil.  
E-mail: placida@marilia.unesp.br

## I INTRODUÇÃO

Com o entendimento de que “a representação da informação é necessária em qualquer ambiente para proporcionar uma recuperação mais eficiente” (SANTOS; ALVES, 2009, p. 1), os desenvolvedores da Web têm procurado construir mecanismos e estruturas que possibilitem a atribuição de metadados aos recursos, visando à otimização do acesso à

informação e à recuperação de recursos nas mais variadas áreas do conhecimento.

A Web é um espaço distribuído de informação que provê uma abundância de fontes de informação heterogêneas em muitos formatos. A troca de informações na Web, no entanto, só é possível quando há acordo no estabelecimento de formatos de dados padrão (*standard data formats*) e na exploração dos *hyperlinks* que tornam os recursos distribuídos em estruturas do tipo Web. Vale destacar que essas estruturas não são somente utilizadas pelos agentes humanos para a leitura e a navegação,

---

<sup>1</sup> Pesquisa financiada pela Fundação de Amparo à Pesquisa do Estado de São Paulo - FAPESP (número do processo: 2011/15085-6).

mas também por mecanismos de busca para a coleta e o *ranking* de páginas.

A falta de estruturação, isto é, a falta de descrição das informações presentes nas páginas Web, fez com que os mecanismos de busca clássicos se tornassem insuficientes para gerenciar a quantidade sempre crescente de conteúdo.

Mesmo com a variedade de ferramentas disponíveis e apesar de estarem sempre em constante atualização em suas técnicas de busca, muitas vezes as ferramentas atuais não conseguem atender de modo satisfatório seus usuários. Apesar de toda tecnologia algumas limitações ainda ocorrem. Os principais empecilhos estão relacionados com a crescente quantidade de informações disponibilizadas; com as próprias limitações das ferramentas de busca (suas técnicas de busca e indexação dos sites) e com a falta de representação (tratamento) adequado dos recursos informacionais na rede [...] (ALVES, 2005, p. 25).

Nesse sentido, há uma tendência rumo a uma maior estruturação das informações disponibilizadas na Web. Estruturação essa que diz respeito à representação e à descrição dos recursos informacionais.

Da necessidade de estruturar e de descrever o conteúdo das páginas ou mesmo de representar e de descrever uma entidade não presente no ambiente digital, surge o projeto da Web Semântica. Segundo Santos e Alves (2009),

as tecnologias da Web Semântica convergem para a área de Ciência da Informação, estabelecendo uma estreita relação na questão da representação do conhecimento, principalmente no que se refere ao uso de metadados considerados essenciais no estabelecimento dos requisitos para uma boa representação dos recursos informacionais na rede.

De acordo com Catarino e Souza (2012, p. 78), “o alicerce da Web Semântica está num modelo de descrição dos recursos da Web, o *Resource Description Framework* (RDF)” e, por lidar com a descrição de recursos, a Catalogação descritiva está intimamente relacionada a esse tema.

Segundo Ortega (2011), o termo Catalogação descritiva se refere aos aspectos

da descrição formal dos recursos, incluindo a descrição física e a descrição dos elementos que identificam os recursos, sendo que a atividade de descrição é conhecida na área de Biblioteconomia pelo nome da disciplina Catalogação.

Dessa forma, concorda-se com Catarino e Souza (2012, p. 86) ao afirmarem que “a base da Web Semântica é a representação descritiva dos recursos a partir de um modelo em que os registros de metadados são representados por um conjunto de declarações [...]”.

Nesse contexto, este estudo objetiva apresentar brevemente o modelo *Resource Description Framework*- RDF como um importante mecanismo para a representação e a descrição de recursos. Considera-se que o entendimento do RDF é crucial ao ramo da Catalogação, uma vez que tal modelo será utilizado para a criação do novo padrão de intercâmbio de registros bibliográficos que substituirá o ainda utilizado *Machine Readable Cataloging* (MARC 21) (LIBRARY OF CONGRESS, 2011, 2012).

Segundo a Library of Congress (2011, tradução nossa, p. 3),

O projeto para a nova estrutura bibliográfica [*new bibliographic framework project*] estará focado no ambiente Web, nos princípios e mecanismos do *Linked Data* e no *Resource Description Framework* (RDF) como um modelo de dados básico. [...] é esperado que a utilização do RDF e de outros desenvolvimentos do W3C (*World Wide Web Consortium*) possibilitem a integração dos dados de bibliotecas e de outros dados de patrimônio cultural na Web para, assim, permitir um maior acesso dos usuários às informações.

Desse modo, apresenta-se a seguir o modelo RDF, sua origem e suas características.

## 2 O MODELO DE DADOS RDF

Considerando que a Web oferece acesso sem precedentes às informações distribuídas globalmente e que os metadados – dados estruturados sobre dados – melhoram a descoberta e o acesso a tais informações, o uso efetivo de metadados entre aplicativos requer convenções comuns acerca de semântica, de sintaxe e de estrutura (MILLER, 1998).

A semântica dos metadados, ou seja, seu significado, é definida por cada comunidade de descrição de recursos (agências catalogadoras, por exemplo) que busca atender a sua necessidade específica de descrição. A sintaxe, disposição sistemática dos elementos de dados, facilita a troca e o uso de metadados entre vários aplicativos, enquanto que a estrutura pode ser considerada como uma restrição formal sobre a sintaxe para a representação consistente da semântica (MILLER, 1998).

Segundo Miller (1998), o RDF, desenvolvido pelo *World Wide Web Consortium-W3C*, é uma infraestrutura que permite a codificação, o intercâmbio e o reuso de metadados estruturados. Esta infraestrutura permite a interoperabilidade de metadados por meio da concepção de mecanismos que suportam convenções comuns de semântica, de sintaxe e de estrutura. De acordo com o autor, o RDF não estipula a semântica para cada comunidade de descrição de recursos, mas sim oferece a capacidade para essas comunidades definirem os elementos de metadados conforme as suas necessidades específicas de descrição.

O RDF pode utilizar-se da *Extensible Markup Language-XML* como sintaxe comum para o intercâmbio e o processamento de metadados. Utilizando-se da XML, o RDF impõe uma estrutura que proporciona a expressão não ambígua da semântica e, desse modo, possibilita a codificação, o intercâmbio e o processamento consistente de metadados padronizados (MILLER, 1998).

Além disso, Miller (1998) destaca que o modelo RDF suporta o uso de convenções que facilitam a interoperabilidade modular entre conjuntos separados de elementos de metadados. Essas convenções incluem mecanismos padrão para a representação semântica que se baseiam em um simples, porém poderoso, modelo de dados. Adicionalmente, o *RDF Schema*, uma extensão do RDF, proporciona um meio para a publicação tanto de vocabulários legíveis por humanos quanto de vocabulários legíveis por máquinas.

Vocabulário, nesse contexto, é definido como um conjunto de propriedades ou termos descritivos, isto é, elementos de metadados, definidos por uma comunidade de descrição de recursos. Para o W3C (2012, tradução nossa, p. 1),

Na Web Semântica, vocabulários definem os conceitos e os relacionamentos (também referidos como “termos”) utilizados para descrever e representar uma área de interesse. Vocabulários são utilizados para classificar os termos que podem ser usados em uma aplicação particular, caracterizam as relações possíveis e definem possíveis restrições sobre o uso desses termos. Na prática, os vocabulários podem ser muito complexos (com vários milhares de termos) ou muito simples (descrevendo um ou dois conceitos apenas).

Fazendo-se uso de termos de vocabulários por meio de *Uniform Resource Identifiers - URI*, o RDF é um modelo de dados (*data model*) formal para a descrição de recursos, sendo que seu principal objetivo consiste em possibilitar que aplicações computacionais troquem (meta)dados enquanto preservam seu significado original (FURGERI, 2006; HITZLER; KRÖTZSCH; RUDOLPH, 2009).

Diferentemente da HTML e da XML, a principal intenção não é exibir os recursos corretamente, mas permitir o posterior processamento e a recombinação das informações que estão contidas nesses recursos, ou seja, os dados que compõem as suas representações. Dessa forma, o RDF é conseqüentemente e frequentemente visto como o modelo básico de representação para o desenvolvimento da Web Semântica (HITZLER; KRÖTZSCH; RUDOLPH, 2009, p. 19, tradução nossa).

Recentemente, a importância do RDF tem sido percebida pelas comunidades de descrição de recursos que, desse modo, o tem adotado para a formalização de seus termos descritivos. Como exemplo, cita-se a iniciativa Dublin Core (padrão ISO 15836-2009, ANSI/NISO Z39.85-2001, IETF RFC 5013-2007), comunidade internacional voltada para a descrição simples com foco na descoberta de recursos, que tem adotado o RDF desde 2002 (BECKETT; MILLER; BRICKLEY, 2002; NILSSON et al., 2008).

Zaidan (2011, p. 1) explica que o “RDF representa metadados no formato de sentenças sobre propriedades e relacionamentos entre itens [...]. Todos esses itens são os recursos existentes

e podem ser virtualmente qualquer objeto (texto, figura, música, vídeo, dentre outros) [...]”. Por esse motivo, Dziekaniak e Kirinus (2004, p. 29) afirmam que o RDF é “um modelo de representação para descrição semântica de recursos”.

Para melhor entender o contexto em que o RDF surgiu, apresenta-se a seguir um breve histórico de seu desenvolvimento.

### 3 ORIGEM DO RDF

Segundo Miller (1998), a história dos metadados no W3C começou em 1995 com o desenvolvimento do padrão *Platform for Internet Content Selection- PICS* que visava a possibilitar a classificação e a descrição do conteúdo de páginas Web. Com esse padrão seria possível descrever, por exemplo, se uma página Web continha conteúdo sexual, nudez, violência, palavras de baixo calão, etc. para que os pais, a partir de configurações do navegador, pudessem ter um controle sobre o que seus filhos poderiam acessar na Web. O desenvolvimento desse padrão, naquela época, foi motivado principalmente pela antecipação de restrições sobre o conteúdo da Internet nos EUA.

No entanto, no decorrer do desenvolvimento do PICS, percebeu-se a necessidade de fornecer descrições mais abrangentes acerca das páginas e, a partir de uma série de reuniões com a comunidade de bibliotecas digitais, limitações nas especificações do PICS foram identificadas e requisitos funcionais foram delineados para resolver o problema mais geral da associação de informações descritivas aos recursos presentes na Web. Com a percepção de que a questão da descrição de recursos era muito mais ampla do que a proposta feita inicialmente no contexto do PICS, o W3C criou um grupo de trabalho intitulado *Resource Description Framework- RDF* para discutir especificamente uma estrutura de descrição de recursos que fosse suficientemente abrangente para cobrir as necessidades das várias comunidades de descrição interessadas.

Nesse contexto, o processo de criação e de desenvolvimento do RDF na década de 1990 foi influenciado por várias linguagens, vocabulários de descrição já existentes e áreas do conhecimento. Uma das comunidades de descrição que influenciou a criação do RDF foi a

da iniciativa Dublin Core (MILLER, 1998) e uma linguagem que influenciou a sua criação foi a XML.

Segundo Dziekaniak e Kirinus, (2004, p. 27) “as principais influências vieram das comunidades de padronização da web (HTML, XML e SGML), da Biblioteconomia (metadados de catalogação), da representação do conhecimento (ontologias), da programação orientada a objetos, da linguagem de modelagem, entre outras”. Para Zaidan (2011), as duas grandes comunidades que influenciaram na criação do RDF foram a comunidade de bibliotecas digitais e a de representação do conhecimento.

Em 1999 uma especificação oficial sobre o RDF foi publicada pelo W3C (W3C, 1999), embora a ênfase naquela época estivesse apenas na representação de metadados sobre recursos presentes na Web e não na descrição de recursos em geral, como é o caso da especificação publicada em 2004 (W3C, 2004). O termo metadados geralmente se refere a “conjuntos de atributos, mais especificamente dados referenciais, que representam o conteúdo informacional de um recurso” (ALVES, 2005, p. 115). Em 1999 tais recursos eram considerados principalmente como páginas Web as quais o RDF poderia ajudar, por exemplo, na descrição de informações acerca da autoria ou do *copyright* dessas páginas. Posteriormente, com a visão da Web Semântica, o RDF foi estendido para representar também a informação semântica em geral, indo além das simples descrições de recursos Web como sujeitos primários de tais descrições. Isso gerou, em 2004, a publicação da especificação retrabalhada e estendida do RDF (HITZLER; KRÖTZSCH; RUDOLPH, 2009; W3C, 2004).

O modelo e a especificação da sintaxe RDF foram propostos em fevereiro de 1999 pelo W3C, com o intuito de possibilitar uma maior interoperabilidade no ambiente Web, oferecendo um padrão aberto para a descrição de recursos. Deste modo, o padrão RDF possibilita uma ampla gama de aplicações, permitindo que sejam feitas declarações a respeito de praticamente qualquer tipo de objeto, desde que este possa ser identificado a partir de um URI (RAMALHO, 2006, p. 70).

Devido à possibilidade de descrever recursos que não estão necessariamente

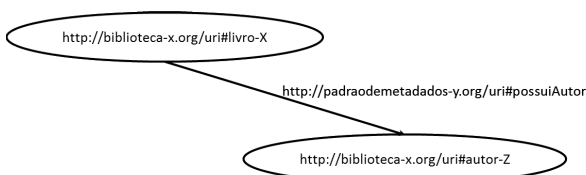


disponíveis na Web – um livro impresso, por exemplo – o RDF também pode ser utilizado para a troca de (meta)dados em diversas áreas de aplicação específicas, como é o caso da área de Biblioteconomia e, mais especificamente, da Catalogação.

#### 4 GRAFOSRDF

Um documento RDF é estruturado em forma de grafo direcionado (também chamado de grafo orientado, grafo dirigido, dígrafo ou quiver), isto é, um conjunto de nós (ou vértices) que estão ligados por arestas direcionadas (setas). Em RDF, tanto os nós quanto as arestas são rotulados com identificadores que os distinguem. A Figura 1 ilustra um exemplo de um grafo RDF simples com apenas dois nós e uma aresta.

**Figura 1** – Um grafo RDF simples descrevendo o relacionamento entre um livro e seu autor



Fonte: Adaptada de Hitzler, Krötzsch e Rudolph (2009, p. 20).

De acordo com Hitzler, Krötzsch e Rudolph (2009), é importante notar que, enquanto a informação codificada em XML apresenta uma estrutura em formato de árvore, o RDF apresenta a informação em forma de um grafo. O formato de árvore é perfeitamente adequado para organizar informações em documentos eletrônicos nos quais são frequentemente encontradas estruturas estritamente hierárquicas (documento > capítulo > corpo do texto > parágrafo > linha). Além disso, informações estruturadas em formato de árvore podem ser buscadas diretamente e serem processadas eficientemente. Mas, se o formato de árvore já é utilizado em XML e é tão eficiente, por que o RDF está estruturado em forma de grafos?

Uma razão importante para essa decisão é que o RDF não foi criado para a tarefa de estruturar documentos, mas sim para descrever relacionamentos genéricos entre objetos de interesse, ou seja, entre recursos de um domínio,

criando declarações acerca de tais recursos (FURGERI, 2006; HITZLER; KRÖTZSCH; RUDOLPH, 2009).

A Figura 1 ilustra um grafo que expressa que o livro X possui como autor o autor Z. Percebe-se que o relacionamento entre o livro X e o autor Z é uma informação que não pertence a um relacionamento do tipo hierárquico (em forma de árvore) entre os recursos livro X e autor Z. Assim, o modelo de dados RDF considera tais relacionamentos como blocos básicos de informação. Quando há vários relacionamentos como esses, as imagens que se formam naturalmente geram grafos, não estruturas em árvore, como as estruturas hierárquicas presentes em documentos XML (HITZLER; KRÖTZSCH; RUDOLPH, 2009).

Segundo os autores, outra razão pela escolha da utilização de grafos ao invés de árvores no modelo de dados RDF se deu pelo objetivo de tornar o RDF uma linguagem de descrição para os dados presentes na Web e em outras redes eletrônicas, isto é, as informações disponíveis nesses ambientes são normalmente armazenadas e gerenciadas por meios descentralizados e, desse modo, torna-se mais fácil combinar dados que estão em RDF a partir de múltiplas fontes. Por exemplo, os grafos RDF de um site podem ser reunidos com os grafos RDF de outros sites, gerando um grafo maior que pode ou não fornecer novas informações úteis. Desse modo, é possível reunir os registros bibliográficos em RDF de uma determinada biblioteca, por exemplo, com os registros bibliográficos em RDF de outras bibliotecas, possibilitando, portanto, a criação de um catálogo coletivo.

Segundo Furgeri (2006, p. 237), o RDF “elimina o problema da representação da informação em forma de árvore, criando uma estrutura mais flexível em forma de grafos, possibilitando a formação de uma cadeia de informações e estabelecendo uma rede de conhecimento”.

#### 5 URIs: IDENTIFICADORES EM RDF

Para representar um recurso em qualquer ambiente primeiramente é necessário identificá-lo. Na Web, a identificação de recursos é, de modo geral, realizada por meio de uma cadeia de

caracteres chamada de *Uniform Resource Identifier* (URI).

Ray (2001) aponta que existem duas categorias principais de URI: a primeira identifica um recurso com exclusividade com base em sua localização, enquanto que a segunda categoria apenas atribui ao recurso um nome exclusivo. Segundo o autor, um URI começa com um esquema - um nome curto que especifica como o recurso está sendo identificado. "Normalmente, ele é um protocolo de comunicações como HTTP ou FTP. Isso é seguido por um sinal de dois pontos (:) e uma sequência de dados que identifica o recurso com exclusividade" (RAY, 2001, p. 77).

Na Web Semântica é o URI que torna o recurso único e não confundível com outros recursos. Segundo Ramalho (2006, p. 71), os URIs possibilitam "uma maneira global e única de nomear itens".

URIs englobam os *Uniform Resource Locators*- URL, isto é, os endereços Web que são utilizados para acessar os recursos online. Cada URL é um URI válido e, portanto, pode ser utilizado como um identificador em documentos que descrevem recursos disponíveis na Web. No entanto, Hitzler, Krötzsch e Rudolph (2009) chamam a atenção para o fato de que em várias aplicações o objetivo não é simplesmente a troca de informações ou metadados sobre recursos Web, mas sobre vários e diferentes tipos de recursos que estão, inclusive, fora da Web. Por meio da Web, é possível descrever qualquer objeto ou recurso que tenha uma clara identidade no contexto de uma dada aplicação: livros, pessoas, locais, editoras, eventos, relacionamentos entre essas coisas, conceitos abstratos e muito mais. Tais recursos podem, obviamente, não ser recuperados online e, portanto, seus URIs são utilizados exclusivamente para fins de identificação única.

Mesmo aqueles URIs que não se referem a recursos disponíveis na Web podem ser construídos segundo a estrutura padrão de endereçamento Web (esquema://autoridade.com/caminho/a/ser/percorrido?consulta#fragmento).

Hitzler, Krötzsch e Rudolph (2009) afirmam que a principal característica de qualquer URI é a sua parte inicial, ou seja, seu esquema. Os esquemas como o *http* são tipicamente associados ao protocolo para a transmissão de informações na Web, entretanto, pode-se encontrar também tais

esquemas em muitos URIs que não se referem a uma localização na Web. Dessa forma, os detalhes do protocolo não são, obviamente, relevantes quando se utiliza um URI apenas como um nome, isto é, como um identificador. O URI <http://biblioteca-x.org/uri#livro-X>, como exemplo para essa situação, pode ser utilizado para se referir a um determinado livro em uma determinada biblioteca. O importante, no entanto, não é a recuperação ou não de algo por um navegador a partir desse URI, nem mesmo se o quê é recuperado tem ou não alguma relação com o livro em questão, mas sim a própria identificação do recurso.

## 6 LITERAIS:VALORES DE DADOS EM RDF

Como dito anteriormente, URIs permitem nomear e identificar recursos abstratos, mesmo aqueles recursos que não podem ser processados diretamente por um computador (pessoas, livros impressos, lugares, editoras, sintomas, etc.). Nesses casos, tais URIs são tratados como identificadores, referências ou nomes aos objetos pretendidos. Uma vez que os URIs podem ser tratados como nomes, a atual interpretação pretendida de um particular URI não é dada de nenhuma maneira formal e, portanto, ferramentas específicas podem ter seus próprios meios para interpretar URIs (HITZLER; KRÖTZSCH; RUDOLPH, 2009).

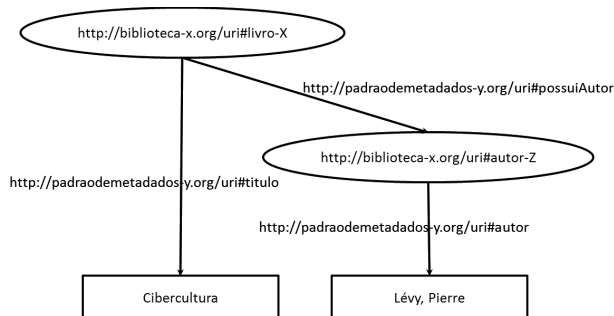
Nesse contexto, um serviço Web qualquer pode reconhecer URIs que se referem a livros e tratá-los de alguma forma especial, exibindo, por exemplo, detalhes de publicação ou mesmo os locais em que esses livros estão disponíveis para empréstimo. Segundo Hitzler, Krötzsch e Rudolph (2009), esse nível de liberdade é útil e inevitável quando se lida com recursos arbitrários. Porém, a situação se torna diferente quando se lida com valores de dados concretos como números, datas, valores absolutos ou cadeias de caracteres. Nota-se que em registros bibliográficos os dados encontrados são frequentemente cadeias de caracteres (*strings*).

Nesses casos, espera-se que cada aplicação tenha um entendimento mínimo dos significados concretos de tais valores. O número 36, por exemplo, tem a mesma interpretação numérica em qualquer contexto (HITZLER; KRÖTZSCH; RUDOLPH, 2009). Em RDF, esses

valores de dados são representados pelo que os profissionais da Ciência da Computação chamam de valores literais, ou apenas, literais. Literais são valores de dados de um certo tipo de dados (*datatype*). O valor de cada literal é geralmente descrito como uma sequência de caracteres, tais como a cadeia de caracteres composta pelos símbolos 3 e 6 do exemplo anterior. A interpretação de tais sequências realizada pela máquina é, então, baseada em um tipo de dados específico. Declarar o tipo de dados é essencial para o entendimento do significado pretendido: as cadeias de caracteres 36 e 036, por exemplo, referem-se ao mesmo número natural, no entanto, são diferentes cadeias de caracteres de texto (HITZLER; KRÖTZSCH; RUDOLPH, 2009).

Como pode ser visto na Figura 2, retângulos são utilizados para distinguir valores literais de URIs nos grafos RDF (FURGERI, 2006, p. 236).

**Figura 2** - Um grafo RDF com literais para a descrição de valores de dados



Fonte: Adaptada de Hitzler, Krötzsch e Rudolph (2009, p. 24).

Outra característica importante dos grafos RDF é que valores literais não podem ser utilizados como origem de arestas. Na prática, isso significa que não se pode fazer declarações acerca de literais, ou seja, as triplas permitidas em RDF seguem o padrão recurso-propriedade-valor, nunca valor-propriedade-valor, ou valor-propriedade-recurso. Assim, Hitzler, Krötzsch e Rudolph (2009) alertam que essa restrição precisa ser levada em consideração quando se planeja a modelagem de dados em RDF.

## 7 SINTAXE PARA O RDF

O modelo RDF é descrito em forma de grafos por meio de diagramas. Esse modo

de representar os dados em RDF facilita o aprendizado do modelo e a leitura por humanos, além de servir como um modo preciso para a modelagem conceitual de um domínio. Mas é evidente que os computadores não são adequados para o processamento e o intercâmbio de grafos uns com os outros. Para humanos, entender grafos visualmente funciona muito bem se os grafos são pequenos, com poucos nós, arestas e literais. Mas os gigantescos bancos de dados atuais contêm milhares de dados que gerariam grafos absurdamente grandes.

Para representar um conjunto de dados modelados em RDF, existem meios que não se utilizam de diagramas, mas sim de cadeias de caracteres que podem ser processadas por máquinas.

Para que isso ocorra, um grafo RDF precisa ser dividido em partes menores que podem ser armazenadas uma por uma. Essa transformação de estruturas de dados complexas para cadeias de caracteres lineares é chamada de serialização (HITZLER; KRÖTZSCH; RUDOLPH, 2009).

### 7.1 Triplas RDF

De acordo Hitzler, Krötzsch e Rudolph (2009), o primeiro passo para serializar um grafo RDF é isolar cada uma de suas relações, o que conduz a um recurso, uma propriedade e um valor para cada relação.

Tomando como exemplo a Figura 2, tem-se o recurso “<http://biblioteca-x.org/uri#livro-X>”, a propriedade “<http://padraometadados-y.org/uri#possuiAutor>” e o valor “<http://biblioteca-x.org/uri#autor-Z>”. Essas três partes distintas são chamadas de sujeito, predicado e objeto, respectivamente.

Percebe-se que cada grafo RDF pode ser completamente descrito por meio de seus nós e arestas e convertido em sujeitos, predicados e objetos. Cada conjunto sujeito-predicado-objeto é considerado uma tripla RDF, ou seja, uma declaração.

Furgeri (2006) explica que o recurso é o sujeito de uma declaração. Pode ser um website, um livro, um CD, etc. Um artigo científico é um exemplo de recurso. A propriedade é o predicado de uma declaração, é um atributo utilizado para descrever um recurso. Um artigo científico pode conter diversas propriedades: nome do autor, título do artigo,

data de publicação, etc. O valor é o objeto de uma declaração e representa o conteúdo das propriedades. Seguindo o exemplo, trata-se dos conteúdos referentes ao nome do autor, ao título do artigo e a data de publicação.

Cabe observar que o sujeito e o predicado precisam ser URIs, enquanto que o objeto pode ser um URI, isto é, outro recurso, ou um valor literal (HITZLER; KRÖTZSCH; RUDOLPH, 2009).

Em relação à aparência de um grafo RDF, Furgeri (2006, p. 236) aponta que normalmente “o sujeito (recurso) é um nó [...] em forma de elipse, o predicado (propriedade) é um arco

com uma seta apontando para o objeto (valor), cuja representação é um retângulo (se for um literal) ou uma nova elipse (se representar outro recurso)”.

## 7.2 A serialização XML do RDF

Atualmente, a principal sintaxe utilizada para serializar um grafo RDF é a sintaxe XML (RDF/XML). Outras sintaxes possíveis são: Notation3 (N3), N-Triples, Turtle e JSON.

O grafo RDF da Figura 2 pode ser serializado com a sintaxe XML como mostrado na Figura 3.

**Figura 3** –Serialização em XML de um grafo RDF que descreveo relacionamento entre um livro e um autor, o título do livro e o nome do autor.

```
<?xml version="1.0" encoding="utf-8"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:cod="http://padraodemetadados-y.org/uri#"

  <rdf:Description rdf:about="http://biblioteca-x.org/uri#livro-X">
    <cod:titulo>Cibercultura</cod:titulo>
    <cod:possuiAutor>
      <rdf:Description rdf:about="http://biblioteca-x.org/uri#autor-Z">
        <cod:nomeDoAutor>Lévy, Pierre</cod:nomeDoAutor>
      </rdf:Description>
    </cod:possuiAutor>
  </rdf:Description>
</rdf:RDF>
```

Fonte: Elaborada pelos autores.

Uma vez que este estudo visa apenas à apresentação do modelo RDF, sua origem suas principais características, não serão apresentados os detalhes técnicos de serialização ou de sintaxe XML para a codificação e o intercâmbio de documentos RDF.

## 8 VANTAGENS DA UTILIZAÇÃO DO RDF

Souza e Alvarenga (2004) constataam alguns benefícios providos pela utilização do RDF, como, por exemplo, a viabilização de um ambiente consistente para a publicação e utilização de metadados por meioda infraestrutura da XML, o estabelecimento de uma sintaxe padronizada para a descrição de recursos com suas propriedades e respectivos valores e

a possibilidade de aplicações agirem de modo inteligente e automatizado sobre as informações publicadas.

Zaidan (2011) considera que o RDF é independente de domínio e, como é composto pela tripla sujeito-predicado-objeto, consegue fazer uma representação primitiva com vistas a uma criação maior, tendo a finalidade de embutir semântica (enriquecimento semântico) e fornecendo a interoperabilidade e a semântica para metadados de modo a facilitar buscas por recursos, já que, até então, esses recursos estavam sendo procurados por mecanismos de busca textual simples.

Furgeri (2006, p. 237) destaca que o RDF

resolve o problema da diversidade na representação da informação que ocorre em XML, criando ligações únicas



entre recursos e [com o auxílio do RDF Schema] estabelecendo vocabulários por meio de namespaces. A RDF elimina o problema da limitação do tamanho da estrutura enfrentada pela XML, criando ponteiros que unem documentos com estruturas menores.

Segundo o autor, o RDF é mais indicado para a criação de metadados, “pois um recurso é referenciado a um objeto por meio de um predicado com significado próprio” (FURGERI, 2006, p. 237). O autor aponta também que o RDF possui atribuições que lhe conferem diversas possibilidades de uso na área de Ciência da Informação, uma vez que esse modelo de dados permite criar relações entre documentos e estabelecer significado aos termos. O “RDF provê uma estrutura mais flexível que XML, aproximando-se da forma como os seres humanos relacionam informações, isto é, por associações” (FURGERI, 2006, p. 237).

Para Moura (2001 apud DZIEKANIAK; KIRINUS, 2004), na área de descoberta de recursos, o RDF possibilita o desenvolvimento de mecanismos de pesquisa mais eficientes. Na área de Catalogação, o RDF pode ser utilizado para a descrição de recursos de informação (LIBRARY OF CONGRESS, 2011, 2012; CASTRO, 2012) e na área de agentes inteligentes o RDF pode facilitar o intercâmbio de informações.

## 9 CONSIDERAÇÕES FINAIS

A partir dos apontamentos apresentados, pode-se considerar que a Ciência da Informação, especialmente em razão da Catalogação descritiva, está intimamente relacionada com um dos alicerces da Web Semântica: o modelo para a descrição de recursos *Resource Description Framework*- RDF.

Por meio de uma revisão de literatura acerca do modelo RDF, sua origem e suas principais características, percebe-se a sua importância para a atividade de descrição de quaisquer recursos, sejam eles digitais ou não, bibliográficos ou não.

O modelo RDF oferece a possibilidade para as comunidades de descrição de recursos definirem a semântica de seus metadados de maneira formal, isto é, definindo o significado dos elementos de metadados, conforme as suas necessidades específicas de descrição, em um modelo processável por máquinas.

Utilizando-se da XML como sintaxe comum para o intercâmbio e o processamento de metadados, o RDF colabora positivamente para a interoperabilidade entre os vários sistemas de informação e de descrição existentes contribuindo, desse modo, para a construção de mecanismos de busca mais integrados que permitirão a oferta de serviços mais especializados aos seus usuários.

Um dos desafios percebidos para a utilização prática do RDF pelas comunidades de descrição de recursos é a correta criação e utilização de URIs que servirão de sujeitos (identificadores que representam recursos), predicados (identificadores que representam as propriedades/atributos dos recursos) e objetos (valores de propriedades que podem ser, eles mesmos, outros sujeitos) nas declarações constituintes das descrições em RDF.

De modo geral, considera-se que o modelo RDF pode ser utilizado para representar recursos de maneira simples, flexível e interpretável, podendo conter representações abrangentes, do ponto de vista cognitivo, e sintéticas para a tomada de decisão dos usuários sobre os recursos buscados. Além disso, o modelo RDF é acessível tanto para as máquinas (incluindo o simples processamento de dados, a leitura e a tomada de decisão por agentes inteligentes) quanto para os seres humanos que se utilizarão de interfaces para a consulta e apresentação dos dados contidos nesses documentos ou bases de dados modelados de acordo com o RDF.

## AGRADECIMENTOS

Os autores agradecem à FAPESP pelo financiamento desta pesquisa (processo nº 2011/15085-6).

## THE RESOURCE DESCRIPTION FRAMEWORK (RDF) DATA MODEL AND ITS ROLE IN THE RESOURCE DESCRIPTION

### Abstract

The Resource Description Framework - RDF is a data model that represents and describes resources aiming to interoperability and access to information in digital environments. In this paper, we present the RDF data model, its origins and its main features. Since this is a theoretical and exploratory research with a descriptive purpose, we used literature review as a methodology for our bibliographical investigation. It was possible to demonstrate the importance of RDF for description of any resources, whether they are digital or not, bibliographic or not. Furthermore, RDF provides for resource description communities the opportunity to define the semantics of their metadata in a formal way, i.e., to determine the meaning of the metadata elements according to their specific needs for description, and in a way readable by machines. Using the Extensible Markup Language - XML syntax for metadata exchange and processing, RDF contributes positively to interoperability between different information and description systems, thus contributing to the construction of integrated search engines that will enable specialized services to its users.

### Keywords:

Resource Description Framework - RDF. Resource Description. Semantic Web. Information and technology.

### REFERÊNCIAS

- ALVES, R. C. V. **Web semântica: uma análise focada no uso de metadados**. 2005. 180 f. Dissertação (Mestrado em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2005. Disponível em: <[http://www.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/alves\\_rcv\\_me\\_mar.pdf](http://www.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/alves_rcv_me_mar.pdf)>. Acesso em: 2 mar. 2013.
- BECKETT, D.; MILLER, E.; BRICKLEY, D. **Expressing Simple Dublin Core in RDF/XML**. Ohio: Dublin Core Metadata Initiative, 2002. Disponível em: <<http://www.dublincore.org/documents/dcmes-xml>>. Acesso em: 22jan. 2013.
- CASTRO, F. F. de. **Elementos de interoperabilidade na catalogação descritiva: configurações contemporâneas para a modelagem de ambientes informacionais digitais**. 201f. Tese (Doutorado em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2012. Disponível em: <[http://www.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/Castro,%20F.F.\\_doutorado\\_CI\\_2012.pdf](http://www.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/Castro,%20F.F._doutorado_CI_2012.pdf)>. Acesso em: 15 dez. 2012.
- CATARINO, M. E.; SOUZA, T. B. de. 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/viewarticle.php?id=472>>. Acesso em: 10jan. 2013.
- DZIEKANIAK, G. V.; KIRINUS, J. B. Web Semântica. **Encontros Bibli**, Florianópolis, n. 18, 2º sem. 2004. Disponível em: <<http://www.periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2004v9n18p20>>. Acesso em: 26fev. 2013.
- FURGERI, S. O papel das linguagens de marcação para a Ciência da Informação. **TransInformação**, Campinas, v. 18, n. 3, p. 225-239, set./dez. 2006. Disponível em: <<http://revistas.puc-campinas.edu.br/transinfo/viewarticle.php?id=184>>. Acesso em: 13jan. 2013.
- HITZLER, P.; KRÖTZSCH, M.; RUDOLPH, S. **Foundations of Semantic Web technologies**. Boca Raton: CRC Press, 2010.
- LIBRARY OF CONGRESS. A Bibliographic Framework for the Digital Age. *Bibliographic*

*Framework Transition Initiative: News and Announcements*, Oct. 2011. Disponível em: <<http://www.loc.gov/bibframe/news/framework-103111.html>>. Acesso em: 20 set. 2012.

LIBRARY OF CONGRESS. *Bibliographic Framework as a Web of Data: Linked Data Model and Supporting Services*. Washington, 2012. Disponível em: <<http://www.loc.gov/marc/transition/pdf/marclid-report-11-21-2012.pdf>>. Acesso em: 25 nov. 2012.

MILLER, E. An Introduction to the Resource Description Framework. *D-Lib Magazine*, v. 4, n. 5, May, 1998. Disponível em: <<http://www.dlib.org/dlib/may98/miller/05miller.html>>. Acesso em: 16 jan. 2013.

NILSSON, M. et al. **Expressing Dublin Core metadata using the Resource Description Framework (RDF)**. Ohio: Dublin Core Metadata Initiative, 2008. Disponível em: <<http://www.dublincore.org/documents/dc-rdf>>. Acesso em: 5 nov. 2012.

ORTEGA, C. D. Do princípio monográfico à unidade documentária: exploração dos fundamentos da catalogação. *Liinc em Revista*, v. 7, n. 1, p. 43-60, 2011. Disponível em: <<http://revista.ibict.br/liinc/index.php/liinc/article/viewFile/402/263>>. Acesso em: 15 jan. 2013.

RAMALHO, R. A. S. **Web Semântica: aspectos interdisciplinares da gestão de recursos informacionais no âmbito da Ciência da Informação**. 2006. 120 f. Dissertação (Mestrado em Ciência da Informação). Faculdade de Filosofia e Ciências - Universidade Estadual Paulista, Marília, 2006. Disponível em: <[\[www.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/ramalho\\\_ras\\\_me\\\_mar.pdf\]\(http://www.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/ramalho\_ras\_me\_mar.pdf\)>. Acesso em: 10 jan. 2013.](http://</a></p></div><div data-bbox=)

RAY, E. **Aprendendo XML**. Rio de Janeiro: Campus; O'Reilly, 2001.

SANTOS, P. L. V. A. da C.; ALVES, R. C. V. Metadados e Web Semântica para estruturação da Web 2.0 e Web 3.0. **DataGramZero**, Rio de Janeiro, v. 10, n. 6, dez. 2009. Disponível em: <[http://www.dgz.org.br/dez09/Art\\_04.htm](http://www.dgz.org.br/dez09/Art_04.htm)>. Acesso em: 25 nov. 2012.

SOUZA, R. R.; ALVARENGA, L. A Web Semântica e suas contribuições para a Ciência da Informação. **Ciência da Informação**, Brasília, v. 33, n. 1, p. 132-141, jan./abr. 2004. Disponível em: <<http://dx.doi.org/10.1590/S0100-19652004000100016>>. Acesso em: 12 dez. 2012.

W3C. **Vocabularies**. Cambridge: W3C, 2012. Disponível em: <<http://www.w3.org/standards/semanticweb/ontology>>. Acesso em: 23 out. 2012.

W3C. **Resource Description Framework (RDF) Model and Syntax Specification**. Cambridge: W3C, 1999. Disponível em: <<http://www.w3.org/TR/1999/REC-rdf-syntax-19990222>>. Acesso em: 15 nov. 2012.

W3C. **RDF Primer**. Cambridge: W3C, 2004. Disponível em: <<http://www.w3.org/TR/rdf-primer>>. Acesso em: 15 nov. 2012.

ZAIDAN, F. XML, RDF e OWL. *IT Web*, 2011. Disponível em: <<http://itweb.com.br/blogs/xml-rdf-e-owl-para-saber-um-pouco-mais-sobre-a-web-semantica-1>>. Acesso em: 10 out. 2012.