



XVII Encontro Nacional de Pesquisa em Ciência da Informação (XVII ENANCIB)

GT 7 – Produção e Comunicação da Informação em Ciência, Tecnologia & Inovação

**UTILIZAÇÃO DA WEB SEMANTICA E RDF EM ESTUDOS MÉTRICOS DA
INFORMAÇÃO: aplicação na base Brapci**

***USE OF THE SEMANTIC WEB AND RDF IN THE INFORMATION METRIC
STUDIES: APPLICATION IN BRAPCI BASE***

Rene Faustino Gabriel Junior¹

Modalidade da apresentação: Comunicação Oral

Resumo: No domínio dos estudos métricos da informação, as bases de dados estão tendo dificuldades em atender as demandas de informações dos pesquisadores, seja pela limitação dos metadados disponibilizados ou pela dificuldade de exportar grande número de registros. O objetivo deste estudo é investigar se o formato FRBR e RDF podem ser aplicados a Brapci de forma a contribuir para a disponibilização de dados e geração de indicadores para os estudos métricos da informação. O estudo iniciou com uma pesquisa exploratório buscando identificar padrões de ontologias passíveis de serem aplicadas ao contexto de artigos científicos, com base nesses padrões foi desenvolvido um programa que exportou todos os dados da Brapci para um formato de triplas n3 e rodar no software *GraphDB Free*, de forma a possibilitar a execução das *queries* de consulta e obter respostas à algumas perguntas norteadoras de teste. A base foi validada comparando com um estudo bibliométrico que utilizou a Brapci e a construção das consultas objetivando a resolução das questões propostas. Os resultados foram resolvidos adequadamente, demonstrando a viabilidade da aplicação nos processos de recuperação e produção de indicadores nos estudos métricos da informação. Os resultados obtidos foram promissores, porém ainda existe a dificuldade da linguagem SPARQL ser muito difícil de compreensão para os pesquisadores da Ciência da Informação, necessitando de melhorias na interface com o usuário.

Palavras-chave: Web Semântica. Resource Description Framework. Bases de Dados. Métricas da informação e Comunicação. Bibliometria.

Abstract: *In the field of metric studies of information, the bases are having trouble meeting the demands of information from researchers. Lower number of metadata and the difficulty of exporting*

¹ UFRGS

large number of records cause this limitation. The aim of this study is to investigate whether FRBR and RDF format to Brapci in order to contribute to the availability of data and generation of indicators for the metric information studies. The study began with an exploratory research seeking to identify patterns insusceptible ontologies to the papers context, based on these standards developed a program that exported all data Brapci for triple format n3 and rotate GraphDB Free software, in order to enable the implementation of the consultation queries and get answers to some test guiding questions. The base validated by comparing with a bibliometric study using the Brapci and construction of consultations aimed at resolving the issues proposed. The results were resolved appropriately, demonstrating the feasibility of application in the recovery process and production indicators in metric studies of information. The results were promising, but there is still the difficulty of the SPARQL language is very difficult to understand for students of Information Science, requiring improvements in the user interface.

Keywords: Web Semantic. Resource Description Framework. Data Bases. Metrics. Bibliometric.

1 INTRODUÇÃO

Durante décadas o MARC tem sido o formato de metadados que as bibliotecas adotaram como representação descritiva de seus acervos (BRANTLEY, 2015). Com o início da era digital o formato MARC está limitando a representação dos documentos produzidos principalmente em ambiente virtual, por não contemplar todas as possibilidades de relações dos documentos e entre documentos. Diversas instituições como a *Library of Congress*, *International Federation of Library Associations and Institutions* (IFLA), *World Wide Web Consortium* (W3C) entre outras estão buscando um modelo que possa desenvolver um alicerce para o futuro da descrição bibliográfica dentro do mundo conectado (Web), e utilizar os princípios e mecanismos de Dados Lincados (*Linked Data*) como modelo de dados básico para a integração dos acervos de bibliotecas e de outros dados do patrimônio cultural na Web de forma a permitir maior acesso dos usuários a informações.

Neste contexto, a IFLA configura uma recomendação para reestruturar os registros bibliográficos de maneira a refletir a estrutura conceitual de buscas de informação, levando em conta a diversidade de usuários, materiais, suporte físico e formatos e oferece uma nova perspectiva da estrutura e nos relacionamentos dos registros bibliográficos com o FRBR. (MORENO; ARELLANO, 2005).

Por se tratar de um modelo conceitual o FRBR precisa de meios para se concretizar em práticas dentro das bibliotecas e sistemas de informação. Na busca de um novo formato que seja compatível com o FRBR, e que possa ser compreendido pelo ser humano e pelos computadores, emerge o padrão *Resource Description Framework* (RDF), padrão esse

compatível com as propostas da Web Semântica (WS) e integração de informações em Dados Lincados.

Como o processo de representação descritiva, os estudos métricos da informação estão sendo limitados pela modelagem das bases de dados bibliográficas, que são a principal fonte de informação de pesquisa. Pelo número limitado de metadados, os estudos resumem-se em investigações sobre a produção científica, produção em colaboração, análises de redes e análises de citação dentro de um *corpus* específico.

A maioria dos estudos métricos de informação² utilizam dados e registros disponibilizados pelas bases como a *Web of Science*, Scopus, Scielo, Plataforma Lattes, Brapci, Base ABCDM, Redalyc, DOAJ entre outras, ou acabam formando sua própria *corpora* para análise. Destaca-se que a maior fração de tempo da pesquisa, nesses tipos de estudos, é consumido para higienização dos dados, seja pela eliminação de erros ou para normalização, o que demanda muitos recursos a fim de propiciar o agrupamento e concatenação dos dados e produzir os indicadores necessários, reservando o menor tempo para os resultados e análises.

Os atuais sistemas de armazenamento de dados bibliográficos foram construídos utilizando a Modelagem Entidade Relacionamento, esta metodologia empregada no desenvolvimento de sistemas limita as relações entre as entidades, pois são necessárias as definições prévias dos atributos que irão descrevê-las. A Brapci também tem essas limitações, pois foi concebida utilizando o padrão de representação descritiva baseado nesse modelo. Qualquer alteração nos atributos das entidades (novos metadados) provoca grande impacto nos sistemas, carecendo de análise de analistas e programados a fim de minimizar os problemas em todo o sistema.

Neste contexto, um novo desafio se instaura na Ciência da Informação (CI), com a busca de novas formas de organizar e reaproveitar todo conteúdo hoje disponível e a ser disponibilizado, de forma a proporcionar acesso ao conteúdo ao seu usuário. A CI tem no domínio da catalogação o desafio de representar os diversos tipos de materiais a partir das novas tecnologias, o desenvolvimento de interfaces mais amigáveis tanto para o profissional catalogador quanto para o usuário final e a definição de parâmetros e padronizações que permitam um nível global de cooperação. O FRBR, mais especificamente o RDF, é uma das mais promissoras delas, porém, por se tratar de um modelo conceitual é necessário o desenvolvimento de metodologias e produtos concretos para sua avaliação e efetivação

2 Entende-se estudos métricos da informação incluem-se os estudos bibliométricos, cientométricos, altmétricos, patente métricos e webométricos.

(FUSCO, 2011), bem como difundir entre as bibliotecas, bibliotecários e os usuários.

Uma outra vertente que instiga os pesquisadores é como esses mecanismos podem contribuir para a efetividade dos estudos métricos da informação, visto que com utilização do Dados Lincados, o universo de pesquisa se amplia consideravelmente, não mais limitando as hipóteses dos pesquisadores.

O objetivo deste estudo é investigar se o formato RDF aplicado a Brapci pode contribuir para a geração de indicadores para os estudos métricos da informação. Para desenvolver o objetivo principal constituem-se os seguintes objetivos secundários: identificar na literatura quais as ontologias utilizadas para a representação de artigos científicos; desenvolver um algoritmo de conversão dos dados da Brapci para o formato RDF com base nas ontologias identificadas; aplicar na base RDF consultas para a resolução de pergunta de pesquisa. De forma a nortear a investigação, foram propostas algumas perguntas que serão aplicadas a *corpora* de análise, as perguntas a serem respondidas pela base RDF foram:

- a) Quais autores publicaram trabalhos em 1972 e na década de 2010?
- b) Quem são os colaboradores de Leilah Santiago Bufrem nos estudos descritivos?
- c) Quais as opções metodológicas que o Suzana Pinheiro Machado Mueller mais utilizou em seus estudos publicados?
- d) Quais os estudos de caso foram aplicados em bibliotecas?
- e) Quais os temas (palavras-chave) mais publicados entre 2010 e 2015?

A escolha da Brapci se deve pela facilidade de acesso do pesquisador responsável, e pela circunstância especial de que ela já mantém dados estruturados na modelagem Entidade Relacionamento (ER). A Brapci também tem incorporada em seu espaço documentário a maioria das publicações em CI no Brasil, permitindo ao pesquisador uma visão de conjunto da produção na área, e ao mesmo tempo revelando as especificidades do domínio científico. As informações das publicações dos autores nas bases de dados seguem estruturas de dados seguindo os formatos bibliográfico compatíveis com o padrão de metadados MARC e Dublin Core.

A hipótese do estudo parte de que um dos grandes ganhos do uso dos dados estruturados em RDF está na forma de perguntar para o sistema, quando no sistema tradicional define-se uma estratégia de busca utilizando palavras-chave a fim de formar um *corpus* para a análise métrica da informação, na WS pode-se além de recuperar pelos termos, realizar perguntas a serem respondidas pela linguagem SPARQL, fazendo com que o sistema crie inferências lógicas.

2 REFERENCIAL TEÓRICO

Os padrões de metadados Dublin Core e MARC são formatos conhecidos e associados a padrões específicos para facilitar a descrição, acesso e uso dos recursos informacionais. A principal finalidade da descrição por meio dos metadados é contribuir para uma busca e recuperação eficiente e eficaz da informação. Todo o contexto da representação da informação e do conhecimento visa garantir uma precisa recuperação de modo a satisfazer as necessidades informacionais dos usuários (PINHO; NASCIMENTO; MELO, 2015).

De acordo com Catarino e Souza (2012), no panorama atual, mudanças significativas estão sendo provocadas pelas tecnologias de informação e comunicação, tendo um impacto significativo na representação descritiva de documentos. Essas alterações estão fazendo com que profissionais da informação sejam instigados a ter olhares diferenciados no tratamento dos recursos informacionais, e adoção de novas práticas na organização da informação, principalmente na Web.

Neste novo contexto deve-se manter os conceitos já consolidados pela CI, que conforme Mey (1995), o uso de códigos normalizados na representação descritiva é fundamental por regularizar e padronizar procedimentos em sistemas de informação, especialmente de médio e grande porte, possibilitando maior consistência, eficiência e qualidade ao sistema, ao fornecer registros coerentes entre si, fundamental para viabilizar a cooperação catalográfica entre Unidades de Informação, e que no contexto atual, seu compartilhamento na Web.

2.1 WEB SEMANTICA

A World Wide Web (Web) foi inicialmente concebida como uma construção de páginas HTML e hiperlink para atender a distribuição de conteúdo. No início ainda não havia a necessidade de seguir padronizações para compartilhamento de informações, o que inviabilizou que os conteúdos fossem interpretados pelos computadores. A ideia básica da Web Semântica (WS) é que agentes inteligentes (outros computadores) fossem capazes de inferir novos conhecimentos a partir da informação digital e suas relações semânticas. O termo Web Semântica foi cunhado por Tim Berners-Lee, James Hendler e Ora Lassila na publicação do artigo de mesmo nome, na revista *Scientific American* em maio de 2001. Neste

trabalho os autores explicam que WS não é um novo conceito de Web, mas uma extensão da atual, na qual a informação tenha um significado bem definido, permitindo que exista interação entre computadores e pessoas (BERNERS-LEE; et al., 2001).

Na proposta de Berners-Lee, Hendelrs, Larissa (2001), para que a Web Semântica funcione os computadores devem ter acesso a coleções estruturadas de informações e conjuntos de regras de inferência que eles possam usar para conduzir um raciocínio automatizado (NAKANO; et al, 2015). Para este fim é necessário contextualizar as informações contidas nestes documentos, e uma das formas é por meio dos metadados.

A W3C, desenvolvedora do projeto WS, esclarece que se trata de uma nova geração da Web, cujo intento é desenvolver tecnologias, linguagens, padrões e recomendações que tornem a informação legível também pelas máquinas, de forma que os computadores possam ter acesso a coleções estruturadas de informações (dados e metadados) e ao conjunto de regras que ajudem no processo de inferência automática, contribuindo para novas descobertas em todos os campos.

Uma das grandes vantagens do uso da WS está ligada à inferência lógica de dados, ou seja, possibilitar a afirmação de “verdade” em proposições em decorrência de sua ligação com outras proposições já reconhecidas como verdadeiras. Consiste, portanto, em derivar conclusões a partir de premissas conhecidas ou decididamente verdadeiras (NAKANO; et.al, 2015).

De forma geral o modelo de WS já está bem desenvolvido e documentado na área da computação, sendo um bom exemplo a rede de dados “*Linked Data*” ou Dados Lincados (W3C, 2006), que já integrava em 2015, 1.014 padrões de *Datasets*. Esses padrões são utilizados para representar os atributos (metadados) das diversas classes (tipos de fontes, objetos, coisas) de informações em ambiente digital, e seguem os princípios definidos pela W3C. A WS é construída totalmente em ambiente digital, onde cada informação é representada por uma única *Uniform Resource Identifiers* (URI) para cada recurso. Este endereço eletrônico deve seguir padrões abertos, norteados pelas definições do RDF (W3C, 2016).

Um de seus criadores, Berners-Lee (2006) define que a Web Semântica não é apenas sobre a colocação de dados na web, mas principalmente como fazer ligações, de modo que uma pessoa ou máquina pode explorar a web de dados. Com dados lincados, fazendo com que com uma pouca fração da informação, seja possível encontrar outros, relacionando-os. Ao contrário da web do hipertexto, onde as ligações são relações âncoras em documentos de hipertexto escritas em HTML, na WS os dados são descritos por RDF. O URI identifica

qualquer tipo de objeto (recurso) ou conceito. Com isso Berners-Lee (2006), para possibilitar seu desenvolvimento, deve-se obedecer a quatro princípios: a) usar URIs como nomes para os itens; b) usar URIs HTTP para que as pessoas possam consultar esses nomes; c) quando alguém consulta uma URI, prover informação RDF útil; e d) incluir sentenças RDFs com links para outras URIs, a fim de permitir que itens relacionados possam ser descobertos.

2.2 RDF: o fundamento da web semântica

O RDF é considerado o fundamento da Web Semântica e que está relacionado com a representação descritiva dos recursos da Web. Segundo **Allemang** e Hendler (2011), o RDF, RDFS e OWL são as linguagens básicas de representação da WS. Ainda para os autores, o RDF é um modelo para a descrição de recursos e definem como está constituído um "sistema de modelagem de dados".

Para Manola e Miller (2004), o RDF é uma forma de representação dos metadados dos recursos Web, tais como título, autor, data de modificação de uma página da Web, *copyright* e informações sobre licenças de documentos Web ou informações sobre o compartilhamento.

A W3C (2015) descreve o RDF como a Descrição de Recursos Funcionais, uma especificação desenvolvida para possibilitar a modelagem de metadados dentro dos diversos formatos já existentes (sintaxes). O formato RDF geralmente é associada ao *Extensible Markup Language* (XML), utilizando a sintaxe RDF/XML, porém pode também ser utilizados outros formatos como o RDFa, *Turtle*, *N-Triples(N3)*. Esses modelos utilizam também estruturas para descrever seus recursos, conceitos, termos e relacionamentos pois utilizam notações como o RDF *Schema* e a *Web Ontology Language* (OWL) nos diversos domínios do conhecimento (W3C, 2001; CATARINO; SOUZA, 2012).

Essencialmente, com o RDF, cada obra, autor, editor, coautor, assunto e volume tem sua própria URI, denominado de recurso, e dentro dela tem a descrição de suas classes, de forma a contextualizar a informação tanto para visualização humana como para ser processada pelos computadores. O resultado é a expansão dos pontos de acesso para que os usuários possam navegar pelo catálogo por muitos vetores diferentes, em busca da obra com base em critérios de sua própria escolha. (BRANTLEY, 2015)

De forma a representar as estruturas de informações, a WS são constituídas de triplas no modelo “Recurso, propriedade e Recurso”, o modelo é representado na Figura 1, esse modelo é chamado de *Resource Description Framework* (RDF) ou em tradução livre, Estrutura de Descrição de Recursos.

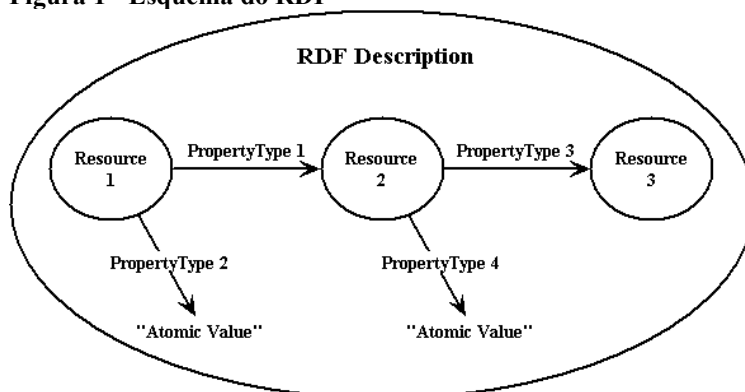
Com o objetivo de codificar, trocar e reutilizar metadados no ambiente da Web foi desenvolvido a linguagem RDF, com uma terminologia desenvolvida pela W3C, elaborada a partir do encapsulamento XML, visando prover intercâmbio de informações entre aplicações sem a perda do significado (FURGERI, 2006). Para isso, define um vocabulário padrão para descrever coisas ou objetos e permitir a troca de metadados entre aplicações. O RDF permite criar declarações sobre esses objetos por meio de propriedades que representam um relacionamento entre recursos (MEISSNER, 2004). Uma declaração é realizada por meio de triplas do tipo “recurso-propriedade-valor”, apresentado na Fig. 1, onde:

a) **Recurso** é o sujeito de uma declaração. Pode ser um website ou parte dele, ou ainda um objeto não acessível via Web (livro, CD, entre outros). Um recurso pode ser acessado e reconhecido de forma única por meio de um URI.

b) **Propriedade** é o predicado de uma declaração. Trata-se de um atributo usado para descrever um recurso. Um recurso pode conter diversas propriedades: nome do autor, título do artigo, data de publicação, entre outros.

c) **Valor** é o objeto de uma declaração. Representa o conteúdo das propriedades. Sendo os conteúdos referentes ao nome do autor, título do artigo, data de publicação, entre outros.

Figura 1 - Esquema do RDF



Fonte: Miller (1998).

A nomenclatura da descrição de “recurso”, “propriedade” e “valor” podem variar conforme o autor, mas o conceito do RDF é igual para a descrição de qualquer tripla.

Ferreira e Santos (2013) apresentam brevemente o modelo RDF como um importante mecanismo para a representação e a descrição de recursos, considerando que o entendimento do RDF é crucial ao ramo da representação descritiva, 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).

A CI já compreende os padrões de RDF e suas aplicações na WS, visto o grande número de trabalhos já publicados nas revistas e livros da área, porém ainda são poucas as tentativas de interagir e desenvolver catálogos e bases de dados com estes conceitos (RAMALHO, VIDOTTI; FUJITA ,2007). As bases bibliográficas atualmente disponíveis foram construídas no formato MARC ou Dublin Core, tendo um baixo nível semântico e interoperabilidade semântica.

Para possibilitar a consulta de dados representados em RDF, utiliza-se o SPARQL, que seria o equivalente a linguagem SQL para os bancos de dados relacionais. Sendo então, o SPARQL é um conjunto de especificações que fornecem linguagens e protocolos para consultar e manipular o conteúdo gráfico RDF na Web ou em estoques de RDF. O padrão compreende as seguintes especificações. O SPARQL possibilita que uma *Query* possa ser usada para formular consultas que variam de padrão gráfico simples de correspondência para consultas complexas envolvendo diversos estoque de RDF distribuídos na Web (SPARQL, 2013). As especificações do SPARQL atualmente estão na versão 1.1, esse protocolo é uma versão com muitas alterações substanciais e evoluções em relação a primeira SPARQL 1.0, publicada pelo W3C como linguagem de consulta em janeiro de 2008.

As especificações de cada estoque de RDF pode ser definida pelas ontologias (OWL), desta forma, uma ontologia define os termos e relacionamentos básicos comprimindo o vocabulário em metas tópicos convenientes, assim como regras para combinação de termos e relacionamentos, para que com isto possa ser definidas extensões do vocabulário (GOMÉZ-PERÉZ, 2004).

2.3 Pesquisas no Brasil

Em uma busca não exaustiva, buscou-se identificar os estudos de WS publicados no Brasil, destacam-se as pesquisas de Catarino e Souza (2012), que contém o resultado de uma pesquisa bibliográfica, cujo objetivo foi verificar como as práticas da representação descritiva estão relacionadas com a proposta de organização dos recursos da Web no contexto da Web Semântica. Os autores consideram que a transformação dos catálogos das bibliotecas e de outras bases de dados para Dados Lincados permitirá uma infinidade de aplicações e facilidades para o usuário na busca de informações, cita como algumas instituições como LC e DCMI, já estão realizando algumas ações iniciais, mas é imprescindível que outros estudos

sejam realizados, tanto por pesquisadores da área de Ciência da Informação (em especial a Organização da Informação) quanto do W3C, de forma a propor ações sistematizadas e globais para que os dados bibliográficos possam estar inseridos no conjunto dos Dados Lincados.

Em outro estudo, Santarem Segundo (2014) descreve a falta de publicações na área de Ciência da Informação que relatem implementações e uso do SPARQL, concluindo que a falta de estudos pode estar associada a dificuldade em entender como pode ser realizado o processo de experimentação do protocolo.

Em uma revisão de literatura sobre o tema Web Semântica sob a égide da Ciência da Informação, Nakano e outros autores (2015) concluíram que em sua maioria, os estudos realizados na área de CI ainda estão focados nos estudos teórico e filosóficos a fim de nortear o planejamento, execução e padronização dos sistemas, ou seja, trabalhar de forma colaborativa e em conjunto com outras áreas do conhecimento.

O único estudo aplicado identificado na literatura brasileira foi publicado por Monteiro e Jacyntho (2016) da Ciência da Computação, os autores estudaram o uso da Web Semântica e dos princípios Dados Lincados propostos pelo (W3C) no desenvolvimento de uma aplicação Web de gestão semântica de documentos digitalizados. Porém a aplicação está em fase inicial, não tendo conclusões sobre a vantagem de uso na recuperação de informação pela limitação de utilização da linguagem SPARQL.

Para a execução do aqui proposta, foi utilizado a Brapci com objeto de estudo. A Brapci é o produto de informação do projeto de pesquisa Opções metodológicas em pesquisa: a contribuição da área da informação para a produção de saberes no ensino superior, 1 cujo objetivo é “subsidiar estudos e propostas na área de Ciência da Informação, fundamentando-se em atividades planejadas institucionalmente”. A Brapci ressalta que o propósito do projeto é contribuir “para estudos analíticos e descritivos sobre a produção editorial de uma área em desenvolvimento, ao subsidiar, com uma ferramenta dinâmica, os alunos, professores e pesquisadores da área” (BRAPCI, 2015).

3 METODOLOGIA

A pesquisa foi iniciada por uma busca exploratória nas bases de dados Scopus, *Web of Science*, Brapci, Google Scholar na localização de RDFschema, ontologias e classes para

representar artigos científicos no modelo RDF. Não foi possível eleger uma única estratégia de busca, havendo a necessidade de realizar diversas combinações com os termos “OWL”, “RDF”, “bibliography” e “Web Semantic” e suas variações em português, a fim de localizar trabalhos sobre as práticas.

O resultado da pesquisa não apresentou um conjunto de *datasets* que pudessem representar o modelo de dados da Brapci, considerou-se que a base da estrutura seria empregada com o uso de vocabulários para descrição estruturada da representação descritiva bibliográfica, comunicação científica, e de autoridades. Optou-se, baseado na proposta de Hladka, Mynarz e Sklenak (2012), em combinar os vários *datasets* para encontrar as descrições mais adequadas para os dados.

Eleitos as classes (*datasets*) para cada elemento foi desenvolvido um programa em linguagem PHP que acesse os dados diretamente do sistema gerenciador de banco de dados (SGBD) gerando um único arquivo no formato de triplas (n3). Ao executar o programa, o sistema gerou os dados dos 15.882 trabalhos de revistas e 1.832 trabalhos de eventos disponibilizando o título, resumo, palavras-chave, autores, nome da publicação, dados sobre a edição, e das metodologias utilizadas na pesquisa, sendo este último disponível em apenas 5% do total de artigos indexados. O arquivo final gerado tem o tamanho aproximado de 27MBytes e pode ser acessado no link <<http://www.brapci.ufpr.br/brapci/brapci.n3>>.

O arquivo com as triplas foi importado para o programa *Graph Database Free* (GraphDB Free) desenvolvido pela Ontotext na Bulgária, o software leve feito em Java e com uma interface acessível foi instalado em um computador pessoal, possibilitando a consulta ao estoque RDF constituído. Os dados exportados pela Brapci para o software criaram uma base semântica denominada de BrapciRDF, com base nela foram construídas *queries* de consulta em SPARQL³ objetivando a obtenção das melhores respostas as perguntas a *priori* formuladas. Uma das dificuldades foi compreender a linguagem de consulta SPARQL.

Para validar o funcionamento da base BrapciRDF em consultas SPARQL foi utilizado o estudo de Freire (2012) com o tema “responsabilidade social”, que utilizou a Brapci como fonte de dados, comparando com os resultados obtidos pelas consultas, os resultados foram transcritos para análise. Após a validação da base de teste foi aplicado as *queries* de consulta de forma a tentar responder as cinco questões previamente propostas. Os dados obtidos foram transcritos nos resultados da pesquisa.

³ SPARQL é uma linguagem de consulta semântica para bancos de dados capaz de recuperar e manipular dados armazenados em *Resource Description Framework*.

4 RESULTADOS

Para gerar o arquivo das triplas foram eleitas as ontologias recorrentes na literatura para representação descritiva de artigos científicos e anais de eventos. Ressaltando que esta primeira proposta de geração de triplas, os dados foram linkados com as fontes da própria Brapci, não existindo elos externos na Web ou ao Dados Linkados, pois este processo necessita de uma reestruturação interna da modelagem da Brapci. O que foi realizado foi somente a exportação dos dados já existente (de forma relacional) em triplas (formato n3). A **Tabela 1** apresenta as ontologias utilizadas, os prefixos da descrição das classes, o nome da ontologia e o link de sua documentação. Não foi localizado na literatura uma ontologia que representasse as opções metodologias das pesquisas, desta forma foi construída uma classe utilizando o software *Protégê* e incorporado ao padrão das triplas. Esta ontologia foi desenvolvida baseada nos resultados dos estudos de Freitas (2012) e do projeto de pesquisa produtividade de Bufrem 2013/2016, que tem como objetivo as análises das opções metodológicas utilizadas nos estudos da CI no Brasil.

Tabela 1 - Ontologias para representação dos dados da Brapci

Prefix	Nome	URL
bibo:	Bibliographic Ontology Specification	http://purl.org/ontology/bibo/
brapci:	Opções metodológicas	http://www.brapci.ufpr.br/rdf/schema/brapci/
dc:	Doublin Core	http://purl.org/dc/terms/
dcterms:	DCMI Metadata Terms	http://purl.org/dc/terms/
foaf:	Friend of a Friend	http://xmlns.com/foaf/0.1/
frbr:	Functional. Requirements for Bibliographic Records	http://vocab.org/frbr/core
rdf:	Resource Description Framework	http://www.w3.org/1999/02/22-rdf-syntax-ns#
time:	Date and Time	http://www.w3.org/2006/time#

Fonte: autor (2016).

A Tabela 1 demonstra que a representação bibliográfica já está contempladas nos modelos RDF, ressaltando ainda que outras descrições podem ser incorporadas a proposta, como o modelo SKOS, que já foram descritas. Observa-se ainda nessa tabela a possibilidade de incorporação de novas descrições de ontologias (ou classes) como foi proposta (brapci:).

Com a importação dos dados do arquivo no software GraphoDB Free foi necessário validar a base comparando com um estudo já realizado que utilizou o mesmo *corpus*. Desta forma foi identificado o trabalho de Freire (2013) apresentado no Enancib com o tema de “responsabilidade social”.

Em sua pesquisa, Freire (2013) utilizou como estratégia de busca os termos “responsabilidade” e “social” separados, e a busca foi realizada no título, palavras-chave e

resumo entre o período de 1972 e 2013. Para atender essa mesma estratégia foi criado e executado a *query* representada no Quadro 1.

Quadro 1 - Consulta SPARQL para o termo “responsabilidade” e “social”

```
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX dct: <http://purl.org/dc/terms/>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX foaf: <http://xmlns.com/foaf/0.1/>
PREFIX brapci: < http://www.brapci.ufpr.br/brapci/index.php/rdf/schema/brapci/>
PREFIX dc: <http://purl.org/dc/terms/>
PREFIX bibo: <http://purl.org/ontology/bibo/>
select DISTINCT ?titulo ?id {
  ?id a bibo:AcademicArticle .
  ?id dc:title ?titulo .
  ?id ?b ?key .
  ?id time:year ?ano .
FILTER (regex(str(?key), "responsabilidade", "i")
        && regex(str(?key), "social", "i")
        && ?ano < "2014") .
}
```

Em tradução da *query* para ser compreendida por seres humanos, podemos descrever que ela solicita ao sistema SPARQL que recupera a identificação (id) de todos os Artigos Acadêmicos (bibo:AcademicArticle) e armazenar na variável de mesmo nome “?id”, em uma nova rodada de consulta, o SPARQL recupera, conforme padrão Dublin Core, o título do Artigo Acadêmico (dc:title) e guarda na variável “?titulo”, na última sequência lógica de consulta, o sistema recupera todas as palavras-chave (?key) utilizadas nos Artigos. Identificados todos os trabalhos, com o ID, título e palavras-chave faz-se necessário a aplicação de filtros para delimitar o universo, para isso utiliza-se os parâmetros “FILTER” cujos trabalhos contêm nas palavras-chave (?key) as ocorrências dos termos “responsabilidade” e “social” em qualquer parte das palavras-chave, e o resultado apresentado será o título e o ID do trabalho (?titulo ?id). Para equiparar os resultados com a pesquisa de Freire (2013) foi necessário limitar a busca até 2013, com a inclusão de um novo elemento na consulta que retorna o ano da publicação (time:year) e de um FILTER para os trabalhos publicados antes de 2014.

O resultado da busca se mostrou mais efetivo do que o realizado pela autora. Freire recuperou em sua busca 78 trabalhos sobre o tema, enquanto com a mesma estratégia usando o SPARQL foram identificados 103 trabalhos. O mesmo ocorreu com o número de autores, 50 do artigo original e 116 deste estudo, e o total de periódicos, 15 do artigo original e 16 deste estudo. Algumas diferenças podem ser justificadas incorporação de novos artigos após a pesquisa da autora, ou ainda pelas revistas ainda não terem disponibilizados os trabalhos no período da coleta de dados. Entretanto os resultados se mostraram mais abrangentes, com

maior quantidade de documentos recuperados, e muito próximos, considerando que a consulta SPARQL consegue responder exaustivamente. Os resultados são apresentados na Tabela 2.

Tabela 2 - Comparação entre o estudo de Freire (2013) e consulta de SPARQL

Item de análise	FREIRE (2013)	SPARQL
Total de trabalhos recuperados	78	103
Total de periódicos localizados (entre 1999-2012)	15	16
Autores que publicaram sobre o tema	50	116

Esta comparação ajustou a validar o uso das triplas na recuperação de informação, e por consequência o andamento do estudo. Ainda utilizando o mesmo estoque de RDF foram então formuladas as *queries* de consultas objetivando responder as cinco perguntas *a priori* formuladas.

A primeira pergunta a ser respondida pela BrapciRDF é “Quais autores publicaram trabalhos em 1972 e na década de 2010?”, essa pergunta busca identificar os autores pioneiros das revistas da Ciência da Informação brasileira em seu primeiro ano e que ainda publicaram algum trabalho na década de 2010. Um estudo deste tipo demandaria a identificação dos autores que publicaram em 1972, e comparar com os autores que publicaram entre os anos de 2000 e 2009, ressaltando que a partir de 2005, houve um avanço das publicações eletrônicas, fazendo com o número de autores aumentasse consideravelmente nesta década.

A construção da *query* de consulta demandou um tempo até a compreensão da lógica de funcionamento do SPARQL, porém depois de assimilado sua lógica, o resultado da consulta bem rápido. O sistema retornou o nome de quatro autores nesta busca, com base nesta informação foi realizado uma consulta na base tradicional da Brapci para melhor contextualizar esses resultados. Identificou-se que o sistema respondeu adequadamente, foram identificados: Fonseca, que publicou suas memórias na Revista Ibero-Americana no terceiro volume da revista em 2010; Polke, cujo trabalho remete a uma entrevista realizada com a profa., transcrita e publicada na revista Perspectiva em Ciência da Informação em 2010. Zaher, que fez a apresentação do XI Enancib em Brasília, e a Hagar Espanha Gomes, que tem trabalhos publicados nesses quase 45 anos da CI no Brasil, contribuindo com mais de três trabalhos nesta década. O Quadro 2 apresenta esses resultados, e Quadro 3 a *query* de consulta utilizada para esta recuperação.

Quadro 2 - Autores que publicaram em 1972 e na década de 2010

Autor	Comentário
FONSECA, Edson Nery da	Memórias publicadas na Revista Ibero-Americana em 2010.
GOMES, Hagar Espanha	Três publicações indexada na Brapci na década de 2010.
POLKE, Ana Maria Athayde	Publicada entrevista com a profa. Polke em 2010 na Perspectiva
ZAHER, Célia Ribeiro	Publicou a Apresentação do XI Enancib na Revista Liinc em 2011

Quadro 3 - Query de consulta para identificar os autores que publicaram em 1972 e na década 2010

```
SELECT DISTINCT ?nome {
  ?a a bibo:AcademicArticle .
  ?a dc:creator ?autorc .
  ?a time:year ?ano .
  ?autorc foaf:name ?nome .
  ?a2 dc:creator ?autorc2 .
  ?a2 time:year ?ano2 .
  FILTER(?autorc = ?autorc2)
  FILTER(?ano >= "1972" && ?ano <= "1972" && ?ano2 >= "2010" && ?ano2 <= "2020" && ?ano != ?ano2) .
}
```

A implicações lógicas na montagem desta consulta, demonstra em parte o desafio e as potencialidades desde novo ambiente de pesquisa.

Ao tentar responder a segunda questão empregando os métodos tradicionais, será necessário a identificação inicialmente dos pesquisadores que são coautores da pesquisadora, e posteriormente analisar todos os trabalhos e identificar os correspondentes a opção metodológica que se busca. Desta forma ao tentar responder a questão “Quem são os colaboradores de Leilah Santiago Bufrem nos estudos descritivos? ”, permite criar relações lógica facilitando a recuperação dos informação.

A linguagem SPARQL permite a combinação de elementos em uma mesma consulta, podendo vincular dados como posição geográfica, instituição, descritor, afiliação institucional, data em uma única consulta. Esse procedimento foi realizado para a responder esta pergunta, combinando o autor pesquisado, com seus colaboradores e com opção metodológica utilizada, sendo que cada informação era buscada na propriedade correspondente (ex, autor, opção metodológica). No processo tradicional a busca deveria ser feita em todos os campos, retornando documentos muitas vezes não importantes. O Quadro 4 apresenta a query de consulta e o Quadro 5 os resultados obtidos.

Quadro 4 - Query SPAQRL para identidicar os colabores de Bufrem nos estudos descritivos

```
select DISTINCT ?nome ?nome_colaborador ? titulo where {
  ?a a bibo:AcademicArticle .
  ?a dc:creator ?autorc .
  ?a brapci:metodology ?opcoes .
  ?a dc:title ?titulo .
  ?opcoes brapci:name ?opcoes_nome .
  ?autorc foaf:name ?nome .
  ?a dc:creator ?colaboradores .
  ?colaboradores foaf:name ?nome_colaborador .
  FILTER (regex(str(?nome),"Bufrem","i") && regex(str(?opcoes_nome),"descritiva","i"))
  FILTER (?nome != ?nome_colaborador)
}
```

O quadro 4 apresenta a forma de consulta, observa-se que existe uma complexidade maior na estrutura, porém o sistema responde de forma eficiente e com um tempo de resposta muito baixo, mesmo pesquisando nos mais de 12.000 registros bibliográficos

disponibilizados. Foram utilizados os filtros de nome do autor, e das opções metodológicas, podendo criar outras variações de forma simples.

Quadro 5 - Autores que colaboraram com Bufrem nos estudos descritivos

	Nome do Coautor	Título do trabalho*
1	BREDA, Sônia Maria	Ética e formação profissional: uma leitura da produção científica em Ciência da Informação (1970-2006)
2	SORRIBAS, Tidra Viana	Ética e formação profissional: uma leitura da produção científica em Ciência da Informação (1970-2006)
3	FREITAS, Juliana Lazzarotto	Ética e formação profissional: uma leitura da produção científica em Ciência da Informação (1970-2006)
4	ARAUJO, Ana Carolina Arantes	Informação para negócios: aspectos da literatura científica nacional em revistas da área de ciência da informação
5	PRATES, Yara	O saber científico registrado e as práticas de mensuração da informação
6	FREITAS, Juliana Lazzarotto	Configuração epistemológica da Ciência da Informação na literatura periódica brasileira por meio de análise de citações (1972-2008)
7	ARBOIT, Aline Elis	Configuração epistemológica da Ciência da Informação na literatura periódica brasileira por meio de análise de citações (1972-2008)
8	GABRIEL JUNIOR, Rene Faustino	Práticas de coautoria no processo de comunicação científica na pós-graduação em Ciência da Informação no Brasil
9	GONCALVES, Viviane	Práticas de coautoria no processo de comunicação científica na pós-graduação em Ciência da Informação no Brasil

* Os resultados apresentados referem-se somente aos trabalhos já categorizados na base, que corresponde a 5% do total de trabalhos indexados.

O número de trabalhos pelas metodologias está diretamente relacionado ao total de trabalhos já categorizados pelas metodologias, que hoje estão em menos de 5% de toda base.

Ainda utilizando o tema de opções metodológicas, suscitou a curiosidade de saber quais as técnicas, tipos de estudos e metodologias que Suzana Mueller utilizou em suas pesquisas. Com o SPARQL é possível realizar esse tipo de consulta, como representado na Quadro 6 e os resultados apresentado na Tabela 3. A consulta possibilita gerar totalizadores de recursos, propriedades ou valores.

Tabela 3 - Opções metodológicas de MUELLER, Suzana Pinheiro Machado

Opções metodológicas	Total
Análise ou Projeto de Sistema ou Programa	1
Explicativa	1
Pesquisa Documental	1
Estudos Métricos de Informação	2
Exploratória	2
Análise Documental	3
Descritiva	3
Pesquisa Bibliográfica	3

Observa-se que a Suzana Mueller trabalha muito com a pesquisa bibliográfica, análise documental e pesquisas descritivas, tendo um foco também nos estudos métricos de informação.

Quadro 6 - Query SPARQL para identificar as opções metodológicas de um autor

```

select DISTINCT (count( ?opcoes_nome ) as ?total) ?opcoes_nome ?total where {

  ?a a bibo:AcademicArticle .
  ?a dc:creator ?autorc .
  ?a brapci:metodology ?opcoes .
  ?opcoes brapci:name ?opcoes_nome .
  ?autorc foaf:name ?nome .
  FILTER (?nome = "MUELLER, Suzana Pinheiro Machado")

}

```

A combinação das várias ontologias e a seleção das diversas fontes de consultas são uma das grandes vantagens da linguagem SPARQL para RDF, porém a complexidade da consulta depende muito da prática de quem está criando as consultas para que o RDF possa responder.

A próxima questão tem como tema os estudos de caso em uma temática específica, formulada como “Quais os estudos de caso foram aplicados em bibliotecas?” A consulta foi montada de forma a sistema localizar o “estudo de caso” como opção metodológica e identificar o termo “biblioteca” nas palavras-chave, foi executado a query apresentada no Quadro 6.

Quadro - 6 Query de consulta para estudo de caso em biblioteca.

```

select ?title ?key where {
  ?a a bibo:AcademicArticle .
  ?a dc:title ?title .
  ?a dc:term ?key .
  ?a brapci:metodology ?opcoes .
  ?opcoes brapci:name ?opcoes_nome .
  FILTER (regex(str(?opcoes_nome), "Estudo de Caso", "i") && regex(str(?key), "Biblioteca", "i") ) .
}

```

Parte dos trabalhos recuperados é apresentado na Tabela 4, A quantidade de trabalhos foi de 23, porém como o objetivo deste estudo foi apenas demonstrar o recurso, na Tabela são apresentados o resultado dos primeiros oito registros de um total de 44 registros identificados com esta estratégia.

Tabela 3 - Estudos de casos aplicados a bibliotecas

#	Título do trabalho	Opção metodológica	Palavra-chave
1	Educação contínua: atitudes e experiências dos bibliotecários do Sistema de bibliotecas da UNICAMP	Estudo de Caso	"Biblioteca universitária"@pt-br
2	Avaliação dos processos de automação em bibliotecas universitárias	Estudo de Caso	"Biblioteca universitária"@pt-br
3	Serviço ARUANDA: uma introdução ao sistema	Estudo de Caso	"Biblioteca universitária"@pt-br
4	Technology and Research in a Global Networked University Digital Library (NUDL)	Estudo de Caso	"biblioteca digital"@pt-br
5	Centro de informação jurídica eletrônico e virtual	Estudo de Caso	"biblioteca digital"@pt-br
6	A universalização da informação	Estudo de	"Biblioteca"@pt-br

		Caso	
7	Aspectos que interferem na construção da acessibilidade em bibliotecas universitárias	Estudo de Caso	"Biblioteca universitária"@pt-br
8	Integração e interoperabilidade no acesso a recursos informacionais eletrônicos em C&T: a proposta da Biblioteca Digital Brasileira (existem mais registros não apresentados)	Estudo de Caso	"biblioteca digital"@pt-br

Pode-se observar que a busca foi realizada de forma léxica, fazendo com que o conteúdo recuperado muitas vezes não seja pertinente ao tema desejado, no futuro muito próximo, poderemos realizar esta mesma consulta, não mais de forma léxica, mas de forma semântica, possibilitando a identificação do tipo de biblioteca que queremos recuperar por URIs.

Para responder a última pergunta proposta, foi utilizada consulta do Quadro 7, respondendo “Quais os temas mais publicados entre 2010 e 2015?”, esse tipo de pergunta não foi possível responder com a base BrapciRDF sem a carga semântica e vocabulários controlados nas descrições dos termos que representam os conteúdos. Esta certamente será a contribuição da CI na construção de vocabulários dentro dos domínios, de forma a remover a ambiguidade dos descritores, aumentando a carga semântica dentro de um domínio. A query construída apresentou os mesmos resultados da consulta realizada no banco de dados SQL.

Tabela 5 - Frequência dos descritores entre os trabalhos de 2010 e 2015.

Descritor	Frequência	Descritor	Frequência
Biblioteconomia	321	Comunicação científica	63
Gestão do conhecimento	128	Tecnologia da informação	54
Comunicação	112	Memória	53
Informação	107	Produção científica	47
Gestão da informação	97	Estudo de usuários	43
Arquivologia	80	Ontologia	41
Bibliometria	72	Redes sociais	40

Fonte: Dados da pesquisa (2016)

Quadro - 7 Query de consulta para as palavras-chave com maior frequência

```

select ?key ( count(?key) as ?total) ?total where {
    ?a a bibo:AcademicArticle .
    ?a dc:term ?key .
    ?a time:year ?ano .
    FILTER ($ano >= "2010" && $ano <= "2015")
}

```

5 CONSIDERAÇÕES FINAIS

O trabalho apresentou os resultados de um estudo aplicado utilizando o RDF sobre os registros bibliográficos da base Brapci, observou-se que a CI já tem uma boa compreensão sobre os conceitos da Web Semântica, RDF e do modelo conceitual FRBR para aplicações na

área. Porém as maiores dificuldades estão relacionadas as tecnologias, linguagens de programação e desenvolvimento de sistemas, domínios relacionados a Ciência da Computação (CC). Observa-se que está existindo uma aproximação da CI e da CC, principalmente pelo número de trabalhos dessa área publicados nas revistas de CI, com foco no desenvolvimento de novos sistemas de informação baseados nos conceitos da Web Semântica, havendo uma aproximação entre as áreas e por consequência a complementação dos saberes em prol de um objetivo comum, muito benéfico para ambas as áreas.

O trabalho partiu de perguntas básicas que os investigadores podem obter de seu *corpus* de pesquisa, e que se a WS teria condições de respondê-las. Os mesmos resultados poderiam ser obtidos no processo tradicional, porém empregando métodos mais trabalhosos e que demandam maior consumo de tempo.

Dos objetivos propostos neste estudo, considera-se que foram atingidos satisfatoriamente, desde a identificação de um referencial teórico que oferecesse a fundamentação para construção de um modelo baseado em RDF, como na escolha dos *datasets* compatíveis com a aplicação (artigos científicos) até a extração dos resultados do *corpus* estudado. A maior dificuldade da pesquisa foi dominar a linguagem SPARQL, que é pouco amigável, e possibilita múltiplas combinações, variáveis e filtros quando aplicada nas triplas: recurso, propriedade e valor.

No Brasil as aplicações da WS, RDF e FRBR, principalmente no meio acadêmico, ainda são incipientes e são voltadas a estudos de cunho teóricos e filosóficos. Desta forma, esse estudo foi um dos pioneiros na aplicação da WS em estudos métricos da informação em âmbito acadêmico. Dos resultados obtidos, o estudo demonstrou-se promissor, mas considera ainda necessário uma melhor carga semântica da fonte de informação, com a utilização de vocabulários controlados e controles de autoridades que garantam precisão na recuperação da informação, outra dificuldade encontrada refere-se a interface de consulta do SPARQL tenha melhor usabilidade, pois a compreensão de suas *queries* é de difícil compreensão para a maioria dos profissionais de informação e os usuários dessas bases.

REFERÊNCIAS

- ALLEMANG, D.; HENDLER, J. **Semantic web for the working ontologist: effective modeling in RDFS and OWL**. Amsterdam: Morgan Kaufmann, 2011.
- BERNERS-LEE, T. Linked Data, 2006. Disponível em: <<https://www.w3.org/DesignIssues/LinkedData.html>>. Acesso em 4 de ago. 2016.
- BERNERS-LEE, T.; HENDLER, J.; LASSILA, O. The semantic web. **Scientific American**, v.284, n.5, p.34-43, 2001.
- BOICA, A. L.; OLIVEIRA, L. H. M. de. Conversão de metadados do padrão Dublin Core para o RDF. **Global Science and Technology**, v. 1, n. 2, p. 8-13, dez./mar. 2008.
- BRANTLEY, P. Fun with metadata: with MARC showing its age, new formats such as RDF show promise. **Publishers Weekly**, v. 15 n. 54, 2015.
- CATARINO, M. E.; SOUZA, T. B. A representação descritiva no contexto da web semântica. **Transinformação**, v. 24, n. 2, p. 77-90, 2012. DOI: 10.1590/S0103-37862012000200001.
- FERREIRA, J. A.; SANTOS, P. C. L. V. A. C. O modelo de dados resource description framework (rdf) e o seu papel na descrição de recursos. **Informação & Sociedade: Estudos**, v. 23, n. 2, p. 13-23, 2013.
- FREIRE, I. M. Temática <responsabilidade social> na literatura da ciência da informação indexada pela Brapci. **Anais...**, ENANCIB, XIV. Encontro Nacional de Pesquisa em Ciência da Informação. Florianópolis, 2013.
- FREITAS, J. L. **Opções metodológicas em pesquisas na área de ciência da informação: contribuições a uma análise de domínio**. Dissertação... (Mestrado em Ciência, Gestão e Tecnologia da Informação) Universidade Federal do Paraná, Curitiba, 2012.
- FURGERI, S. R. O papel das linguagens de marcação para a ciência da informação. **Transinformação**, v. 18, n. 3, p. 225-239, 2006.
- FUSCO, E. **Aplicação dos FRBR na modelagem de catálogos bibliográficos digitais**. São Paulo: Cultura Acadêmica, 2011.
- HLADKA, J.; MYNARZ, J.; SKLENAK, V. Experience with transformation of bibliographic data into linked data. **Journal of Systems Integration**, vol. 3, n. 1, 2012. DOI: 10.20470/jsi.v3i1.106.
- LIBRARY OF CONGRESS. Bibliographic Framework as a Web of Data: Linked Data Model and Supporting Services. Washington, 2012. Disponível em: <<https://www.loc.gov/bibframe/pdf/marclid-report-11-21-2012.pdf>>. Acesso em: 5 ago. 2016.
- MEY, E.S.A. **Introdução à catalogação**. Brasília: Briquet de Lemos, 1995.
- MILLER, E. An Introduction to the Resource Description Framework. **D-Lib Magazine**, may 1998.
- MONTEIRO, L. L. P.; JACYNTHO, M. D. A. Dados ligados. digitalização. documentos digitais. metadados. rdf. web semântica. **Transinformação**, v. 28, n. 2, p. 241-251, 2016.10.1590/2318-08892016000200010. DOI:10.1590/2318-08892016000200010.
- MORENO, F. P.; ARELLANO, M. N. M. R. Requisitos funcionais para registros bibliográficos - FRBR: uma apresentação. *Revista Digital de Biblioteconomia & Ciência da Informação*, v. 3, n. 1, p. 20-38, 2005.
- MORENO, F. P.; BRASCHER, M. MARC, MARCXML E FRBR: relações encontradas na

literatura. **Informação e Sociedade**: Estudos, v.17, n.3, p.13-25, set./dez. 2007.

NAKANO, N. et al. Web semântica no repositório Brapci: uma revisão da literatura. **Brazilian Journal of Information Science**, v. 9, n. 2, p. 26-41, 2015.

PINHO, F. B. A.; NASCIMENTO, B. L. S. C.; MELO, W. L. As dimensões ôntica, epistêmia e documental na representação da informação e do conhecimento. **Revista ACB: Biblioteconomia em Santa Catarina**, v. 20, n. 1, p. 112-123, 2015.

ROWLEY, J. **Informática para bibliotecas**. 3.ed. Brasília: Briquet de Lemos/Livros, 1994.

SANTAREM-SEGUNDO, J. E. Web semântica: introdução a recuperação de dados usando sparql. **Anais...**, ENANCIB, XV. Encontro Nacional de Pesquisa em Ciência da Informação. Belo Horizonte, 2014.

SPARQL Sparql 1.1 Overview. W3C, 2013. Disponível em
<<http://www.w3.org/TR/2013/REC-sparql11-overview-20130321/>> Acesso em: 13 jul. 2016

WORD WIDE WEB CONSORTIUM (W3C) W3C semantic web activity. 2001. Disponível em: <<http://www.w3.org/2001/sw/>>. Acesso em: 2 ago. 2016a.

WORD WIDE WEB CONSORTIUM (W3C). Resource Description Framework (RDF). Disponível em: <<https://www.w3.org/RDF/>>. Acesso em: 2 ago. 2016b.