



Perspectivas de investigação

Uma investigação sobre o uso de vocabulários e formalismos: em busca de um caminho para representação semântica de patrimônio digital e ativos de informação cultural

Claudio Jose Silva Ribeiro

Universidade Federal do Estado do Rio de Janeiro
Brasil · claudio.ribeiro@unirio.br

Resumo: Os temas explosão informacional e a avalanche de dados têm trazido reflexos nas discussões e debates sobre a gestão da informação na atualidade. Estes reflexos decorrem não apenas pela variedade de informações que são incorporadas diariamente, mas também pelo alto volume de dados que podem ser coletados e publicados de forma imediata. Com a proposta da Web Semântica, formulada por Tim Bernes-Lee há dez anos, surgiram algumas alternativas para representação e para efetuar correlações entre as informações na Web. Estas alternativas tornam disponíveis um quantitativo cada vez maior de dados e informação, incrementando a capacidade de criar novos conhecimentos a partir das conexões efetuadas. O patrimônio digital e os demais ativos de informação sobre a cultura estão começando a fazer parte desta coleção, pois bibliotecas e museus estão publicando seus objetos e unidades documentais na Web, trazendo outros desafios para o profissional da Ciência da Informação. Este relato de pesquisa discute o uso de vocabulários e formalismos, pois estes possuem um papel essencial no processo de representação, permitindo desenvolver estas representações tanto no nível das ideias quanto no nível da sintaxe. Reúne propostas de representação já disponíveis em bibliotecas e museus, que permitem a correlação de uma grande variedade de informações a partir da junção de conjuntos de dados publicados na Web. Apresenta a abordagem de Big Data, como um caminho que viabiliza esta junção, bem como permite a criação de ligações nunca antes pensadas pelos profissionais de Ciência da Informação destas instituições. Por fim, insere uma coleção de indagações sobre a capacidade de representação do patrimônio cultural, utilizando os fundamentos da Web Semântica e dos argumentos, propriedades, silogismos e inferências disponíveis na lógica atual de representação.

Palavras-chave: Formalismos; Web Semântica; Patrimônio Digital; Ativo de Informação Cultural; Representação do Conhecimento; Big Data.

Abstract: The information explosion and data deluge themes have had an impact on discussions and debates regarding contemporary information management. These impacts are due not only to the variety of information incorporated on a daily basis but also the great volume of data that can be collected and published immediately. Following the proposal of the Semantic Web, formulated by Tim Berners Lee ten year ago, some alternatives have appeared for representing and correlating information on the Web. These alternatives make an increasingly large quantity of data and information available, enhancing the capacity to create new knowledge from the connections undertaken. Digital heritage and other cultural information assets are becoming part of this collection, given that libraries and museums are publishing their objects and documental units on the Web, thus presenting Information Science professionals with fresh challenges. This research report discusses the use of vocabularies and formalisms, as they play an essential role in the representation process by enabling these representations to be developed at both the ideas and syntax levels. It congregates proposals for representation that are already used in libraries and museums, and which allow the correlation of a great variety of information by joining sets of data published on the Web. The Big Data approach is used to make the latter possible as well as to enable the creation of links that Information Science professionals from these institutions could never have previously

imagined. Finally, it examines a series of questions regarding the ability to represent cultural heritage using the foundations concepts of the Semantic Web and the arguments, properties, syllogisms and inferences available in the contemporary logic of representation.

Keywords: Formalisms; Semantic Web; Digital Heritage; Cultural Information Assets; Knowledge Representation; Big Data.

Introdução

A sociedade contemporânea está vivenciando um conjunto de transformações que estão sendo impulsionadas não apenas pelo desenvolvimento da tecnologia, mas também por novos pressupostos na economia e na política social do planeta. A informação passou a ser um bem que está cada vez mais se tornando um insumo para a geração de riquezas, pois a sua utilização em tempo exato está permitindo que os usuários e gestores possam angariar vantagens com o seu uso. (Ribeiro, 2008). Mayer-Schonberger e Cukier (2013) observam nos capítulos introdutórios de seu livro que na atualidade estamos vivenciando o momento onde os "dados falam" e que precisamos conseguir reunir todos estes insumos capazes de gerar conhecimento. Assim, pode-se afirmar que estes aspectos vêm ocasionando mudanças no processo de gestão da informação.

Novos elementos são incorporados cotidianamente ao cenário de gestão da informação. O uso frequente de dispositivos móveis tem gerado uma onda que vem inundando a sociedade. É possível observar que Celulares, Tablets, aparelhos sensores, além do forte uso de localização por GPS, atuam como geradores e consumidores de dados e informação. Aplicações para uso de taxis, aplicações para compartilhamento de viagens em automóveis e para evitar tráfego nas cidades, automação de aparelhos eletrodomésticos como geladeiras, microondas e aparelhos de som, fazem com que os usuários de informação assumam o papel tanto de consumidores quanto de produtores de informação. (Ribeiro, 2014a).

Todo este cenário vem trazendo reflexo na atuação do profissional que trabalha com representação e recuperação da informação. O usuário final, que antes apenas consumia informação com algum grau de interação, passou a procurar novos caminhos para interagir, buscando informações em outras partes do mundo e fazendo conexões com estas informações.

Museus, bibliotecas e instituições encarregadas de gerir arquivos públicos e privados, têm procurado caminhos para tornar as suas coleções cada vez mais inseridas neste contexto. Pode-se afirmar que interligar dados e informações de Patrimônio Cultural que estão disponíveis na Web contemporânea não é o principal problema. Archer (2014) convalida esta questão quando observa em sua apresentação que na atualidade não existem problemas para fazer ligações entre os conjuntos de informação disponíveis na Internet. Archer continua e observa que os esforços desenvolvidos em harvesting de metadados, fazendo o uso do padrão *Dublin Core*, além da possibilidade de promover conexões por meio de links externos usando o formalismo RDF (*Resource Description Framework*), tornam possível a conexão entre unidades documentárias de bibliotecas, museus e arquivos – nomeados neste relato de ativos de informação cultural - em forma digital e com a apresentação em *Web Browser* em tempo real. Baca e Gill (2015) também endossam esta percepção, quando apontam a tendência existente no ambiente de informação em cultura, para unificar sistemas *online* para buscas cruzadas com o uso de repositórios e coleções heterogêneas.

Dentro desse contexto, a pergunta de partida que impulsionou o desenvolvimento desta investigação foi: "Dado o volume e a variedade de dados, informações e objetos sobre cultura presentes para a sociedade na atualidade, os vocabulários e formalismos disponíveis na Web, são suficientes para descrever estes ativos de informação cultural?"

Assim, este relato reúne propostas e discute o uso destes elementos que possuem um papel essencial no processo de representação. Ademais, apoiado na estruturação de projetos de *Big Data*, este ensaio propõe uma jornada que poderá auxiliar no

tratamento das múltiplas fontes envolvidas no cenário das informações sobre cultura na atualidade.

Avalanche da dados e informação como catalisadora do processo de transformação

Estamos imersos em um mar de registros sobre todas as nossas ações. A Internet das Coisas (IoT – *Internet Of Things*) é uma realidade e os aparatos tecnológicos produzem algumas dezenas de Zettabytes¹ de informação, pois pessoas, instituições e dispositivos se comunicam, gerando e capturando dados que são catalogados, indexados, armazenados e disponibilizados na nuvem. (Ribeiro, 2014a).

Adicionalmente, tanto os governos quanto a sociedade organizada vêm se estruturando na disseminação de informações para consumo pelos indivíduos. São registros sobre o desempenho da educação, serviços ao cidadão, indicadores de saúde e bem estar, investimentos, economia, impostos, despesas, dentre outros, que incrementam a ideia de avalanche de dados. (Ribeiro & Pereira, 2015).

As instituições encarregadas de organizar e gerir o conhecimento produzido pela humanidade, notadamente Museus, Bibliotecas e Instituições de Arquivo precisam vencer mais esse desafio, pois necessitam lidar com essa massa de informação que vem sendo produzida e acumulada. Além dos ativos de informação culturais que já fazem parte das coleções destas instituições, existe uma grande quantidade de novas informações sendo geradas e que podem ser incorporadas segundo o conceito de patrimônio digital, entendido aqui como ativos de informação que transitam no ciberespaço e que possuem representatividade junto à história contemporânea. (Dodebei, 2007).

Na visão de Marcondes (2012, p. 174):

“A Web vem oferecendo possibilidades extraordinárias para que arquivos, bibliotecas, centros de documentação e museus possam ampliar sua oferta de serviços e produtos a seus usuários. Grande parte dessas instituições vêm disponibilizando seus acervos através da Web. [...]”

“[...] Para arquivos, bibliotecas e museus, a disponibilização de seus acervos através da Web está se constituindo, cada vez mais, na forma corrente de prestação de seus serviços a uma gama muito mais ampla de usuários.”

Ao lançar um olhar especialmente sobre Museus, segundo Maturana, Ortega & Lopez-Sola (2013), estima-se que existam cerca de 55.000 museus espalhados pelo mundo. Portanto, representar essas informações, de maneira que seja possível conectar os ativos de informação existentes entre eles, permitindo que os usuários construam novas ligações e criando outros conhecimentos é o desafio que se coloca ao profissional de informação.

Por outro lado, é possível inferir que a dinâmica para organização deste alto volume de informação está se estruturando e também vem impulsionando o desenvolvimento de novos estudos. O profissional que trabalha com a Organização da Informação começa a desenvolver investigações no campo conhecido como *Big Data*, onde a noção de instantaneidade e disponibilidade vem modificando a forma de atuação nos processos de gestão de informação sobre a cultura. (Ribeiro, 2014a).

Maturana, Ortega & Lopez-Sola (2013) continuam e observam que nas bibliotecas, arquivos e museus, estão presentes muitos ativos de informação cultural. Estes ativos são objetos de memória e que tem um enorme potencial para a construção de ligações com outros elementos na Web de Dados (*Web of data*). O conhecimento registrado e armazenado nestas instituições, precisa de conexões para desenvolver novos laços na direção da construção de conhecimento para a sociedade. No entanto, estes autores continuam e indicam que a construção dessas conexões é praticamente inviável sem o uso do aparato que suporta a Web Semântica, o que reforça o desenvolvimento desta proposta de investigação.

¹ Para facilitar o entendimento da característica ligada ao volume, a unidade Zettabyte é equivalente, em valores aproximados, a 1 mil Exabytes, ou a 1 milhão de Petabytes, ou ainda, a 1 bilhão de Terabytes

A necessidade de propor trabalhos que investiguem a possibilidade de aumentar a expressividade destas ligações, assume um papel de destaque quando se busca a integração massiva desses ativos de informação cultural. (Maturana, 2014).

Representando unidades de informação sobre ativos culturais na Web

Ribeiro (2008) indica que para o desenvolvimento de representações de unidades de informação em domínios, deve-se partir da investigação sobre a relação entre a linguagem, o entendimento do contexto e a percepção da verdade. Esta relação deve ser mediada tanto pela sintaxe quanto pela semântica e pragmática, como dimensões da verdade linguística, e que podem auxiliar na especificação da representação. Apel (2000) esclarece que enquanto a sintaxe busca observar a relação entre os signos, a semântica e a pragmática objetivam cobrir tanto os fatos designados quanto a percepção dos usuários sobre estes fatos. Portanto, mediado pela linguagem, o processo de representação ganha destaque na medida em que objetiva designar corretamente as qualidades conhecidas à luz de seus significados, que são percebidos, apreendidos e entendidos.

O processo de representação lida com informação percebida, registrada em algum suporte e disponível para uso, pois, caso contrário, as informações poderiam se tornar propriedade individual e inacessível. Resgatando-se o conceito filosófico para representação, nota-se sua aproximação com a percepção e o entendimento, pois na tradição filosófica, a palavra "representação" refere-se aos diversos modos de apreensão de um objeto.

No entanto, faz-se necessário uma breve discussão sobre este processo, pois ele precisa lançar mão de instrumentos que auxiliem a descrever e interpretar o objeto e/ou fenômeno sob análise. Nesta direção, percebe-se que tanto na Ciência da Informação quanto na Ciência da Computação, este entendimento e a sua natural representação passa por uma abordagem que lança mão de descrições lógicas, que fundamentam a elaboração das representações.

Em decorrência da existência de diferentes formas de expressão, é possível deduzir que as linguagens naturais representem grande parte do conhecimento existente. Entretanto, para fazer uso deste conhecimento, armazenando-o em qualquer meio digital e computadorizado, será preciso utilizar descrições precisas e bem formuladas, pois estes ambientes foram projetados para lidar com assertivas lógicas e formalismos estruturados com rigor matemático. Assim, o uso de lógica proposicional com sentenças e predicados, o uso de silogismos, além de ontologias, são definições importantes na formulação de qualquer proposta para representação do conhecimento. (Sowa, 2000).

A identificação das propostas de representação parte da definição, pelo observador, do domínio sob análise. Este observador, pode trabalhar com diferentes níveis de representação, contemplando as várias óticas que ele deve lançar sobre o contexto. (Ribeiro, 2012a). Estas óticas devem estar em acordo com níveis de abstração, que podem ser utilizados para o desenvolvimento da representação desejada. Na visão de Ron Brachman (1979 como citado em Sowa, 2000, p. 187), estes níveis são:

- **Nível Linguístico:** Nível inicial, compreende o uso de expressões em linguagem natural, conceitos e palavras;
- **Nível Conceitual:** Nível onde são materializadas as relações semânticas e algumas relações linguísticas, tanto com objetos e ações quanto com papéis e responsabilidades;
- **Nível Epistemológico:** Nível onde ocorrem as definições por tipos, subtipos, além de possíveis relações sobre heranças. Algumas relações adicionais podem ser observadas quando derivadas das definições obtidas;
- **Nível Lógico:** Nível de estruturas lógicas, proposições, predicados, variáveis, operadores booleanos e sentenças;
- **Nível de Implementação:** Nível de estruturas de dados, apontadores, listas e códigos em linguagem de programação.

Em cada representação e/ou nível de abstração, a escolha das propriedades que serão utilizadas passa pelo entendimento dos elementos que auxiliam na organização dos ativos de informação. Portanto, é necessário apresentar as definições de alguns elementos que poderão nos auxiliar na escolha dessas propriedades: os conjuntos, as coleções, os tipos e as categorias. (Sowa, 2000).

- Conjuntos: estrutura que deve ser utilizada na descrição de agrupamentos discretos (ex: animais, números inteiros, frutas, etc.), não servindo para uso em agrupamentos contínuos (ex: não existe um conjunto de água, nem de ar). A estrutura faz uso de dois tipos de operadores:
 - Operador 1: *MemberOf*
 - Operador 2: *SubsetOf*
- Coleções: estrutura mais simples que deve ser utilizada em agrupamentos contínuos. Possui apenas um tipo de operador:
 - Operador 1: *SubsetOf*
- Tipos: são estruturas que contêm uma especificação básica para um conjunto ou uma coleção de entidades que estão ou estarão presentes em um dado domínio ou contexto. São características imutáveis que independem das características das instâncias.
- Categorias: apesar de Aristóteles ter tornado categorias e predicados palavras com significados semelhantes, na atualidade estas têm se diferenciado. Predicado está mais próximo do significado proposto por Aristóteles para o entendimento de características das coisas. Já a categoria está mais próxima do processo de classificação, realizada de acordo com as características ou predicados essenciais.

Para a representação das propriedades e predicados nos níveis escolhidos, faz-se necessária a escolha do tipo de especificação, aspecto fundamental para a confecção de sentenças bem formuladas. São dois tipos de especificação: informais e formais. O uso de especificações informais, onde o problema é descrito em linguagem textual mas sem a preocupação em relatar detalhes necessários para a operacionalização, pode ser adequado para representações de alto nível de abstração, entretanto, ocasionará um esforço maior para sua interpretação, uma vez que este processo irá requerer um conhecimento prévio sobre o assunto. Já o uso de especificações formais, onde o problema é descrito em linguagem textual, mas com a incorporação de lógica e regras para a realização da tarefa em análise, pode ser mais apropriado para representações mais próximas dos níveis epistemológicos, lógicos e de implementação (Sowa, 2000).

O recorte adotado neste ensaio buscou trabalhar com especificações formais e com os formalismos RDF, OWL, Europeana/EDM, SKOS, Dublin Core, METS, MODS, FRBR e EAD. Cabe destacar que estes elementos não esgotam o tema da representação, no entanto, no processo de revisão bibliográfica estes foram percebidos como os mais relevantes na literatura investigada. Estão brevemente apresentados a seguir:

RDF

Resource Description Framework é um formalismo para descrição de recursos de informação. Pode ser entendido como uma arquitetura de metadados voltada para a descrição de recursos na Web (Klein, 2001). Podendo ser descrito em XML, possui a característica de permitir o processamento e a interpretação por máquinas, com utilização em diferentes áreas de aplicação: descrição de recursos de informação e seus componentes, elaboração de mapas de navegação e estruturação de sítios, comércio eletrônico, bem como no desenvolvimento de agentes de software. Possibilita a representação por meio de grafos, mas pode também ter a respectiva tradução em código entendível por computadores utilizando XML (RDF/XML) (W3c, 2004).

Com RDF as linguagens para descrição adquirem uma gramática comum (Baker, 2012). Seu modelo principal é composto pelos seguintes construtos (Rdf Primer, 2014)

Quadro1: Elementos RDF

namespace: rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"	
Construto	Exemplo de sintaxe
Class (uma classe)	rdf:class
Property (de classe)	rdf:property
Type (uma propriedade)	rdf:type
SubClassOf (uma propriedade)	rdf:SubClassOf
SubPropertyOf (uma propriedade)	rdf:SubPropertyOf
Domain (uma propriedade)	rdf:domain
Range (uma propriedade)	rdf:range

Fonte: elaborado pelo autor

OWL

A *Web Ontology Language* (OWL), foi desenvolvida pelo *Word Wide Web Consortium* (W3C) para ser um formalismo para criação de ontologias para a *Web Semântica*. Permite a representação de elementos tais como classes, propriedades, hierarquias, relações entre classes e valores, possibilitando a integração com outros elementos descritos em *RDF* e *SKOS*. Provê também a capacidade de descrever assertivas como forma de resolver problemas ligados à inferências lógicas e raciocínio (*Word Wide Web Consortium*, 2012).

Seu modelo principal está estruturado a partir de (Owl, 2015; Allemang & Hendler, 2011):

Quadro2: Elementos Owl

namespace: owl="http://www.w3.org/2002/07/owl#"	
Construtos	Exemplo de sintaxe
Entities (classes, propriedades e individuais)	owl:Class; a: Person
Axioms (assertivas sobre o domínio em análise)	owl:hasValue; owl:Restriction
Expressions (representando noções complexas)	owl:unionOf; owl:complementOf

Fonte: elaborado pelo autor

Europeana/EDM

Fruto de uma necessidade de trazer uma convergência para os esforços em utilizar diferentes padrões e formalismos na disponibilização de informações em meios digitais, o projeto *Europeana* surge como uma alternativa para preencher as lacunas de representação, possibilitando o reuso de outros mecanismos de descrição, além de contribuir para a construção de novos conhecimento originários destes ambientes culturais. Pode ser entendido como um framework que possibilita a interligação entre diferentes conteúdos e ativos culturais digitais. (*Europeana Primer*, 2013).

Nesse sentido, como evolução dos esforços desenvolvidos para representação no bojo do *Europeana Semantic Elements (ESE)*, surge então o *Europeana Data Model (EDM)*, como elemento facilitador para o desenvolvimento e compartilhamento de ativos de informação sobre a cultura, possibilitando alternativas para representação no âmbito da *Web Semantica*. (*Europeana Primer*, 2013). Quando estes elementos são utilizados em conjunto com recursos em *Linked Open Data*, possibilitam a melhoria do conteúdo disponível na *Web*, na medida em que permitem o enriquecimento de informações. Esta melhoria é possível com o cruzamento e a incorporação de novas informações originadas em outros provedores de conteúdo de ativos de informação cultural, participantes ou não da rede *Europeana*.

Seu modelo principal está estruturado a partir de (*Europeana Primer*, 2013; *Europeana Definition*, 2014)

Quadro3: Elementos Europeana/EDM

namespace: edm="http://www.europeana.eu/schemas/edm/"	
Construtos	Exemplo de sintaxe
EDM Classes	edm:Agent; edm:EuropeanaAggregation; edm:EuropeanaObject; edm:Event; edm:InformationResource; edm:Non- InformtionResource; edm:PhysicalThing; edm:Place; edm:ProvidedCHO; edm:TimeSpan; edm:WebResource;

EDM Properties	edm:aggregatedCHO; edm:begin; edm:country; edm:currentLocation; edm:datasetName; edm:dataProvider; edm:end; edm:europaProxy; edm:happenedAt; edm:hasMet; edm:hasType; edm:hasView; edm:incorporates; edm:isAnnotationOf; edm:isDerivativeOf; edm:isNextInSequence; edm:isRelatedTo; edm:isRepresentationOf; edm:isShowAt; edm:isShownBy; edm:isSimilarTo; edm:isSuccessorOf; edm:landingPage; edm:language; edm:object; edm:occuredAt; edm:preview; edm:provider; edm:realizes; edm:rights; edm:type; edm:ugc; edm:userTag; edm:wasPresentAt; edm:year
----------------	--

Fonte: elaborado pelo autor

SKOS

Por meio do vocabulário SKOS (*Simple Knowledge Organization System*) é possível representar estruturas de organização do conhecimento, de maneira semi-formal no ambiente Web, especialmente tesouros, taxonomias e esquemas de classificação. Em função da forte ligação entre SKOS, RDF e OWL, seu uso torna a representação processável por máquinas e facilita a troca de dados e informação entre aplicações de software no ambiente Web, além de prover algum nível de registro semântico sobre os ativos de informação (Skos Primer, 2009)

SKOS estabelece um conjunto de padrões na representação de conceitos e esquemas conceituais, além das relações de vizinhança necessárias para compreender melhor um conceito (*broader*, *narrower* e *related*) (Putkey, 2011):

Quadro4: Elementos SKOS

namespace: skos= "http://www.w3.org/2004/02/skos/core#"	
Construto	Exemplo de sintaxe
Concept scheme	skos:ConceptScheme
Top Concept	skos:hasTopConcept
Concept	skos:Concept
A broader term (BT)	skos:broader
A narrower term (NT)	skos:narrower
A related term (RT)	skos:related

Fonte: elaborado pelo autor

Dublin Core (DC)

A proposta do *Dublin Core* busca solucionar a demanda para trazer alguma semântica à descrição de recursos na Internet. O padrão consiste em um conjunto de metadados para descrição de objetos digitais, que traz a proposta de um "cartão virtual" para registrar os recursos que estarão sendo disponibilizados. O *Dublin Core* se tornou um padrão registrado como padrão ANSI e norma ISO 15836-2003.

Composto por 15 elementos considerados como núcleo da proposta, por utilizar elementos com estrutura XML permite a extensão e criação de novos elementos de maneira muito simples, tornando a proposta bastante flexível (Ribeiro, 2008).

Seu modelo principal é composto pelos seguintes termos (*Dublin Core*, 2012):

Quadro5: Elementos DC

namespace: dc="http://purl.org/dc/terms/"	
Construto	Exemplo de sintaxe
Contributor	dc:contributor
Creator	dc:creator
Description	dc:description
Identifier	dc:identifier
Publisher	dc:publisher
Rights	dc:rights
Subject	dc:subject
Type	dc:type
Coverage	dc:coverage
Date	dc:date
Format	dc:format
Language	dc:language
Relation	dc:relation

Source	dc:source
Title	dc:title

Fonte: elaborado pelo autor

FRBR

O modelo FRBR (*Functional Requirements for Bibliographic Records*) apresenta os requisitos mínimos que os registros bibliográficos em formato eletrônico devem possuir. É baseado no modelo Entidade-Relacionamento proposto por Peter Chen (Chen, 1976), fazendo uso dos conceitos de entidades, relacionamentos e atributos para representar recursos de informação. A concepção do modelo FRBR derivou de uma percepção da necessidade de compartilhamento de dados bibliográficos, o que culminou com geração da proposta pela IFLA (*International Federation of Library Associations and Institutions*). (Moreno, 1998).

No âmbito do compartilhamento de informações sobre ativos de informação cultural, cabe ressaltar que uma das propostas de evolução deste modelo está ligado à compatibilidade com o modelo de referência CIDOC-CRM, proposto pelo *Comité International pour la Documentation*. Este esforço procurou representar a visão unificada sobre as entidades do patrimônio histórico-cultural: as bibliotecas, os museus e arquivos. (Lima, 2010a).

Seu modelo principal é composto por classes em RDF que contemplam os Grupos 1, 2 e 3 do modelo FRBR. Com o intuito de esclarecer a composição dos elementos, para este relato houve um recorte sobre os elementos e serão apresentados apenas os 20 primeiros, apesar do processo de análise ter contemplado todos os 69 elementos (Frbr, 2009):

Quadro7: Extrato dos elementos FRBR

namespace: frbr="http://purl.org/vocab/frbr/core#"	
Construto	Exemplo de sintaxe
abridgement	frbr:abridgement
abridgement of	frbr:abridgementOf
adaption	frbr:adaption
adaption of	frbr:adaptionOf
alternate	frbr:alternate
alternate of	frbr:alternateOf
arrangement	frbr:arrangement
arrangement of	frbr:arrangementOf
classical work	frbr:ClassicalWork
complement	frbr:complement
complement of	frbr:complementOf
concept	frbr:concept
corporate body	frbr:CorporateBody
creator	frbr:creator
creator of	frbr:creatorOf
data	frbr:Data
embodiment	frbr:embodiment
embodiment of	frbr:embodimentOf
endeavour	frbr:endeavour
event	frbr:event

Fonte: elaborado pelo autor

METS

Sigla para Metadata Encoding & Transmission Standard. Esta proposta cobre os registros sobre partes descritivas, de administração e estruturação dos recursos de uma biblioteca digital. Usa XML para descrever os registros e tem o intuito de gerenciar o acesso, uso e preservação de objetos digitais (Cundiff, 2004). Sua estrutura usa seções que podem ser detalhados em elementos e atributos, organizados segundo (Mets, 2010):

Quadro8: Elementos METS

namespace: mets="http://www.loc.gov/METS/"	
Construto	Exemplo de sintaxe
root	<mets>
Header	<metsHdr>

descriptive metadata Section	<dmdSec>
administrative metadata Section	<amdSec>
file Section	<file>
structural Map Section	<structMap>
structural Link Section	<structLink>
behavior Section	<behavior Sec>

Fonte: elaborado pelo autor

MODS

Sigla para *Metadata Object Description Schema*. Em conjunto com o MARC21, esta proposta permite a catalogação de recursos produzidos em suporte eletrônico, por meio de uso de XML. As suas maiores vantagens são: simplicidade, marcações (tags) orientadas ao usuário e perfeita integração ao padrão MARC21. MODS pode definir elementos principais, elementos subordinados e atributos. (McCallum, 2004). Sua estrutura usa conceitos que podem ser detalhados em elementos de alto nível e seus atributos, organizados segundo:

Quadro9: Elementos MODS

namespace: mods="http://www.loc.gov/mods/v3"	
Construto	Exemplo de sintaxe
titleInfo	<titleInfo>
language	<language>
note	<note>
location	<location>
name	<name>
physicalDescription	<physicalDescription>
subject	<subject>
accessCondition	<accessCondition>
typeOfResource	<typeOfResource>
abstract	<abstract>
classification	<classification>
part	<part>
genre	<genre>
tableOfContents	<tableOfContents>
relatedItem	<relatedItem>
extension	<extension>
originInfo	<originInfo>
targetAudience	<targetAudience>
identifier	<identifier>
recordInfo	<recordInfo>

Fonte: elaborado pelo autor

EAD

Encoded Archival Description (EAD) é uma estrutura de dados padrão que possibilita aos arquivistas implementarem as bases do conceito da proveniência, bem como, desenvolver as descrições que podem facilitar nas pesquisas pela internet. Desenvolvido a partir dos conceitos originários em SGML (*Standard Generalized Markup Language*), possibilita o registro de informações sobre a autoria e sobre o conteúdo do arquivo e permite a criação de metadados para material arquivístico. Possibilita descrever não só as coleções completas mas também todas as partes desta coleção (Frea, 2003).

De forma análoga ao modelo FRBR, com o intuito de esclarecer a composição dos elementos, para este relato houve um recorte sobre os elementos e serão apresentados apenas os 20 primeiros, apesar do processo de análise ter contemplado todos os 146 elementos. (Ead, 2002)²

Quadro10: Extrato dos elementos EAD

namespace: ead="urn:isbn:1-931666-22-9"	
Construto	Exemplo de Sintaxe
Abbreviation	<abbr>

² Em 28 de Agosto de 2015 foi publicada a versão EAD3 (Ead, 2015), contendo revisões nos construtos, Tendo em vista a data de envio deste relato, não foi possível atualizar a análise para contemplar esta nova versão.

Abstract	<abstract>
Conditions Governing Access	<accessrestrict>
Accruals	<accruals>
Acquisition Information	<acqinfo>
Address	<address>
Address Line	<addressline>
Alternative Form Available	<altformavail>
Appraisal Information	<appraisal>
Arc	<arc>
Archival Description	<archdesc>
Archival Description Group	<archdescgrp>
Archival Reference	<archref>
Arrangement	<arrangement>
Author	<author>
Bibliography	<bibliography>
Bibliographic Reference	<bibref>
Bibliographic Series	<bibseries>
Biography or History	<bioghist>
Block Quote	<blockquote>

Fonte: elaborado pelo autor

A junção destes elementos em um espaço semântico na Web, de forma a que usuários consumam dados mesclados, ligados e consolidados, originários de múltiplas fontes, dependerá da capacidade do profissional de informação executar as respectivas representações usando formalismos e vocabulários disponíveis. (Baker, 2011). Neste sentido, outros conceitos precisam fazer parte do cotidiano do profissional que lida com representação da informação, pois ele passa a lidar com problemas de descrição e uso de lógica até então pouco conhecidos por ele.

Na visão de Barreto (2014), os estoques e fluxos na Web atual começam a ser renomeados para *Big Data*. Esta vertente está apoiada em características básicas conhecidas como os 5 (cinco) V's (Figura 1) que devem provocar mudanças no perfil do profissional que trabalha com representação de informação. Este profissional está sendo impulsionado a investigar e representar elementos que possuem características dinâmicas ainda pouco exploradas nas Ciências Sociais. (Ribeiro, 2014a).

Figura 1: Os 5 V's girando em torno de *Big Data*



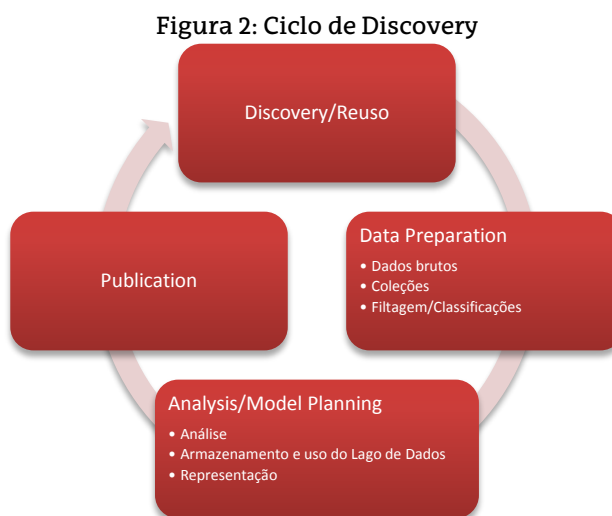
Elaborado pelo autor a partir de Hashem et. al. (2015) e Ibm (2015).

A necessidade de conhecer: o Volume de Dados e informação; o Valor que estes elementos têm para contribuir dentro dos objetivos estabelecidos; a Velocidade em que estes Dados surgem e necessitam de interpretação e representação; a Veracidade ou a credibilidade passada por estes Dados; e por último a Variedade decorrente das múltiplas fontes envolvidas; contribuem para o enriquecimento das atividades ligadas à representação, com o surgimento de outras formas de representar ativos de informação culturais.

Assim, o recorte adotado para o encaminhamento desta investigação parte do ciclo de *Discovery*, pois não é objetivo deste relato explorar o desenvolvimento de projetos em *Big Data* mas sim apresentar alternativas para o tratamento e a representação de

grandes volumes de dados e uma enorme variedade de informação. Cabe esclarecer que a organização de esforços em fases e etapas, onde o profissional de informação deve atuar nas atividades para descrever melhor os elementos deste domínio, poderá iniciar a jornada de representação, de forma a viabilizar o desenvolvimento de outras formas de descrição.

Um dos objetivos finais para o ciclo de *Discovery* é a identificação e adoção de representações para os conjuntos que serão carregados e armazenados no lago de dados. Este processo transcorre ao longo das fases de *Data Preparation* e *Analysis/Model Planning* (Figura 2). Estas fases poderão gerar produtos intermediários no trabalho de análise, tais como: Modelos Conceituais Gráficos; relação com as coleções; visões para categorizações; relação de propriedades para realizar as classificações; dentre outros. Estes produtos possibilitarão a melhor delimitação do domínio a ser modelado e representado, podendo ser validados pela comunidade de usuários que estão envolvidos no domínio alvo de análise. Esta dinâmica contribui para a melhoria da qualidade do produto final obtido. (Ribeiro, 2008).



Elaborado pelo autor a partir de Tavares (2014) e Khan et. al. (2014).

Neste sentido, propõe-se para o profissional da informação a adoção do ciclo de *Discovery* como caminho para desenvolver os processos de representação das múltiplas conexões possíveis.

Metodologia

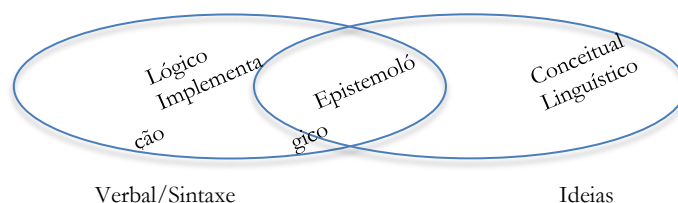
Esta pesquisa foi desenvolvida a partir de levantamento e revisão bibliográfica acompanhada de um estudo exploratório. Para Gil (2002) a pesquisa bibliográfica se desenvolve a partir de material presente em livros e material científico. O estudo exploratório “[...] tem o objetivo de reunir dados, informações, padrões, ideias ou hipóteses sobre um problema [...]”. (Braga, 2007, p. 25)

A partir da revisão bibliográfica foi reunido o *corpus* do conhecimento que trouxe a fundamentação para fomentar a discussão teórica sobre o processo de representação. Cabe ressaltar que os instrumentos presentes neste ensaio não esgotam o tema e correspondem a um recorte estabelecido para os objetivos da pesquisa do autor que está em curso na Unirio.

De modo a auxiliar a condução desse processo investigativo, apoiado em Ranganathan (1967 como citado em Lima (2010b, p. 1); Lima (2013)) foram estabelecidas duas subdivisões para os níveis de abstração encontrados em Sowa (2000). A primeira delas é denominada visão das ideias, onde foram categorizados os níveis Linguístico e Conceitual. Os níveis Epistemológico, Lógico e de Implementação foram categorizados na visão dita verbal ou da sintaxe. Vale destacar que em decorrência do detalhamento de características semânticas, o nível epistemológico poderia ser categorizado tanto no nível das ideias quanto no nível verbal/sintaxe

(Figura 1). O pressuposto adotado nesta pesquisa foi de que as definições de tipos, subtipos e herança se aproximam mais do nível verbal do que no nível das ideias.

Figura 3: Níveis de abstração para representação



Fonte: Elaborado pelo autor

Portanto, a reunião dos elementos identificados em um quadro de referencia, contendo o formalismo identificado, a classificação segundo os níveis de abstração e o tipo de especificação provida, auxiliou na estruturação da discussão. O quadro possibilitou evidenciar a cobertura dos formalismos em relação às visões de abstração (campo das ideias e no campo verbal), além da cobertura da sintaxe para caracterizar as propriedades dos ativos de informação.

Mecanismos de representação: análise e consolidação

Tendo em mente um processo de recuperação de informação eficaz, percebe-se que existem diferentes desafios na representação dos ativos de informação ligados à cultura. Neste sentido, a incorporação da temática sobre a avalanche de dados enriquece a discussão pois remete o trabalho de representação para além da atividades cotidianas do profissional de informação. A variedade de fontes, bem como o grande volume de dados e informação disponíveis na internet pode levar o usuário a fazer conexões nunca antes pensadas pelos profissionais envolvidos no processo de representação das informações. (Ribeiro, 2014).

Assim, pode-se inferir que os primeiros movimentos para o processo de representação passam pelo conhecimento do ambiente, além da estruturação e categorização das fontes. Parte-se da fase de *Discovery*, onde nas fases *Data Preparation* e *Analysis/Model Planning*, o profissional encarregado pela atividade de levantamento e representação deverá utilizar vocabulários e formalismos, de acordo com o objetivo da representação. A possibilidade de dotar de parcela semântica os dados que serão trabalhados, auxiliará os processos para o tratamento e recuperação de dados e informação.

Para isso, é preciso lançar mão de mecanismos que auxiliem na representação dos ativos de informação. Percebe-se que existem diferentes esforços que buscam convergir os resultados nos trabalhos de representação, por meio do uso dos formalismos apresentados. O uso destes têm permitido descrições para o nível de abstração caracterizado como verbal/sintático, de maneira bastante adequada aos requisitos de representação dos ativos de informação. O quadro abaixo traz uma análise, fazendo uma comparação sobre sua aplicação no processo de representação. (Quadro 11).

Quadro 11: Quadro analítico dos níveis de representação.

Formalismo	Nível de Sowa	Nível de Abstração	Tipo de Especificação provida
RDF	Epistemológico, Lógico e Implementação	Ideias e Verbal/Sintaxe	Formal. Conjuntos, coleções, tipos e categorias
OWL	Conceitual, Epistemológico, Lógico e Implementação	Ideias e Verbal/Sintaxe	Formal. Conjuntos, coleções, tipos e categorias
Europeana	Conceitual, Epistemológico, Lógico e	Ideias e Verbal/Sintaxe	Formal. Conjuntos, coleções, tipos e categorias

	Implementação		
SKOS	Lógico; Implementação	Verbal/ Sintaxe	Formal. Conjuntos, coleções, tipos e categorias
Dublin Core	Lógico; Implementação	Verbal/ Sintaxe	Formal. Tipo
FRBR	Epistemológico, Lógico e Implementação	Ideias e Verbal/ Sintaxe	Formal. Conjuntos, coleções, tipos e categorias
METS	Lógico e Implementação	Verba/ Sintaxe	Formal
MODS	Lógico e Implementação	Verbal/ Sintaxe	Formal. Coleções e tipos
EAD	Lógico e Implementação	Verbal/ Sintaxe	Formal. Coleções e tipos

Fonte: elaborado pelo autor

Discussão

As comunicações online entre diferentes aparatos tecnológicos, bem como as transações de negócio executadas entre empresas, podem ser solucionadas na dimensão da Descrição Lógica apresentada. No entanto, as interações entre os entes humanos, bem como entre esses entes e aparatos tecnológicos, vão além do uso dos aspectos da Lógica de representação e ficam limitadas muitas vezes pela lógica das máquinas. (Veltman, 2006). Em relação as interações homem-máquina e homem-homem, verificou-se que tanto a identificação da necessidade de informação quanto a representação dos ativos de informação cultural identificadas ainda estão distantes de soluções que expressem as múltiplas facetas da cultura.

A representação precisa ser tratada de forma múltipla, pois "Culture is multi-lingual and multi-cultural". (Veltman, 2002, p.2). Baca e Gill (2015) convalidam a percepção trazida por Veltman quando nos dizem que os ativos de informação sobre a cultura são de natureza multilíngue, multicultural, interdisciplinar e se manifestam em diferentes formatos.

Com o intuito de compreender os princípios que podem nortear a discussão desta temática, buscou-se tratar a representação segundo uma visão apoiada nos fundamentos do aprendizado. Na antiguidade, o estudante das Ciências e das Artes iniciava a vida escolar buscando vencer primeiramente os caminhos dos desafios da mente - o *Trivium*, contemplando a Gramática, Dialética e Retórica - para então depois se dedicar ao mundo das coisas - o *Quadrivium*, contemplando a Geometria, Aritmética, Astronomia e Música - e aos estudos em teologia, direito canônico e medicina. As preocupações sobre o entendimento, a semântica e o significado estavam ligadas não apenas a Lógica que era parte da Dialética. Este era o caminho trilhado na busca de uma vida intelectual madura. (Joseph, 2008). Portanto, para discutir o processo da representação no âmbito da cultura do mundo, bem como de sua história, parte-se do pressuposto que é preciso considerar não apenas a Dialética, mas todas as 7 artes ligadas à formação do saber.

Com base na análise do Quadro 11, pode-se verificar que boa parte da representação e do reuso estão centradas em questões ligadas ao nível verbal/sintaxe (Dialética/*Trivium*). As representações foram consideradas como formais, sendo que RDF, OWL, Europeana, SKOS e FRBR possuem capacidade de descrever mais adequadamente ativos de informação cultural no nível da abstração das ideias, sejam em Museus, Bibliotecas ou Arquivos Públicos e/ou Privados. Todavia, o alerta trazido por Veltman sobre a capacidade de representação para ativos de informação cultural calcada na Lógica formal de representação, permite inferir a necessidade de investigar diferentes caminhos que possam descrever de maneira mais completa estes ativos de informação.

Por outro lado, verifica-se que o desenvolvimento de vocabulários e formalismos vêm apoiando o processo na descrição de unidades de informação. O uso tanto de SKOS quanto de vocabulários gerados em domínios específicos, têm possibilitado uma boa relação entre os signos utilizados na representação e contribuindo para a resolução de problemas ligados à sintaxe e reuso (Ribeiro & Pereira; 2014). Putkey (2011) complementa esta questão explorando a utilização de SKOS e RDF, como forma

de dotar o processo de organização do conhecimento na representação de soluções para ativos de informação cultural.

Além disso, também demonstram a potencialidade destas conexões os esforços de publicação presentes na Alemanha (http://www.dnb.de/EN/Service/DigitaleDienste/LinkedData/linkedata_node.html), na Inglaterra (<http://www.bl.uk/bibliographic/datafree.html>), na Holanda (<http://semanticweb.cs.vu.nl/lod/am/>) e na Espanha (<http://datos.bne.es>). (Ribeiro, 2014b). No âmbito de investigações no Brasil, Marcondes (2012) e Santos Netos et. al. (2013) apontam caminhos para a interligação de registros bibliográficos em ambientes de informação na Web.

Percebeu-se nesta investigação que as dimensões relacionadas ao nível de abstração das ideias (semântico) ainda carecem de desenvolvimento, pois quando tratamos de representação desses ativos de informação ligados ao patrimônio e a cultura, surgem lacunas que precisam ser preenchidas com representações que podem ser caracterizadas como: de alto-nível ou específicas para domínios (Guarino, 1998); ou de fundamentação (Guizzardi & Wagner, 2010); ou ontologias leves (Guizzardi, 2007).

Ramalho, Zafalon e Ouchi (2014, p.4140), convalidam esta percepção quando, ao tratar de registros bibliográficos em sua pesquisa, propõem a geração de “[...] uma ontologia baseada nas entidades, atributos e relacionamentos previstos nos FRBR, cujo escopo considera as tarefas a serem desempenhadas pelos usuários [...]”. Ou seja, a busca de uma melhor compreensão no nível das ideias.

Neste sentido, Europeana pode ser um caminho virtuoso para o encaminhamento desta discussão, pois é “[...] o projeto mais ambicioso de interligação de instituições arquivísticas, bibliotecas e museus [...] [onde é possível] [...]qualificar e enriquecer semanticamente tanto objetos culturais e/ou suas representações digitais quanto suas diferentes relações [...]”.(Marcondes, 2012, p.188).

Entretanto, pode-se afirmar que o projeto Europeana tem seus recursos direcionados prioritariamente para o reuso e compartilhamento de ativos de informação sobre a cultura, lançando mão de outros formalismos também apresentados neste relato, o que motiva novas investigações e reforça a percepção apontada na questão de partida deste relato: os fundamentos utilizados na lógica de representação atual são suficientes para dar conta de representarmos os ativos de informação cultural?

Considerações finais

Ao tratar aspectos de organização e representação, verificou-se nesta investigação que os formalismos possuem capacidade limitada para descrever os ativos de informação cultural, pois deve-se considerar as limitações da Lógica em relação a dificuldades de representar acordos semânticos para propriedades essenciais e acidentais destes ativos (Sowa, 2000, p. 356). Assim, pode-se afirmar que o alerta de Veltman sobre o uso da representação Lógica continua válido e deve ser observado, na medida em que objetiva a melhoria dos processos de representação.

Por outro lado, em relação à dinâmica utilizada para a reunião de dados e informação em um espaço semântico na Web, infere-se que o uso da abordagem de *Big Data* pode ser uma estratégia facilitadora, trazendo orientação para o profissional da informação conseguir lidar com os 5 V's que se relacionam com o cenário adverso que a contemporaneidade apresenta.

Por fim, o trabalho com as camadas de meta-informação é um dos desafios presentes no modelo de Linguagens para Web Semântica de Tim Bernes-Lee. (Ribeiro, 2008). A visão dessa Web é questão essencial para recuperações e representações de ativos de informação, pois precisamos obter ocorrências diferentes de números e palavras soltas sem contexto. Precisamos chegar ao significado de textos, imagens e outras criações do espírito humano (Veltman, 2002). É preciso ir além, mapear e construir outras propostas de representação e aparatos que permitam mover de forma dinâmica entre diferentes linguagens, diferentes níveis de vocabulários, diferentes cronologias, e diferentes métodos de trabalho.

Referências

- Allemang, D., Hendler, J.** (2011). *Semantic Web for the working ontologist: Effective Modeling in RDFS e OWL*. Second Edition. Morgan Kaufmann, Elsevier: USA.
- Apel, K. O.** (2000). *A transformação da Filosofia*. São Paulo: Loyola.
- Archer, P.** (2014). Webinar Phil Archer. Congresso LOD-Brasil. Florianópolis. Recuperado em 01 de junho de 2015 de <https://www.youtube.com/watch?v=EQj12Ar50AE>.
- Baca, M. & Gill, M** (2015). Encoding Multilingual Knowledge Systems in the Digital Age: the Getty Vocabularies. *Knowledge Organization*. 42 (4). p. 232-243.
- Baker, T.** (2012). Libraries, languages of description and linked data: a Dublin Core perspective. *Library Hi Tech*. 30 (1), p. 116-133.
- Braga, K. S.** (2007). Aspectos relevantes para a seleção de metodologia adequada à pesquisa social em Ciência da Informação. In: Mueller, S. P. M. Métodos para a Pesquisa em Ciência da Informação. Brasília. Thesaurus. p. 17-38.
- Cundiff, M. V.** et al. An introduction to the Metadata Encoding and Transmission Standard (METS). *Library Hi Tech*, 22 (1), p. 52-64, 2004. Recuperado em 01 de agosto de 2015 de <http://www.emeraldinsight.com/10.1108/07378830410524495> >
- Chen, P. P.** (1976). The Entity-Relationship Model - toward a unified view of data. *ACM Transactions on Database Systems*, v. 1, n. 1. Mar.
- Dublin Core (2012). Dublin Core Metadata Element Set, Version 1.1. Recuperado em 20 de maio de 2015 de <http://dublincore.org/documents/dces/>.
- Dodebei, V.** (2007). *Digitalização do Patrimônio e Organização do Conhecimento*. Anais do VIII ENANCIB – Encontro Nacional de Pesquisa em Ciência da Informação 28 a 31 de outubro de 2007. Salvador. Bahia. Brasil . Recuperado em 30 de maio de 2015 de <http://enancib.ibict.br/index.php/enancib/viiienancib/paper/viewFile/2815/1943>.
- Ead (2002). Encoded Archival Description Tag Library, Version 2002. Recuperado em 25 de maio de 2015 de http://www2.archivists.org/sites/all/files/EAD2002TL_5-03-V2.pdf
- Ead (2015). Encoded Archival Description Tag Library, Version 2002. Recuperado em 06 de setembro de 2015 de <http://www2.archivists.org/sites/all/files/TagLibrary-VersionEAD3.pdf>
- Europeana Definition. (2014). Definition of Europeana Data Model v. 5.2.6 . Disponível em: http://pro.europeana.eu/files/Europeana_Professional/Share_your_data/Technical_requirements/EDM_Documentation//EDM%20Definition%20v5.2.6_01032015.pdf . Acesso em Julho de 2015.
- Europeana Pimer (2013). Europeana Data Model Primer. Recuperado em 10 de junho de 2015 de: http://pro.europeana.eu/files/Europeana_Professional/Share_your_data/Technical_requirements/EDM_Documentation/EDM_Primer_130714.pdf . Acesso em Agosto de 2015.
- Frean, N.** (2003). Encoded Archival Description on the Internet. *ABI/INFORM Global*, 27 (5), p. 369-370.
- FRBR (2009). Expression of Core FRBR Concepts in RDF. Recuperado em 15 de maio de 2015 de <http://vocab.org/frbr/core.html>
- Gil, A. C.** (2002). *Como elaborar projetos de pesquisa*. 4ª ed. São Paulo: Atlas.
- Guarino, N.** (1998). Formal Ontology in Information Systems. *Proceedings of FOIS'98*. Trento, Italy, 6-8 June. Amsterdam, IOS Press, pp. 3-15
- Guizzardi, G.** (2007). On Ontology, ontologies, Conceptualizations, Modeling Languages, and (Meta)Models, *Frontiers in Artificial Intelligence and Applications, Databases and Information Systems IV*, Olegas Vasilecas, Johan Edler, Albertas Caplinskis (Editors), ISBN 978-1-58603-640-8, IOS Press, Amsterdam.
- Guizzard, G.; Wagner, G.** (2010). Using the Unified Foundational Ontology (UFO) as a Foundation for General Conceptual Modeling Languages. In: R. Poli; M. Healy; A. Kameas (Orgs.); *Theory and Applications of Ontology: Computer Applications*. p.175-196. Germany: Springer Netherlands.
- Hashem, I. A. T., Yaqoob, I., Anua N. B., Mokhtar S., Hani A., Khan S. U.** (2015). The rise of "Big Data" on cloud computing: Review and open research issues. *Information Systems*, 47. p. 98-115

IBM (2015). IBM BigData e Analytics Hub. Why only one of the 5 Vs of big data really matters. Recuperado em Maio 03 de Maio de 2015 de <http://www.ibmbigdatahub.com/blog/why-only-one-5-vs-big-data-really-matters>.

Joseph, M. (2008). *O Trivium: As Artes Liberais da Lógica, Gramática e Retórica*. É Realizações Editora, São Paulo, SP.

Khan, N et. al. (2014). Big data: survey, technologies, opportunities, and challenges.(Report). *The Scientific World Journal*, 14.

Klein, M. (2001). XML, RDF, and Relatives. Tutorial the Semantic Web. *IEEE Intelligent Systems*. March/April. p. 26-28.

Lima, J. A. O (2010a). FRBROO – Conceitos básicos. In : Jaime Robredo; Marisa Bräscher (Orgs.). *Passeios pelo bosque da informação: Estudos sobre representação e organização da informação e do conhecimento*. Brasília DF: IBICT, p. 81-92. Edição eletrônica. Recuperado em 01 de junho de 2015 de: <http://www.ibict.br/publicacoes/eroic.pdf>.

Lima, J. A. O. (2010b). Interoperabilidade Semântica no LexML In: Cláudia do Socorro Ferreira Mesquita; Nazaré Lopes Bretas. (Org.). *Panorama da Interoperabilidade no Brasil*. Brasília: MP/SLTI, p. 74-79. Recuperado em 01 de junho de 2015 de <http://www.governoeletronico.gov.br/anexos/interoperabilidade-semanticano-lexml/>.

Lima, J. A. O. (2013). Uso de Ontologias no Governo. Palestra proferida no Ontobrás 2013 – Seminário de Pesquisa em Ontologias no Brasil. Recuperado em 10 de maio de 2015 de: https://www.google.com.br/url?sa=t&rct=j&q=&esrc=s&source=web&cd=4&ved=0CC8QFjAD&url=http%3A%2F%2Fontobras.eci.ufmg.br%2Fpt-br%2Fevento%2Fmaterial%2Fpalestra-joao-lima-uso-de-ontologias-no-governo%2Fat_download%2Ffile&ei=mIWeVcr4EceYwQT0irWADA&usq=AfQjCNEckkf6ntTdi3usORdwrctVwZ79Q&bvm=bv.96952980,d.Y2I.

Marcondes, C. H. (2012). “Linked Data” – Dados Interligados e interoperabilidade entre arquivos, bibliotecas e museus na Web. *Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação*, 17(34), p.171-192, maio/ago.

Maturana, R. A. (2014). *Big Data y Bibliotecas: convertir datos en conocimiento*. 7ª Jornada RBIC. Big Cultural Data. Recuperado em 20 de maio de 2015 de em: <https://youtu.be/oTmBg8hTwjE>.

Maturana, R. A.; Ortega, M.; Lopez-Sola, S. (2013) *Mismuseos.net: Art After Technology. Putting cultural data to work in a Linked Data platform*. Recuperado em 10 de maio de 2015 de <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.429.2803&rep=rep1&type=pdf>.

Mets Primer (2010). Recuperado em 01 de junho de 2015 de <http://www.loc.gov/standards/mets/METSPrimerRevised.pdf>.

McCallum, S. H. (2004). An introduction to the Metadata Object Description Schema (MODS). *Library Hi Tech*, 22 (1), p. 82-88. Recuperado em 01 de junho de 2015 de <http://www.emeraldinsight.com/10.1108/07378830410524521>.

Moreno, F. P. (2010). Atributos dos Requisitos Funcionais para Registros Bibliográficos (FRBR). In: Jaime Robredo; Marisa Bräscher (Orgs.). *Passeios no Bosque da Informação: Estudos sobre Representação e Organização da Informação e do Conhecimento*. Brasília DF: IBICT. 335 p. ISBN: 978- 85-7013-072-3. Capítulo 5, p. 93-114. Edição eletrônica. Recuperado em 15 de maio de 2015 de: <http://www.ibict.br/publicacoes/eroic.pdf>. (Edição comemorativa dos 10 anos do Grupo de Pesquisa EROIC).

Owl (2015). Recuperado em 20 de maio de 2015 de: <http://www.w3.org/TR/2012/REC-owl2-syntax-20121211/>.

Putkey, T. (2011). Using SKOS to Express Faceted Classification on the Semantic Web. *Library Philosophy and Practice (eJournal)*. Paper 620. Recuperado em 25 de maio de 2015 de: <http://digitalcommons.unl.edu/libphilprac/620>.

RDF 1.1 Primer (2014). Recuperado em 15 de junho de 2015 de: <http://www.w3.org/TR/2014/NOTE-rdf11-primer-20140225/>.

Ribeiro, C. J. S. (2008). *Diretrizes para o projeto de portais de informação: uma proposta interdisciplinar baseada na Análise de Domínio e Arquitetura da Informação*. Tese de Doutorado em Ciência da Informação. Rio de Janeiro. 298f

Ribeiro, C. J. S. (2012a). Entendimento de requisitos de sistema com abordagem orientada ao domínio. *DataGramaZero - Revista de Informação*, 13(2), abr. Recuperado em 08 de maio de 2015 de: http://www.dgz.org.br/abr12/Art_01.htm.

Ribeiro, C. J. S. (2014a). Big Data: os novos desafios para o profissional da informação. *Informação e Tecnologia*, 1, p. 96-105.

Ribeiro, C. J. S. (2014b). Painel 10 anos da Web Semântica: um olhar sobre ativos culturais e a criação do conhecimento para a sociedade. Congresso LOD-Brasil. Florianópolis. Recuperado em 01 de agosto de 2015 de <http://www.inf.ufsc.br/~tite/LODBrasil/Painel2/ClaudioRibeiro.pdf>.

Ribeiro, C. J. S. & Pereira, D. V. (2015). Uso do Vocabulário Controlado do Governo Eletrônico (VCGE) no processo de publicação de Dados Governamentais Abertos: proposta de revisão da classe sobre Previdência Social. *Transinformação*, 27(1), pp. 73-82.

Santos Neto, A. L. et. al. (2013). Tecnologias de dados abertos para interligar bibliotecas, arquivos e museus: um caso machadiano. *Transinformação*, 25(1), p.81-87, jan./abr.

Skos Primer (2009). SKOS Simple Knowledge Organization System Primer. Recuperado em 06 de junho de 2015 de <http://www.w3.org/TR/2009/NOTE-skos-primer-20090818/>.

Sowa, J. F., (2000). *Knowledge Representation*. Ed.. CAed.

Tavares, E. BIG DATA in Business. 12 de maio de 2014. Palestra apresentada no 2o. EMC Summer School on Big Data. EMC/NCE/UFRJ. Rio de Janeiro. 2014. Recuperado em 7 de maio de 2015 de http://2014.emcbigdataschool.nce.ufrj.br/images/presentations/Apresentacao_Elaine_Tavares.pdf.

Veltman, K. (2006). Towards a Semantic Web for Culture. *Journal Of Digital Information*, 4(4). Recuperado em 01 de junho de 2015 de <https://journals.tdl.org/jodi/index.php/jodi/article/view/113/112>

Veltman, K. (2002). *Challenges for Semantic Web*. Recuperado em 01 de junho de 2015 de <http://semanticweb2002.aifb.uni-karlsruhe.de/proceedings/Position/veltmann.pdf>.

World Wide Web Consortium (2012). OWL 2 web ontology language document overview. Recuperado em 01 de junho de 2015 de <http://www.w3.org/TR/owl2-overview/>.