

ORGANIZAÇÃO DE RECURSOS BIBLIOGRÁFICOS E MULTIMÍDIA NA WEB: CONTRIBUIÇÕES INTERDISCIPLINARES

ORGANIZACIÓN DE RECURSOS BIBLIOGRÁFICOS Y MULTIMEDIA EN LA WEB: CONTRIBUCIONES INTERDISCIPLINARES

Daniela Lucas da Silva Lemos^a
Renato Rocha Souza^b

RESUMO

Introdução: O artigo elucida a fertilização cruzada de tecnologias, teorias e metodologias interdisciplinares endereçada ao processo de construção de modelos ontológicos de qualidade centrados no usuário e destinados à representação de recursos bibliográficos e multimídia na Web. **Objetivo:** Apresentar as interlocuções de disciplinas seminais à área da organização da informação e do conhecimento oriundas dos campos da Ciência da Informação e da Ciência da Computação mostrando, em especial, o avanço contínuo da catalogação descritiva para ambientes de informações digitais, propiciando, assim, melhorias nos sistemas de recuperação da informação. **Metodologia:** A pesquisa foi delineada a partir de um levantamento bibliográfico e documental à luz de literatura específica, buscando explorar, descrever e explicar o fenômeno presente no campo da realidade documental em ambientes digitais. **Resultados:** O panorama atual concernente à representação de documentos bibliográficos e multimídia na Web indica a participação efetiva de dois campos de conhecimento – a Ciência da Informação e a Ciência da Computação – juntamente com suas disciplinas e princípios teóricos e metodológicos imprescindíveis para a evolução dos metadados às anotações semânticas. **Conclusões:** Os produtos e processos da organização e representação da informação e do conhecimento apresentados e discutidos no artigo mostraram a importância histórica e interdisciplinar de disciplinas fundadoras bem como tecnológicas no progresso da área de tratamento da informação. Tais disciplinas, como a representação descritiva, representação temática e modelagem conceitual contribuem sobremaneira com seus princípios teóricos e metodológicos no processo de construção de modelos ontológicos destinados a representar semanticamente documentos de natureza bibliográfica e

^a Doutora em Ciência da Informação pela Universidade Federal de Minas Gerais (UFMG). Professora do Departamento de Biblioteconomia da Universidade Federal do Espírito Santo (UFES). E-mail: danielalucas@hotmail.com

^b Doutor em Ciência da Informação pela Universidade Federal de Minas Gerais (UFMG). Professor da Escola de Matemática Aplicada da Fundação Getúlio Vargas (FGV/RJ). E-mail: renato.souza@fgv.br

multimídia na Web.

Descritores: Organização do Conhecimento. Ambiente de Informação Digital. Metadados. Ontologias. Anotação Semântica.

1 INTRODUÇÃO

Na literatura sobre “organização de conhecimento” encontram-se várias definições para este conceito (HODGE, 2000; DAHLBERG, 1993; ABBAS, 2010; HJØRLAND, 2003). Em muitos casos, considera-se a “organização de conhecimento” como equivalente à “organização de informação” (BRÄSCHER; CAFÉ, 2008). As autoras analisam e sintetizam a reflexão sobre os termos afirmando que o processo de organização da informação inclui um conjunto de objetos informacionais que são arranjados sistematicamente em coleções; já a organização do conhecimento lida com conceitos visando a construção de modelos para representar uma dada realidade. Desse modo, têm-se como entendimento que organizar a informação é diferente de organizar o conhecimento, embora muitas vezes seja necessário organizar o conhecimento para organizar a informação. Tal afirmação é compartilhada por Hodge (2000), que considera os sistemas de organização do conhecimento como “o coração de toda biblioteca, museu e arquivo”, uma vez que são “mecanismos de organização da informação”.

Uma definição de senso mais genérico é dada por Hjørland (2003) quando elucida que a organização do conhecimento está ligada à análise dos conceitos e relacionamentos de um domínio de conhecimento, com a conseqüente síntese de um Sistema de Organização do Conhecimento (do inglês, *Knowledge Organization System* - KOS). Segundo Hodge (2000), O termo KOS foi proposto em 1998 pelo Grupo de Trabalho de Sistemas de Organização do Conhecimento em Rede (do inglês, *Networked Knowledge Organization Systems Working Group*) para englobar sistemas de classificação, cabeçalhos de assunto, arquivos de autoridade, redes semânticas e ontologias.

Segundo Hjørland (2003), para a comunidade de Biblioteconomia e Ciência da Informação (do inglês, *Library and Information Science* – LIS) a organização do conhecimento encontra-se na organização da informação de registros bibliográficos, incluindo: i) associação, gerando relacionamentos; ii) representação, gerando pontos de acesso e índices em processos de catalogação e indexação; iii)

classificação, promovendo colocação e ordenação para os documentos; e iv) categorização, gerando esquemas de categorias. E geralmente faz-se uso de instrumentos que orientam processos, atividades e tarefas envolvidas no tratamento da informação. Neste contexto, os registros bibliográficos são produzidos pela subárea denominada Bibliografia. Esta é uma área primordial no guarda-chuva conceitual da Biblioteconomia e Ciências da Informação (HJØRLAND, 2007), a despeito de muitas vezes que consideram o campo obsoleto. Além da organização de listagens de livros correlacionados a um tópico específico (bibliografia enumerativa), compreende a descrição sistemática de livros como objetos físicos (Bibliografia descritiva), ramo no qual se estudam as confecções de representações e referências de materiais bibliográficos. Ainda há vertentes quantitativas em campos derivados, como a bibliometria e aos estudos de análises de citações, que são campos bastante influentes. Originalmente a Bibliografia tinha foco primordial em livros, mas hoje, em contextos de produção multimídia, são cobertas as descrições de mídias diversificadas, como áudio, vídeos, objetos gráficos, bancos de dados, sítios web e outras mídias. É nesta seara que este trabalho se encaixa.

A Ontologia Formal (GUARINO, 1998; GUIZZARDI, 2005) como uma disciplina da filosofia propõe desenvolver um sistema de categorias gerais e suas relações para fins de desenvolvimento de teorias científicas e teorias de senso comum em domínios específicos. Nesse sentido, o ontologista está interessado em todos os domínios reais em um alto nível de abstração (nível ontológico), diferentemente do cientista que busca tratar de questões específicas do campo (nível epistemológico). O interesse da disciplina denominada “Ontologia Aplicada” (*Applied Ontology*) (GUARINO, 1998; SMITH, 2004) concentra-se na aplicação de teorias da Ontologia Formal em domínios específicos. Almeida (2013) assinala que a disciplina Ontologia Aplicada pode trazer benefícios em relação ao progresso de pesquisas em Ciência da Informação no aspecto de teorias associadas à organização do conhecimento, especialmente no que diz respeito a teorias da classificação, assunto seminal do campo.

Desde a década de 90, a comunidade LIS vem pesquisando e fazendo uso

de ontologias (SOERGEL, 1997; VICKERY, 1997; FISCHER, 1998; SOERGEL, 1999; GILCHRIST, 2003; CURRÁS, 2003) na busca de padronizar a linguagem envolvida no domínio e facilitar a comunicação entre usuários e sistemas de recuperação de informação (SRI). A proposta de um SRI é levar o usuário ao encontro de suas necessidades de informação através das tarefas de: i) *representação* das informações contidas nos documentos: usualmente através dos processos de indexação e descrição dos documentos; ii) *armazenamento*: gestão física e/ou lógica dos documentos e de suas representações em algum tipo de base de dados; e iii) *organização e acesso aos itens de informação*: recuperação das informações relevantes e dos próprios documentos armazenados, de forma a satisfazer as necessidades de informações dos usuários (LANCASTER, 2004; BAEZA-YATES; RIBEIRO-NETO, 1999; ABADAL; CODINA, 2005).

Ontologias podem auxiliar o trabalho do profissional da informação no momento da representação dos documentos, pois são sistemas constituídos por um vocabulário e um conjunto de definições consensuais no âmbito de um domínio, que pode se apresentar em uma linguagem lógica. Esta, por sua vez, através de sua semântica formal, poderia eliminar contradições envolvendo os conceitos e as relações entre estes, resultando numa especificação não ambígua do domínio. E, finalmente, com base no formalismo, os usuários envolvidos num SRI poderiam se comunicar sob o mais alto nível de abstração a respeito de um determinado domínio, viabilizando, em função disso, a estratégia de busca e a recuperação da informação. As ontologias funcionariam, desta forma, como uma mediação entre a necessidade do usuário e a representação de documentos.

A correlação entre organização da informação e do conhecimento, técnicas computacionais e aspectos filosóficos, sociais e cognitivos pode contribuir sobremaneira para o desenvolvimento de instrumentos tecnológicos de qualidade, a partir da organização centrada no usuário. A fertilização cruzada de tecnologias, teorias e metodologias interdisciplinares pode contribuir no processo de construção de modelos ontológicos de qualidade endereçados à representação de recursos bibliográficos na Web, e em qualquer tipo de mídia. Nesse sentido, os processos,

produtos e teorias subjacentes envolvidos na organização e na representação da informação e do conhecimento em sistemas de informação podem ser arranjados em disciplinas, tais como representação descritiva, destacando o conceito “metadados”, representação temática, destacando o propósito dos vocabulários controlados; e modelos para a organização da realidade documental, destacando o uso de modelos conceituais bem fundamentados baseados em ontologias.

O propósito do presente artigo é apresentar as interlocuções destas disciplinas alicerçadas nos campos da Ciência da Informação (CI) e da Ciência da Computação (CC) mostrando, em especial, o avanço contínuo da catalogação descritiva para ambientes de informações digitais, propiciando, assim, melhorias nos sistemas de recuperação da informação.

A seção 2 apresenta o percurso metodológico adotado para delinear o esquema das interlocuções interdisciplinares no cumprimento do propósito em organizar recursos bibliográficos e multimídia na Web de dados. A seção 3 tece considerações sobre o entrelaçamento entre disciplinas da CI envolvidas no aspecto da representação e da organização de documentos. A seção 4 apresenta como resultados um panorama atual da área de representação de documentos bibliográficos e multimídia na Web, destacando o modelo conceitual ontológico proposto em Silva (2014) para o domínio de anotação de documentos dessa natureza.

2 METODOLOGIA DE PESQUISA

Esta seção se propõe a apresentar os fundamentos metodológicos empregados no delineamento da pesquisa, visando posicionar teoricamente a pesquisa frente ao tipo de abordagem em relação ao problema, objetivos e procedimentos técnicos utilizados em seu desenvolvimento. Elucida-se também a forma como foi conduzida a investigação nas bases de dados de documentos científicos adequadas à delimitação teórica e empírica do artigo.

Com base na abordagem do *problema*, esta pesquisa pode ser classificada como qualitativa, no sentido de buscar o entendimento do que está por detrás do fenômeno investigado do qual ainda pouco se conhece. Sendo assim, a determinação, a análise e a descrição das teorias, conceitos e relações interdisciplinares envolvidos na pesquisa serão possíveis através de um maior entendimento do fenômeno presente no campo da realidade documental em ambientes digitais.

Com base em seus *objetivos*, esta pesquisa pode ser classificada em: i) exploratória, pois pretende investigar e aprimorar ideias sobre um assunto emergente no campo das representações documentárias envolvendo ontologias para anotação de conteúdo multimídia, em que se requer uma investigação minuciosa na literatura e em casos específicos que fazem uso de recursos nessa perspectiva; ii) descritiva, pois pretende levantar e descrever características do fenômeno a ser investigado (conforme já elucidado acima) à luz da literatura; e iii) explicativa, pois a proposta de um panorama atual de uma determinada área de conhecimento, como a da representação de documentos bibliográficos e multimídia na Web, juntamente com suas interlocuções interdisciplinares alicerçadas em campos específicos requer a elucidação da realidade epistemológica envolvida.

E com base aos *procedimentos técnicos* para coleta e análise dos dados, esta pesquisa pode ser classificada como bibliográfica e documental, pois pretende usar como fonte de consulta materiais já publicados, como artigos científicos, anais de congressos, relatórios técnicos de pesquisa, teses, dissertações, normas e fontes documentais subjacentes a recursos de conhecimento ontológicos e não ontológicos selecionados na pesquisa.

A organização da realidade documental na Web de dados tem sido tratada como objeto de estudos em diferentes campos do conhecimento. Esta pesquisa está concentrada em alguns desses campos, dentre os quais podemos destacar a Ciência da Informação e a Ciência da Computação.

Para a identificação de documentos nos campos supracitados, foram consultadas bases de dados de documentos científicos no portal de periódicos da

Capes¹ e na biblioteca digital Citeseer². Esta última proporciona análises estatísticas das citações para todos os documentos do repositório, o que possibilita ao pesquisador tomar conhecimento da opinião dos pares em relação ao artigo em questão e, em consequência, de sua aceitação pela comunidade. No que diz respeito ao portal de periódicos da Capes, as editoras consultadas foram: i) *Association Computing Machinery*, conhecida como portal ACM; ii) *Journal Multimedia Tools and Applications*; iii) *IEEE MultiMedia*; iv) *Journal of signal processing systems for signal, image, and video technology*; e v) *IEEE Transactions on Circuits and Systems for Video Technology*. A primeira dissemina literatura de diversas áreas da CC, destacando a inteligência artificial, área na qual há várias publicações sobre representação do conhecimento e ontologias. A segunda e a terceira disseminam publicações sobre sistemas multimídia e hipermídia, bibliotecas digitais e Web de dados; além de revisões de literatura sobre ferramentas e aplicações emergentes envolvendo anotações de documentos multimídia com uso de vocabulários semânticos, especialmente ontologias. A quarta e a quinta fornecem informações de cunho tecnológico para processamento de sinais e extração de dados envolvendo imagem, vídeo e áudio; neste caso, a atenção voltou-se a assuntos relacionados à tecnologia MPEG-7, por ser o padrão de metadados multimídia usado na pesquisa. Pesquisas também foram realizadas nas revistas classificadas na categoria Qualis tanto da CI quanto da CC, além da Biblioteca Digital Brasileira de Teses e Dissertações (BDTD³).

3 REPRESENTAÇÃO E ORGANIZAÇÃO DA REALIDADE DOCUMENTAL

Svenonius (2000) expõe que a informação precisa ser descrita para a sua organização e que o produto desse processo descritivo é a representação da informação. Segundo Bräscher e Café (2008), a organização da informação compreende a organização de um conjunto de objetos informacionais que são

¹ <http://www.periodicos.capes.gov.br/>

² <http://citeseer.ist.psu.edu>

³ <http://bdttd.ibict.br>

arranjados sistematicamente em coleções por meio da descrição física (forma) e de conteúdo (assunto), também denominado de processo de catalogação. Neste caso, tem-se a organização da informação em bibliotecas, museus, arquivos, tanto tradicionais quanto eletrônicos. A catalogação, assim, considerada como um processo de tratamento documental pode incluir em seu âmbito a catalogação descritiva, a catalogação de assunto e, ainda, a classificação, como universo notacional de localização física do documento em bibliotecas organizadas pelo arranjo sistemático. Revela-se, portanto, as dimensões de forma e de conteúdo, por meio das denominadas catalogação descritiva e catalogação de assunto (disciplinas também denominadas de, respectivamente, representação descritiva e representação temática, conforme se apresentam na Figura 1).

As atividades envolvidas no processo de catalogação são orientadas geralmente por instrumentos, tais como normas técnicas, vocabulários controlados, manuais de bases de dados, regras, normas e instruções de catalogação, modelos conceituais e tabelas de classificação. Dentre esses instrumentos, torna-se válido destacar as regras e os modelos fundamentais para a catalogação, como os códigos ISBD, o AACR2 e o formato bibliográfico de intercâmbio de dados, o MARC (*Machine-Readable Cataloging*).

Velluci (1998) considera que o advento da Internet abriu espaço para bibliotecas e especialistas em informação na promoção de esforços na melhoria de métodos de descrição, organização e recuperação de objetos digitalizados e acessados remotamente. Nesse contexto, normas e instruções de catalogação emergentes como o RDA ou Descrição e Acesso de Recursos foram concebidos em princípios internacionais de catalogação e harmonizados com dois modelos conceituais, o FRBR e o FRAD, com vistas às atuais necessidades de descrição de recursos informacionais em ambientes digitais, possibilitando: interoperabilidade, compartilhamento de recursos, intercâmbio contínuo e reutilização de metadados. Oliver (2011, p. 8) considera que “o escopo ampliou-se, mas a meta continua a mesma: romper as barreiras que impedem a comunicação sobre recursos bibliográficos. A normalização continua sendo imprescindível”.

Tomando como ponto de partida os modelos FRBR e FRAD, a RDA introduz uma nova abordagem de descrição dos aspectos técnicos e de conteúdo de um recurso, substituindo as designações gerais de materiais e o conceito de tipos de materiais da AACR2 por três elementos: (i) tipo de conteúdo; (ii) tipo de mídia; e (iii) tipo de suporte. O primeiro reflete a forma fundamental de comunicação em que o conteúdo é expresso, como sons, música executada, imagem em movimento bidimensional e tridimensional, dentre outros. O segundo reflete o tipo geral de dispositivo de intermediação necessário para ver, tocar e exibir o conteúdo de um recurso, como áudio, vídeo, computador, dentre outros. E o terceiro elemento é também um atributo do nível de manifestação, porém mais concreto e específico do que o tipo de mídia. Pode ser definido como uma categorização que reflete o formato do meio de armazenamento, como imagem em formato jpeg e bmp; apresentação multimídia em formato SMIL e Flash, dentre outros.

A informação sobre o tipo de conteúdo, mídia e suporte é importante para o usuário, seja como um meio para descobrir recursos, seja para delimitar as buscas. Assim, a informação descrita em RDA é rigorosamente ordenada em informação sobre a expressão (a mensagem) e a informação sobre a manifestação (o suporte), tendo em vista que a RDA tem como base o FRBR, que por sua vez identifica atributos relativos à obra, expressão, manifestação e entidades de itens. Assim, ao melhorar os dados que são registrados e os pontos de acesso que são construídos, a RDA prepara o terreno para o aperfeiçoamento da prática de descobrimento de recursos pelo usuário, seja em ambientes de bibliotecas tradicionais, seja em ambientes da Rede.

Alvarenga (2003) contribui nessa perspectiva afirmando que no novo contexto de produção, organização e recuperação de objetos digitais, as metas de trabalho não se restringem à criação de representações simbólicas dos documentos constantes de um acervo. Compreendem a criação de novas formas de escrita para os hipertextos e a criação dos denominados metadados, muitos dos quais podem ser extraídos diretamente dos próprios objetos.

O Quadro 1 exibido a seguir apresenta um esquema sinóptico de regras tradicionais e emergentes voltadas à catalogação de recursos de informação, além de modelos conceituais endereçados à organização conceitual de dados bibliográficos (ABBAS, 2010; OLIVER, 2011; IFLA, 1998).

Quadro 1 - Esquema sinóptico de normas e modelos para catalogação

Normas e modelos conceituais de catalogação			Informações pontuais			Propósito	Fundamentação
			Período	Descrição	Características		
ISBD ⁶			Final da década de 1969	Descrição Bibliográfica Internacional Padronizada	O padrão cobre obras raras, recursos eletrônicos, materiais cartográficos, recursos contínuos; material não bibliográfico; música impressa.	Servir como padrão internacional que define um conjunto de elementos descritivos para incluir em registros bibliográficos.	Diretrizes do comitê do IFLA.
AACR ⁷ e AACR2			1ª edição em 1967 e 2ª edição em 1978	Código de Catalogação Anglo-Americano	Padronização da prática de catalogação; facilidade de cooperação entre bibliotecas; fundamenta o formato MARC e MARC21 em sistemas informatizados.	Efetuar grande parte da descrição bibliográfica em sistemas de informação.	Princípios da ISBD.
IFLA ⁸	Modelos Conceituais	FRBR ⁹	1997	Requisitos Funcionais para Registros Bibliográficos	Fundamentado no modelo entidade-relacionamento (CHEN, 1976).	Proporcionar uma base teórica-metodológica coerente de modo a construir uma prática de descobrimento de recursos.	Diretrizes do comitê do IFLA.
		FRAD ¹⁰	2009	Requisitos Funcionais para Dados de Autoridade	Fundamentado no modelo entidade-relacionamento (CHEN, 1976).	Incluir dados de autoridade ao modelo FRBR.	Diretrizes do comitê do IFLA.
RDA ¹¹			2004-2010	Descrição e Acesso de Recursos	Norma internacional de conteúdo para o ambiente digital. Utilizada para recursos tradicionais e não tradicionais, analógicos e digitais, dentro e fora da biblioteca.	Substituir o AACR2. Fornecer elementos para descrição e acesso a recursos numa perspectiva mais abrangente.	Instruções baseadas nos modelos FRBR e FRAD. Princípios Internacionais de Catalogação.

Fonte: elaborado pelo autor.

⁶ *International Standard Bibliographic Description.*

⁷ *Anglo-American Cataloging Rules.*

⁸ *Internacional Federation of Library Associations and Institutions [Federação Internacional de Associações e Instituições Bibliográficas].*

⁹ *Functional Requirements for Bibliographic Records.*

¹⁰ *Functional Requirements for Authority Data.*

¹¹ *Resource Description and Access.*

Na representação descritiva de documentos, a criação e uso de metadados torna-se essencial para a padronização e descrição de recursos de informação, pois promove a interpretação uniforme e universal, em qualquer idioma e em qualquer tipo de unidade de informação; além de ser a forma mais comumente empregada para agregar semântica a informações com o propósito de facilitar a busca de recursos (SIQUEIRA; SILVA, 2011). Pode ser considerado produto do desenvolvimento histórico de regras bibliográficas, mantendo, portanto, conexões evolutivas com os Códigos e com as novas metodologias de tratamento da informação em diversas mídias, a exemplo do RDA e das linguagens XML (*eXtensible Markup Language*), RDF (*Resource Description Framework*) e OWL (*Ontology Web Language*) para representações de características de objetos digitais, especialmente na Web.

Contudo, os elementos da descrição de documentos podem ser fornecidos por usuários em um formato livre, isto é, usando-se de linguagem natural, sem um controle de vocabulário. Neste caso, as descrições ficam sujeitas a problemas de heterogeneidade semântica devido à natureza ambígua da linguagem natural como sinonímia, polissemia, homonímias e variações de sintagmas (ANSI, 2005); além de fatores hierárquicos que provocam efeitos negativos na recuperação da informação. Por outro lado, a indexação pode descrever o assunto dos documentos seguindo uma linguagem documentária, construída artificialmente a partir da linguagem natural presente nos documentos, buscando-se obter um vocabulário controlado de um assunto específico (LANCASTER, 2004). Assim, nesse tipo de indexação determinam-se os cabeçalhos a empregar, distinguem-se homônimos, controlam-se sinônimos, preveem-se relações de equivalência, hierárquicas e associativas.

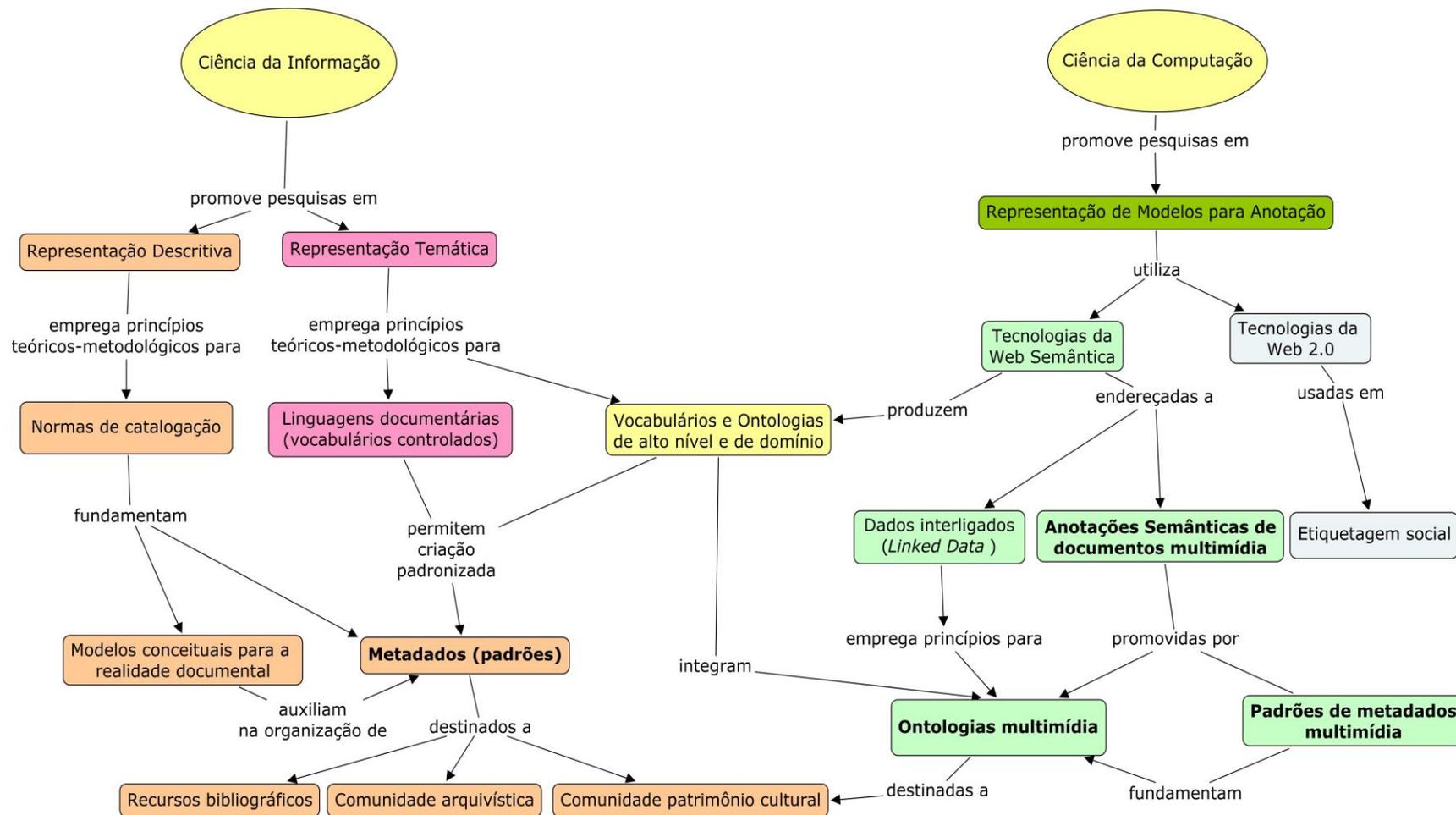
As ontologias possuem os mesmos princípios dos vocabulários controlados, isto é, trabalham com linguagem natural e fazem a delimitação de termos e de relações. Entretanto, a semântica envolvida na terminologia da ontologia se difere dos vocabulários controlados (dentro do contexto das linguagens documentárias) por incluir axiomas formais (através de declarações lógicas) que restringem a utilização do vocabulário. Outra distinção está no fato

de as especificações de relações no contexto das ontologias serem em número superior as dos tesouros, por exemplo.

Desse modo, a representação temática concentra-se na organização de conteúdos de documentos redundando em linguagens documentárias ou vocabulários controlados que podem auxiliar: i) nos processos de análise e descrição de documentos, permitindo a criação padronizada de metadados ao nomear, de forma consistente, os pontos de acesso aos documentos e a informação neles contida; e ii) no processo de busca em um SRI através da padronização e expansão do vocabulário das consultas.

A Figura 1 a seguir apresenta a relação entre as representações descritiva e temática no processo de organização de recursos bibliográficos, apoiando-se em princípios teóricos e metodológicos oriundos da Ciência da Informação na concepção de normas de catalogação, modelos conceituais, linguagens documentárias e ontologias destinadas à criação padronizada e à organização de metadados para as mais variadas necessidades.

Figura 1- Representação de documentos bibliográficos e multimídia na Web - panorama atual



Fonte: elaborado pelo autor.

A necessidade de inclusão de metadados para além de recursos textuais, isto é, metadados destinados à descrição de conteúdo multimídia, especialmente na Web, viabilizou o esforço conjunto entre comunidades e entidades normativas interessadas em fornecer um *framework* comum de metadados para aplicações de mídias inteligentes. São os casos do *World Wide Web Consortium (W3C)* e da *International Organization for Standardization/ International Electrotechnical Commission (ISO/IEC)*, que buscam soluções inteligentes para descrição de conteúdo multimídia processável por máquina e baseada em semântica por meio da Web Semântica (BERNERS-LEE; HENDLER; LASSILA, 2001); além da Interface de Descrição de Conteúdo Multimídia, conhecida como MPEG-7 (MARTÍNEZ; KOENEN; PEREIRA, 2002).

A comunidade de Biblioteca Digital usa metadados como auxílio na catalogação e na recuperação da informação em grandes coleções de documentos. O padrão Dublin Core é o comumente usado na comunidade com os seus 15 elementos de metadados e qualificadores específicos destinados a descrever principalmente proveniência, formato, idioma, direitos autorais e itens físicos.

Para o W3C, a Web Semântica é uma tentativa de produzir resultados de pesquisas em Biblioteca Digital e Representação do Conhecimento que sejam aplicáveis à Web. Nesse sentido, a seção 4 busca elucidar as contribuições da Ciência da Computação (apontadas na Figura 1) no aspecto de modelos e tecnologias voltados à descrição e à indexação de documentos, não necessariamente textuais, mas envolvendo também vídeos, imagens e áudios.

4 MODELOS DE ANOTAÇÃO PARA DOCUMENTOS BIBLIOGRÁFICOS E MULTIMÍDIA NA WEB

Em ambientes digitais, um recurso¹ pode ser descrito por meio de modelos de anotação expressos em *tags*, atributos, relações e ontologias

¹No contexto da Web, um recurso é qualquer artefato identificável por um identificador único (exemplo, um URI - *Uniform Resource Identifier*), como documentos eletrônicos expressados em mídias diversas.

dentre os quais se caracterizam por seus níveis de complexidade estrutural (BÜRGER, 2009). Um elemento de anotação baseado em *tag* caracteriza-se por uma palavra chave não hierárquica ou termo livre associado a um recurso. Um elemento de anotação baseado em “atributo” é representado por um par <AN, AV>, em que AN é o nome do atributo e AV é o seu valor. Já o modelo de anotação de relação é uma extensão do modelo de atributos, permitindo o usuário interligar recursos através de *links*. Exemplos de aplicações que fazem uso desses tipos de modelos são o Delicious (serviço social de compartilhamento de blogs, artigos, músicas, vídeos, etc); o Flickr (serviço de hospedagem de imagens); o CiteUlike (serviço social para compartilhar referências científicas); as redes sociais como Facebook e Picasa; a base de conhecimento Wikidata; e projetos como o Wikipedia.

Instituições que lidam com coleções de patrimônio cultural como bibliotecas, museus, arquivos e centros de documentação, e cujos acervos se encontram em portais ou bases de dados *online*, começam a perceber a importância de incorporar o conhecimento de usuários finais por meio de anotações (expressas em *tags*, atributos e relações) no processo de catalogação. Isso se deve ao fato da quantidade expressiva de recursos digitais a ser gerenciado e à limitação de recursos humanos disponíveis para catalogação. Essa nova forma de anotação é denominada *social tagging*, ou etiquetagem social (SCHANDL *et al.* 2011), em que Hunter (2009) a considera como uma subclasse especializada de sistemas de anotação promovida por tecnologias da Web 2.0. Nesse sentido, as anotações podem ser exploradas para aumentar a acessibilidade e a visibilidade de recursos digitais, que normalmente é de interesse de instituições culturais.

Apesar de serem modelos comumente usados no contexto da Web, as anotações oriundas de *tags*, atributos e relações ficam sujeitas a problemas de heterogeneidade semântica devido à natureza ambígua da linguagem natural, resultando em questões como (i) polissemia, (ii) sinonímia e (iii) fatores hierárquicos. Em (i) os elementos de anotação podem ter interpretação ambígua, por exemplo, o termo “Java” pode ser usado para descrever um recurso sobre a “ilha de Java” ou um recurso sobre a “linguagem de

programação Java”; essa situação pode reduzir a precisão na busca em um SRI. Em (ii), os elementos de anotação podem ser sintaticamente diferentes, mas com o mesmo significado, por exemplo, os nomes de atributos “é imagem de” e “é pintura de” podem ser usados alternativamente por usuários e, portanto, deveriam ser ligados como atributos sinônimos. E em (iii), os termos usados na anotação e na busca são diferentes em suas especificidades, por exemplo, quando o usuário executa a busca com o termo “queijo”, os recursos anotados com o termo “cheddar” (um tipo de queijo), caso não sejam vinculados, não serão encontrados no sistema.

Os problemas supracitados geralmente são endereçados usando-se de instrumentos como vocabulários controlados para identificar de forma unívoca e não ambíguos recursos ou documentos envolvidos em sistemas de recuperação de informação, conforme foi explanado na seção 3.

Ontologias como suporte a anotações podem ser usadas como vocabulários controlados, no entanto numa perspectiva de tratamento semântico, o que permite um usuário descrever e interligar recursos existentes por meio de qualificadores como conceitos, instâncias, propriedades e restrições mantidas entre tais recursos. O modelo é endereçado à anotação semântica de documentos, o que os tornam inteligentes no sentido de possibilitar conhecimento sobre o conteúdo, viabilizando processamento pela máquina (UREN *et al*, 2005). Neste exemplo, para evitar ambiguidade sobre a palavra “Paris” oriunda de um texto, uma anotação semântica poderia relacioná-la a um elemento da ontologia que a identificasse na categoria “Cidade” bem como associá-la à instância “França” pertencente à categoria “País”. Desse modo “Paris” não poderia ser referenciada de outra forma a não ser como uma cidade francesa.

Anotação semântica é uma abordagem subjacente aos conceitos preconizados pela Web Semântica e sua proposta emergente de *Linked Data* que intencionam criar metodologias, tecnologias e padrões de metadados para aumentar o escopo da interoperabilidade e da integração plena de informações heterogêneas entre sistemas de informação por meio de ontologias construídas e disponibilizadas por comunidades diversas. No domínio de anotação

semântica de documentos tanto de natureza bibliográfica quanto multimídia, destaca-se o modelo conceitual de referência proposto em Silva (2014), o qual consegue contribuir na perspectiva de possíveis soluções para o tratamento dos variados tipos de metadados existentes para descrição de acervos com conteúdo multimídia, conforme pode ser conferido nas ponderações que se seguem.

O modelo se destaca como uma solução abrangente para representação de documentos bibliográficos e multimídia (SILVA, 2014; SILVA; SOUZA, 2016; LEMOS; SOUZA, 2017) por ter sido concebido a partir de uma avaliação metódica, bem fundamentada e criteriosa realizada em ontologias nessa perspectiva, o que promoveu as condições necessárias para a seleção e o reúso de recursos de conhecimento apropriados ao domínio.

O propósito do modelo conceitual de referência é representar uma conceituação consensual e compartilhada por uma determinada comunidade para organização e recuperação semântica de documentos de natureza bibliográfica e multimídia. O modelo pode ser considerado um objeto de referência por ser subjacente a esforços de pesquisas voltados a modelos e tecnologias para processamento de metadados envolvendo comunidades da Web Semântica, biblioteca digital, representação do conhecimento e multimídia (MARTÍNEZ; KOENEN; PEREIRA, 2002; NACK; OSSENBRUGGEN; HARDMAN, 2005; SAATHOFF; SCHERP, 2010).

Os princípios teóricos e metodológicos discutidos nesta pesquisa se sustentam num arcabouço interdisciplinar envolvendo os campos da CI e da CC (conforme pode ser visualizado no mapa mental representado na Figura 1) em conjunto com outros que corroboram efetivamente para a sua evolução. São os casos da inteligência artificial, sistemas de informação, engenharia de software, lógica, psicologia cognitiva, filosofia e sociologia que dão condições necessárias à criação de abstrações da realidade e estruturas inteligentes capazes de definir e organizar formalmente as entidades que desempenham variados papéis em domínios de conhecimento específicos. O modelo conceitual de referência foi concebido a partir de tais princípios interdisciplinares quando empregou em sua estrutura central uma ontologia

conceitualista baseada em aspectos cognitivos, filosóficos e linguísticos, a DOLCE+DnS Ultralight (MASOLO *et al.*, 2003), a qual fornece metacategorias abstratas envolvendo eventos, objetos, tempo, espaço, etc, para organizar semanticamente conteúdos advindos de ontologias de domínios específicos. Agregado a essa rica estrutura conceitual, encontram-se também os padrões de projeto de conteúdo de ontologias (GANGEMI, 2005; PRESUTTI; GANGEMI, 2008; GANGEMI; PRESUTTI, 2009) que empregam soluções reusáveis para problemas de modelagem de conteúdo recorrente. No caso específico do modelo proposto, utilizaram-se padrões de projeto para a organização genérica de entidades associadas a conteúdo multimídia, como anotação (ex. uma apresentação multimídia que necessita ser anotada por um conceito semântico advindo de uma ontologia de domínio), segmentação (ex. decomposição de uma obra em capítulos) e coleção (ex. conjunto de imagens coletadas por diferentes pessoas sobre um assunto comum). Tais padrões possuem problemas e soluções comuns que podem ser usados em variadas áreas e domínios que lidam com recursos bibliográficos e multimídia. Assim, os padrões de projeto herdam de sua ontologia fundamental os axiomas e serviços de inferência necessários à manutenção da semântica dos elementos constituintes. A estrutura taxonômica do modelo proposto baseada em uma ontologia de fundamentação e padrões de projeto multimídia propicia solução à questão de *como expressar a estrutura conceitual subjacente à descrição da realidade documental de tipo multimídia*.

Outra característica relevante na estrutura do modelo conceitual de referência e que permite considerações interdisciplinares é o aspecto de evolução dos metadados às anotações semânticas. Conforme explanado na seção 3, o uso de metadados é uma prática antiga na área de biblioteconomia no processo de catalogação e indexação de recursos bibliográficos, especialmente com a adoção de padrões como o MARC e o Dublin Core, além de modelos conceituais para auxiliar a organização de metadados dessa natureza como o FRBR. Entretanto, a necessidade de inclusão de metadados para além de recursos bibliográficos, isto é, metadados endereçados à descrição de conteúdo multimídia, especialmente na Web, viabilizou o esforço

conjunto entre comunidades e entidades normativas interessadas em fornecer um *framework* comum de metadados para aplicações de mídias inteligentes. São os casos do W3C e da ISO que buscam soluções inteligentes para descrição de conteúdo multimídia processável por máquina e baseada em semântica por meio da Web Semântica e da Interface de Descrição de Conteúdo Multimídia, conhecida como MPEG-7. Contudo, observaram-se diferenças pontuais nas propostas dos *frameworks* envolvidos no que diz respeito à descrição de conteúdo multimídia.

A norma ISO MPEG-7 empreendeu esforços na proposição de uma interface comum para descrever material multimídia, refletindo informação sobre o conteúdo. Conforme se percebeu na pesquisa, o MPEG-7 apesar de ser um padrão de descrição recomendado pela comunidade multimídia, principalmente por seu *background* de conhecimento neste domínio, possui limitações de ordem semântica por ser baseado no formato XML *Schema*. Por outro lado, O W3C e comunidades afins vêm empregando esforços de pesquisas para ir além dos padrões de metadados correntes com a adoção de ontologias no domínio de anotação multimídia baseadas em RDF/OWL e no padrão MPEG-7.

A partir das considerações supracitadas, o modelo de referência buscou organizar sistematicamente tipologias de metadados existentes para descrição bibliográfica e de conteúdo multimídia frente a variados contextos e necessidades pela cobertura de requisitos funcionais pré-estabelecidos envolvendo categorias de metadados independentes de conteúdo, dependentes de conteúdo e descritivos de conteúdo. Para tal cobertura, os padrões de metadados MPEG-7 ISO/IEC e Dublin Core (considerados recursos de conhecimento não ontológicos nessa pesquisa) foram selecionados como material de referência para aquisição de conhecimento sobre o domínio bibliográfico e de anotação multimídia. A comparação de várias propostas de ontologias no domínio em pauta frente aos padrões de metadados ISO MPEG-7 e Dublin Core evidenciou características relevantes que podem e devem ser descritas para melhor recuperação de recursos multimídia, principalmente no contexto da Web. A necessidade de integração semântica e disponibilização

global de recursos multimídia na rede é um propósito comum entre as propostas de ontologias pesquisadas.

A primeira categoria foi organizada e endereçada a metadados relacionados à gestão da mídia, incluindo produção, classificação, gestão de direitos a uso e informações técnicas, os quais podem ser aplicados na realização da mídia e no conteúdo desta. Ressalta-se que o padrão de realização e informação (*Information and Realization Pattern*) representa a distinção entre objetos de informação e realizações de informação (MASOLO *et al.*, 2003), similar à proposta de modelagem do FRBR que promove uma percepção mais abrangente sobre o item bibliográfico, permitindo uma distinção entre o conteúdo (semântica) e a forma física do item (dado). Exemplos de objetos de informação são uma fotografia, uma pintura, peças teatrais, uma obra literária ou estruturas narrativas. Cada objeto de informação é realizado no mínimo por uma realização de informação, como, por exemplo, uma apresentação multimídia (objeto de informação) é realizada em formato SMIL e outra em Flash, mas ambos os formatos cobrem a mesma mensagem (conteúdo semântico). Outro exemplo seria o caso de uma imagem digital ser armazenada numa mídia em vários formatos e resoluções. Segundo Saathoff e Scherp (2010), a separação entre objetos de informação e realização da informação torna-se relevante no sentido de fornecer uma distinção clara entre a semântica (conteúdo da mensagem) e o dado (formato do arquivo de mídia).

A segunda categoria foi organizada e endereçada a metadados de nível baixo para aspectos visuais, de localização de regiões espaciais, temporais e espaço temporais no conteúdo da mídia, além de aspectos envolvendo processamento de sinais de áudio. E a terceira categoria foi organizada e endereçada a metadados para segmentos de mídia, anotação de conteúdo semântico, navegação e acesso por meio de sumários de conteúdo e características de áudio de nível alto, incluindo descritores para tratamento de conteúdo falado.

5 CONCLUSÕES

A fundamentação teórica apresentada nesse artigo foi essencial para a formação de um arcabouço teórico e conceitual sobre o assunto “organização da realidade documental” bem como a estruturação de um mapeamento interdisciplinar e bem fundamentado para o delineamento de um panorama atual no campo das representações documentárias.

Os produtos e processos da organização e representação da informação e do conhecimento organizados e discutidos nas seções 3 e 4 mostraram a importância histórica e interdisciplinar de disciplinas fundadoras bem como tecnológicas no progresso da área de tratamento da informação. Tais disciplinas, como a representação descritiva, representação temática e modelagem conceitual contribuem sobremaneira com seus princípios teóricos e metodológicos no processo de construção de modelos ontológicos destinados a representar semanticamente documentos de natureza bibliográfica e multimídia.

Nesse sentido, modelos conceituais ontologicamente bem fundamentados possibilitam a redução da ambiguidade presente em descrições documentais de tais domínios, permitindo um maior rigor na construção de representações descritivas de materiais de cunho documental, ampliando o escopo e da atividade bibliográfica na medida em que se estende para outros suportes documentais. Pôde-se observar com esta pesquisa, uma escassez de ontologias para anotação multimídia que adotam como teoria subjacente uma ontologia de fundamentação, tendo em vista de que esta representa ferramenta conceitual de importância para a concepção de modelos e consequentemente para a criação de ontologias de domínio de qualidade (GUIZZARDI, 2005; GUIZZARDI; FALBO; GUIZZARDI, 2008).

As teorias e métodos envolvendo ontologias como artefato foram importantes para o processo de modelagem envolvendo ontologias para anotação multimídia, tais como conceituação, compromisso ontológico e abstração da realidade. O modelo ontológico voltado à definição e organização de descritores e descrições abrangentes para o domínio pode ser concebido com maior qualidade por meio de princípios interdisciplinares oriundos da CI,

da CC e áreas correlatas. A agregação de esforços científicos multidisciplinares visa apoiar e melhorar processos voltados ao atendimento às necessidades informacionais dos usuários que fazem uso de sistemas de informação cada vez mais sofisticados.

O modelo de referência elucidado na seção 4 juntamente com sua fundamentação alicerçada na Web Semântica e em tecnologias afins forneceu um arcabouço conceitual importante ao entendimento de um assunto emergente envolvendo ontologias para anotação multimídia. Os padrões de metadados selecionados como referência à descrição de documentos desta natureza – MPEG-7 e Dublin Core - foram fundamentais para o delineamento de uma especificação de requisitos abrangente e consolidada em padrões internacionais para a operacionalização de atividades que culminaram na proposição do modelo, a saber: i) identificação e seleção de ontologias no domínio em pauta; ii) análise individual e comparativa das ontologias selecionadas; e iii) proposta de modelagem conceitual envolvendo as ontologias estudadas.

No que diz respeito a aplicabilidades, o modelo de referência pode ser inserido no contexto das Humanidades Digitais (CONEGLIAN; SANTAREM SEGUNDO, 2017), um campo de pesquisa emergente capaz de realizar a junção entre as áreas de tecnologia da informação e comunicação e as ciências humanas, buscando tornar mais efetivo o acesso e a recuperação das informações, por meio da aplicação das tecnologias. Outro contexto no qual o modelo poderia ser aplicado seria o de Dados Abertos Ligados (*Linked Open Data*). O modelo seria uma forma de publicar dados abertos padronizados utilizando as tecnologias da Web Semântica, permitindo consultas globais dos dados e ligações entre dados de diferentes fontes. Por exemplo, poder-se-ia fazer consultas utilizando fontes de dados específicas, além de outras fontes externas como, por exemplo, a base de conhecimento DBpedia². Desse modo, novos conhecimentos podem ser gerados através do cruzamento dos dados, permitindo vários tipos de consultas personalizadas.

² <http://dbpedia.org/About>

Finalmente, o modelo de referência proposto abre oportunidades de pesquisas em CI e CC voltadas a tecnologias eficientes para geração, exposição, descobrimento e consumo de recursos multimídia semanticamente vinculados na Web de dados.

REFERÊNCIAS

ABADAL, E.; CODINA, L. **Bases de datos documentales**: características, funciones y método. Madrid: Síntesis, 2005.

ABBAS, J. **Structures for organizing knowledge**: exploring taxonomies, ontologies, and other schema. New York: Neal-Schuman Publishers, 2010.

ALMEIDA, M. B. Revisiting ontologies: a necessary clarification. **Journal of the American Society of Information Science and Technology**, [S.l.], v. 64, n. 8, p. 1682-1693, 2013.

ALVARENGA, L. Representação do conhecimento da perspectiva da Ciência da Informação em tempo e espaço digitais. **Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação**, Florianópolis, v. 8, n. 15, p. 18-40, 2003. Disponível em: <<http://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2003v8n15p18/5233>>. Acesso em: 22 abr. 2018.

ANSI/NISO Z39.19-2005 (R2010). **Guidelines for the construction, format, and management of monolingual controlled vocabularies**. Baltimore: NISO Press, 2005. 184 p.

BAEZA-YATES, R.; RIBEIRO-NETO, B. **Modern information retrieval**. New York: Addison Wesley, 1999.

BERNERS-LEE, T.; HENDLER, J.; LASSILA, O. The semantic web. **Scientific American**, [S.l.], v. 284, n. 5, p. 34-43, maio 2001.

CAFÉ, L.; BRASCHER, M. Organização da informação ou organização do conhecimento. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO (ENANCIB), 9, 2008. **Anais eletrônicos...**São Paulo, USP, 2008.

BÜRGER, T. *et al.* **INSEMTIVES**: deliverable 2.1.1, report on the state-of-the-art and requirements for annotation representation models. 2009. Disponível em: <<http://eprints.biblio.unitn.it/1808/1/007.pdf>>. Acesso em: 22 abr. 2018.

CHEN, P. P. **The Entity-Relationship Model - Toward a Unified View of Data**. 1976. Disponível em:
<<http://www.inf.unibz.it/~nutt/IDBs1011/IDBPapers/chen-ER-TODS-76.pdf>>.
Acesso em: 22 jun. 2018.

CONEGLIAN, C. S.; SANTAREM SEGUNDO, J. E. Europeana no Linked Open Data: conceitos de Web Semântica na dimensão aplicada das Humanidades Digitais. **Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação**, v. 22, n. 48, p. 88-99, jan./abr. 2017.

CURRÁS, E. **Ontologies, taxonomies and thesauri in systems science and systematics**. Cambridge: Woodhead Publishing, 2003.

DAHLBERG, I. Knowledge organization: its scope and possibilities. **Knowledge Organization**, [S.l.], v. 20, n. 4, p. 211-222, 1993.

FISCHER, D. **From Thesauri towards Ontologies?** In: International Society for Knowledge Organization Conference, 1998, Lille. *Proceedings...* Würzburg: Ergon Verlag, 1998. Structures and Relations in Knowledge Organization.

GANGEMI, A. Ontology Design Patterns for Semantic Web Content. In: International Semantic Web Conference, 2005, Galway. **Proceedings...** Heidelberg: Springer-Verlag Berlin, 2005. p. 262-276. Disponível em:
<<http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=C9EDAD3D6AAC5B93C21D11057C780555?doi=10.1.1.108.7452&rep=rep1&type=pdf>>. Acesso em: 22 abr. 2018.

GANGEMI, A.; PRESUTTI, V. Ontology design patterns. In: STAAB, S.; STUDER, R. (Ed.). **Handbook on ontologies**. 2. ed. Berlin: Springer - Verlag. 2009. p. 221-243. Disponível em:
<<http://books.google.com.br/books?id=W6ZNCaolVbwC&pg=PA221&lpg=PA221&dq=ontology+design+patterns+handbook&source=bl&ots=fAKBdXOS-E&sig=ey-jmLbOKzB-zcMH-YyrKfvavog&hl=pt-BR&sa=X&ei=m1BhU8i2EdHgsATzuYDwDQ&ved=0CFcQ6AEwBA#v=onepage&q=ontology%20design%20patterns%20handbook&f=false>>. Acesso em: 22 abr. 2018.

GILCHRIST, A. Thesauri, taxonomies and ontologies: an etymological note. **Journal of Documentation**, [S.l.], v. 59, n. 1, p. 07-18, 2003.

GUARINO, N. **Formal ontology in information systems**. 1998. Disponível em:
<<http://citeseer.ist.psu.edu/viewdoc/download;jsessionid=E88DA9B5B5A9797C83C1F2E3C907991F?doi=10.1.1.29.1776&rep=rep1&type=pdf>>. Acesso em: 22 abr. 2018.

GUIZZARDI, G. **Ontological foundations for structural conceptual models**. 2005. 416 f. Tese (Doutorado em Ciência da Computação) – Universidade de Twente, Enschede, Holanda, 2005.

GUIZZARDI, G.; FALBO, R. A.; GUIZZARDI, R. S. S. A importância de ontologias de fundamentação para a engenharia de ontologias de domínio: o caso do domínio de processos de software. **IEEE Transactions Latin America**, [S.l.], v. 6, p. 244-251, 2008.

HJØRLAND, B. Fundamentals of knowledge organization. **Knowledge Organization**, [S.l.], v. 30, n. 2, p. 87-111, 2003.

_____. Arguments for 'the bibliographical paradigm'. Some thoughts inspired by the new English edition of the UDC, **Information Research**, v. 12, n. 4, 2007. Disponível em: <<http://InformationR.net/ir/12-4/colis06.html>>. Acesso em: 22 jun. 2018.

HODGE, G. Systems of knowledge organization for digital libraries: beyond traditional authority files. In: **The council on library and information resources**. Washington, DC, 2000. Disponível em: <<http://www.clir.org/pubs/reports/pub91/contents.html>>. Acesso em: 22 abril 2018.

HUNTER, J. Collaborative semantic tagging and annotation systems. In: CRONIN, B. (Ed.). **Annual review of information science and technology**. [S.l.]: American Society for Information Science & Technology, 2009. v. 43, cap. 2.

INTERNATIONAL FEDERATION OF LIBRARY ASSOCIATIONS AND INSTITUTIONS (IFLA). **Functional requirements for bibliographic records**. Munique: K.G. Sauer Verlag, 1998. 144 p. Study Group on the Functional Requirements for Bibliographic Records. Disponível em: <<http://www.ifla.org/files/assets/cataloguing/frbr/frbr.pdf>>. Acesso em: 22 abr. 2018.

JIMENEZ, A. G. Instrumentos de representación del conocimiento: tesauros versus ontologías. **Anales de Documentación**, [S.l.], n. 7, p. 79-95, 2004.

LANCASTER, F. W. **Indexação e resumos: teoria e prática**. Brasília: Briquet de Lemos, 2004.

LEMO, D. L. da S.; SOUZA, R. R. Modelos de Anotação para Representação de Documentos: uma Proposta Ontológica para o Domínio de Anotação Multimídia. In: Encontro Nacional de Pesquisa em Ciência da Informação (ENANCIB), 18, 2017, Marília. **Anais...** Marília: ANCIB, 2017.

MARTÍNEZ, J.; KOENEN, R.; PEREIRA, F. MPEG-7: the generic multimedia content description standard - part 1. **IEEE Multimedia**, [S.l.], v. 9, n. 2, p. 78-87, abr./jun. 2002.

MASOLO, C. *et al.* **Ontology library: wonder web deliverable D18**. Trento, 2003. Disponível em: <<http://www.loa.istc.cnr.it/old/Papers/D18.pdf>>. Acesso em: 22 abr. 2018.

NACK, F.; OSSENBRUGGEN, J. V.; HARDMAN, L. H. That obscure object of desire: multimedia metadata on the web -part 2. **IEEE MultiMedia**, [S.l.], v. 12, n. 1, p. 54-63, 2005.

OLIVER, C. **Introdução à RDA: um guia básico**. Brasília: Briquet de Lemos, 2011.

PRESUTTI, V.; GANGEMI, A. Content ontology design patterns as practical building blocks for web ontologies. In: International Conference on Conceptual Modeling, 2008, Barcelona. **Proceedings...** Heidelberg: Springer – Verlag Berlin, 2008. p. 128-141.

SAATHOFF, C.; SCHERP, A. Unlocking the semantics of multimedia presentations in the web with the multimedia metadata ontology. In: International Conference on World Wide Web, 2010, Raleigh. **Proceedings...** New York: ACM, 2010. p. 831-840.

SCHANDL, B. *et al.* **Linked Data and multimedia: the state of affairs**. **Multimedia Tools and Applications**, [S.l.], p. 01-34, 2011.

SILVA, D. L. da. **Ontologias para representação de documentos multimídia: análise e modelagem**. 2014. 441 f. Tese (Doutorado em Ciência da Informação) – Universidade Federal de Minas Gerais, Escola de Ciência da Informação, Belo Horizonte, 2014.

SILVA, D. L.; SOUZA, R. R. Modelagem Conceitual baseada em Ontologias na Organização de Documentos Multimídia. In: Encontro Nacional de Pesquisa em Ciência da Informação (ENANCIB), 17, 2016, Salvador. **Anais...** Salvador: ANCIB, 2016.

SIQUEIRA, I. C. P.; SILVA, J. F. M. da. Metadados: o fio de Ariadne ou a coragem de Teseu? **Bibliotecas Universitárias: Pesquisas, experiências e perspectivas**, v. 1, n. 1, p. 11-18, jan./jun. 2011.

SMITH, Barry. **Ontology and Information Systems**. 2004. Disponível em: <[http://ontology.buffalo.edu/ontology\(PIC\).pdf](http://ontology.buffalo.edu/ontology(PIC).pdf)> Acesso em: 22 jun. 2018.

SOERGEL, D. **Functions of a thesaurus / classification / ontological knowledge base**. University of Maryland, College of Library and Information Services, 1997.

_____. The rise of ontologies or the reinvention of classification. **Journal of the American Society of Information Science**, [S.l.], v. 50, n. 12, p. 1119-1120, 1999.

SVENONIUS, E. **The intellectual foundations of information organization**. Cambridge: The MIT Press, 2000. 255 p.

UREN, V. *et al.* Semantic annotation for knowledge management: requirements and a survey of the state of the art. **Journal of Web Semantics**, [S.I.], v. 4, n. 1, p. 14-28, 2005.

VELLUCI, S. L. Metadata. **Annual Review of Information Science and Technology (ARIST)**, [S.I.], v. 33, p. 189-222, 1998.

VICKERY, B. C. Ontologies. **Journal of Information Science**, [S.I.], v. 23, n. 4, p. 277-286, 1997.

ORGANIZATION OF BIBLIOGRAPHIC RESOURCES AND MULTIMEDIA ON THE WEB: INTERDISCIPLINARY CONTRIBUTIONS

ABSTRACT

Introduction: The article approaches the cross-fertilization of technologies, theories and interdisciplinary methodologies aiming the process of building high quality ontological models, focused on the user and oriented to the representation of bibliographic and multimedia resources on the Web. **Objective:** Present the intertwining of seminal disciplines to the area of information and knowledge organization, namely the fields of Information Science and Computer Science, showing, in particular, the continuous advance of descriptive cataloging for digital information environments, improvements in information retrieval systems. **Methodology:** This research originated from a bibliographical and documentary survey, backed by specific literature, seeking to explore, describe and explain the phenomenon of documents in digital environments. **Results:** The current panorama that encompasses the representation of bibliographic and multimedia documents on the Web indicates the effective participation of two fields of knowledge - Information Science and Computer Science - together with its disciplines, theoretical and methodological principles, essential for the evolution of the metadata to semantic annotations. **Conclusions:** The products and processes of the organization and representation of information and knowledge showed the historical and interdisciplinary importance of the disciplines in the progress of the area of information treatment. Such disciplines, such as descriptive representation, thematic representation and conceptual modeling contribute greatly with their theoretical and methodological principles in the process of constructing ontological models destined to semantically represent bibliographic and multimedia documents on the Web.

Descriptors: Knowledge Organization. Digital Information Environment. Metadata. Ontologies. Semantic Annotation.

ORGANIZACIÓN DE RECURSOS BIBLIOGRÁFICOS Y MULTIMEDIA EN LA WEB: CONTRIBUCIONES INTERDISCIPLINARES

RESUMEN

Introducción: El artículo elucida la fertilización cruzada de tecnologías, teorías y metodologías interdisciplinarias dirigidas al proceso de construcción de modelos ontológicos de calidad centrados en el usuario y destinados a la representación de recursos bibliográficos y multimedia en la Web. **Objetivo:** Presentar las interlocuciones de disciplinas seminales al área de la organización de la información y del conocimiento oriundas de los campos de la Ciencia de la Información y de la Ciencia de la Computación mostrando, en especial, el avance continuo de la catalogación descriptiva para ambientes de informaciones digitales, propiciando, así, mejoras en los sistemas de recuperación de la información. **Metodología:** La investigación fue delineada a partir de un levantamiento bibliográfico y documental a la luz de literatura específica, buscando explorar, describir y explicar el fenómeno presente en el campo de la realidad documental en ambientes digitales. **Resultados:** El panorama actual concerniente a la representación de documentos bibliográficos y multimedia en la Web indica la participación efectiva de dos campos de conocimiento -la Ciencia de la Información y la Ciencia de la Computación- junto con sus disciplinas y principios teóricos y metodológicos imprescindibles para la evolución de los metadatos a las anotaciones semánticas. **Conclusiones:** Los productos y procesos de la organización y representación de la información y del conocimiento presentados y discutidos en el artículo mostraron la importancia histórica e interdisciplinaria de disciplinas fundadoras así como tecnológicas en el progreso del área de tratamiento de la información. Tales disciplinas, como la representación descriptiva, representación temática y modelado conceptual contribuyen sobremanera con sus principios teóricos y metodológicos en el proceso de construcción de modelos ontológicos destinados a representar semanalmente documentos de naturaleza bibliográfica y multimedia en la Web.

Descriptores: Organización del Conocimiento. Entorno de Información Digital. Metadatos. Ontologías. Anotación semántica.

Recebido: 10.05.2018

Aceito: 25.08.2018