



Organização e representação do conhecimento e da informação na web: teorias e técnicas

Gercina Ângela de Lima^I

<http://orcid.org/0000-0003-0735-3856>

^I Universidade Federal de Minas Gerais, MG, Brasil.

Doutora em Ciência da Informação.

Professor do PPG-GOC/Escola de Ciência da Informação.

<http://dx.doi.org/10.1590/1981-5344/4299>

Apresenta reflexões sobre as teorias e técnicas da área da Organização e Representação do Conhecimento e da Informação em relação às tecnologias da Web, no âmbito da pesquisa no Grupo MHTX e da disciplina Organização do Conhecimento e da Informação, no Programa de Pós-graduação Gestão e Organização do Conhecimento (PPGGOC), da Escola de Ciência da Informação da UFMG. Retrata uma visão didática do conteúdo a partir da necessidade de planejar essa disciplina valendo-se da experiência em pesquisa e ensino, e ao constatar as dificuldades de absorção do conhecimento quando se mantém o diálogo de aprendizagem somente no plano teórico. A partir de uma revisão narrativa da literatura, discorre sobre: (1) o processo de modelagem conceitual para elaboração dos Sistemas de Organização do Conhecimento (SOC), (2) as características de quatro SOC, (3) os processos de catalogação, indexação e classificação na representação da informação e (4) a contribuição dos recursos tecnológicos para representação do conhecimento e da informação na Web, com vistas à interoperabilidade entre os sistemas e a melhoria da recuperação da informação. Conclui que o grande desafio para o desenvolvimento da Representação do Conhecimento e da Informação na Web está na grande quantidade de dados não estruturados e na dificuldade

cada vez maior de acesso a esse volume de informação e sua recuperação. As soluções vislumbradas perpassam (1) pela criação de ontologias para auxiliar na disponibilização de conteúdos que necessitam ser mais bem explicitados na web para serem recuperados; (2) pela disponibilização de informações em diversos formatos e acessos; (3) pelo serviço de referência virtual para ajudar os usuários a localizar as melhores fontes de informação; (4) pelo desenvolvimento de repositórios com recursos de aprendizagem disponíveis em diferentes formatos; (5) pelo armazenamento da informação em nuvens (Cloud Computing), (6) pela melhoria da padronização para tornar o conteúdo acessível para pessoas com deficiência.

Palavras-chave: Organização do Conhecimento. Representação do Conhecimento. Organização da Informação. Representação da Informação. Web Semântica. Sistemas de Organização do Conhecimento. Grupo MHTX.

Organization and Representation of Knowledge and Information on the Web: Theories and Techniques

The paper presents reflections on the theories and techniques of Organization and Representation of Knowledge and Information with focus on the Web environment, within the scope of the MHTX Research Group and the discipline Organization of Knowledge and Information in the Graduate Program in Knowledge Management & Organization (PPG-GOC). It portrays a didactical view of the content from the need to plan and teach this subject using my experience in research and noting students' difficulties of absorbing knowledge when the learning dialogue takes place only in the theoretical plane. From a narrative literature review, it discusses: (1) the conceptual modeling process for the elaboration of Knowledge Organization Systems (KOS), (2) examples from four KOS, (3) the processes of cataloging, indexing, and classification in information representation, (4) the contribution of technological resources to the

representation of knowledge and information on the Web, with a view to interoperability between systems and improved information retrieval. The paper concludes that the major challenge for the development of Knowledge and Information Representation on the Web is the large amount of loosely structured data and the increasing difficulty of accessing and retrieving from this large store. The envisioned solutions are based on (1) creation of ontologies to assist in providing content that needs to be better explicit on the web to be retrieved; (2) availability of information in various formats and modes of access; (3) virtual reference service to help users locate the best sources of information; (4) development of repositories of learning resources in different formats; (5) improved standardization to make content accessible to users with special needs,(6) cloud computing

Key-words: *Knowledge Organization. Knowledge Representation. Organization of Information. Semantic Web. Knowledge Organization Systems. MHTX Research Group.*

Recebido em 27.02.2020 Aceito em 27.02.2020

1 Introdução

A utilização de teorias e técnicas da área da Organização e Representação do Conhecimento e da Informação auxilia na modelagem de um domínio e sua representação, bem como na estruturação de um Sistema de Recuperação da Informação (SRI), tornando a recuperação do conhecimento registrado mais eficiente e eficaz. O surgimento de novas tecnologias, principalmente a Internet, demandou avanços que, muitas vezes, as bases teórico-metodológicas da área não conseguiram acompanhar. Assim, a supracitada área busca possíveis soluções por meio de aportes teóricos em áreas correlatas que contemplem a dimensão interdisciplinar desse campo de estudo, sem perder seu objetivo principal: tratar a informação para o usuário ter uma recuperação eficiente.

Um dos primeiros estudiosos a pensar o mundo e a categorizá-lo foi o filósofo Aristóteles, há mais de dois milênios, com sua inquietude de saber como o homem poderia representar o seu conhecimento. A teoria clássica de Aristóteles é considerada por alguns como a hierarquia perfeita, na qual as categorias são definidas apenas pelas propriedades comuns a todos os seus membros. Por conseguinte, a primeira preocupação em sistematizar o conhecimento, com o intuito de torná-lo

padronizado, via representação das ideias por símbolos que representassem a realidade visível pelo homem veio da Filosofia, mediante os estudos de Aristóteles (LIMA, 2016).

Dessa maneira, tornar-se-ia possível manifestar ideias unindo-as às coisas, por intermédio de um significado convencional, pois estariam integradas de forma que pudessem materializar o pensamento humano, dando aos objetos características e atributos que os representassem. Assim, Aristóteles, em seu *Organón* (ARISTÓTELES, 2010), sistematiza o saber humano propondo um sistema filosófico representado por dez categorias, a partir de proposições que se referem à determinação formal do ser, a saber: Substância, Quantidade, Qualidade, Relação, Duração, Lugar, Ação, Paixão ou Sofrimento, Maneira de ser, e Posição. Outras contribuições deixadas por esse autor foram as quatro distinções entre o gênero, o próprio, o acidente e a definição (MORUJÃO, 2006, p. 304) e o esquema tripartitivo de classificação, no qual o filósofo dividiu o conhecimento em ciências teóricas, práticas e poéticas, representando o pensar, o agir e o produzir, respectivamente (PIEIDADE, 1983). Esses ensinamentos de Aristóteles influenciaram o desenvolvimento dos estudos sobre a organização do conhecimento e são considerados o embasamento teórico para as classificações bibliográficas que começaram a surgir no século IV. A realidade, porém, não pode ser vista sob um único aspecto, advindo da sua percepção, geralmente adquirida por concepções e sensações do indivíduo. Surge aí a necessidade de criar mecanismos de padronização e representação desse conhecimento, tornando-o passível de ser posteriormente compartilhado consensualmente entre os indivíduos, por meio de seu recurso informacional. A aprendizagem humana baseia-se na capacidade de analisar, representar e organizar dados, informações e conhecimento, por isso que nós nos organizamos para recuperá-los.

Este artigo tem como objetivos: (1) apresentar uma revisão narrativa da literatura sobre teorias e técnicas da Organização e Representação do Conhecimento e da Informação, evidenciando seus diferentes propósitos e características, (2) verificar como essas teorias e técnicas podem contribuir para a Web Semântica, (3) analisar como as contribuições tecnológicas podem auxiliar na Organização e Representação do Conhecimento e da Informação, (4) condensar as reflexões sobre o ponto de vista pelo qual essas áreas são abordadas no âmbito da pesquisa no Grupo de Pesquisa MHTX – Modelagem Conceitual para Organização e Representação da Informação Hipertextual- MHTX e, finalmente, (5) dar uma visão geral do conteúdo que subsidia a prática pedagógica no contexto da disciplina Organização do Conhecimento e da Informação, ministrada por mim, no Programa de Pós-graduação Gestão e

Organização do Conhecimento (PPGGOC), da Escola de Ciência da Informação da UFMG.

A disciplina Organização do Conhecimento e da Informação apoia-se no tripé Fundamentos, Metodologia e Prática, conforme apresentado na Figura 1. No decorrer dos estudos, mesclam-se as três abordagens, culminando em um trabalho prático dos alunos, na modelagem conceitual de quatro diferentes Sistemas de Organização do Conhecimento (SOC) do tipo taxonomia, sistema de classificação, tesauros e ontologias.

Figura 1 – O tripé Fundamentos, Metodologia e Prática no ensino da disciplina OC



Fonte: Elaborado pela autora.

Parte-se do pressuposto de que, ao se estudar simultaneamente os fundamentos teóricos que suportam a modelagem e as técnicas, ao mesmo tempo em que se realiza a atividade prática, as características dos processos e dos instrumentos ficariam mais evidenciadas. As autoras Lima e Maculan (2017) apresentam detalhadamente um dos resultados dessas atividades no artigo "Estudo comparativo das estruturas semânticas em diferentes sistemas de organização do conhecimento". A necessidade de planejar e ensinar essa disciplina valeu-se da minha experiência em pesquisa e ensino, ao constatar as dificuldades de absorção do conhecimento tanto de meus orientandos quanto de meus alunos, quando se mantinha o diálogo de aprendizagem somente no plano teórico. Tendo desde o início pesquisado as teorias, técnicas e produtos que se aportam nas áreas de Biblioteconomia e Ciência da Informação (BCI) e Ciência da Computação, notei que, muitas vezes, em pesquisas exploratórias e práticas, necessita-se primeiramente verificar qual resultado será possível alcançar, para, posteriormente, descrever a metodologia adequada; pois, nesse caso, trabalhamos com procedimentos que dependem de experimentos exploratórios, antes de saber se poderão ser utilizados na

metodologia, ou mesmo criamos insumos para viabilizar a metodologia e chegar ao resultado planejado.

Nas próximas seções, apresenta-se primeiramente um panorama da organização do conhecimento e da informação; posteriormente, descreve-se sobre a representação da informação e do conhecimento, sobre as contribuições recíprocas das áreas, e as considerações finais.

2 Organização do Conhecimento e da Informação

A área de Organização do Conhecimento (OC) tem suas raízes históricas na base teórica da Classificação, sendo que, às vezes, pode ser vista como uma área de conhecimento, a qual classifica o conhecimento filosófico, que se inicia com o modelo categorial proposto por Aristóteles; e, às vezes, como classificação bibliográfica, que possui processos e instrumentos para representar conteúdo temático de documentos e aumentar a eficácia na recuperação de informações.

Entretanto, a Organização do Conhecimento, enquanto área de estudo, foi estabelecida mediante estudos realizados por Dahlberg (1993) e pelos membros da *International Society of Knowledge Organization* (ISKO), que, a partir de 1973, decidem reconhecer a OC como área, sugerindo que o conceito "classificação" passasse a ser interpretado tal qual um método que agrupa objetos semelhantes. A partir daí, sob a influência de Dahlberg, o conceito "classificação" passa ser interpretado como um método de classificação, que agrupa objetos semelhantes. Dahlberg (2006) define Organização do Conhecimento como

a ciência que estrutura e organiza sistematicamente unidades do conhecimento (conceitos) segundo seus elementos de conhecimento (características) inerentes e a aplicação desses conceitos e classes de conceitos ordenados a objetos/assuntos (DAHLBERG, 2006, p. 12).

Pode-se dizer que organizar o conhecimento é reunir o que conhecemos em uma estrutura sistematicamente organizada. Assim, como a Classificação se relaciona com a Organização do Conhecimento (OC), a Organização da Informação (OI) também tem uma relação estreita com a OC. O termo Organização e Representação do Conhecimento e da Informação possui uma dispersão terminológica, sendo que algumas vezes são utilizados como complementares, em outras são utilizados de maneira distinta. Esse fato deve-se, em sua maior parte, às diversas áreas que abordam esses conceitos (LARA, 2011, p. 94). Isso pode ser verificado na própria literatura brasileira; frequentemente, há o uso simultâneo dos termos Organização da Informação e Organização do Conhecimento.

Enquanto a Organização do Conhecimento está relacionada à modelagem de um domínio de conhecimento com intuito de construir

representações do conhecimento, tais como os Sistemas de Organização do Conhecimento (SOC), a Organização da Informação compreende processos de representações de objetos informacionais, que se destinam, prioritariamente, a contribuir para uma recuperação eficaz por parte dos usuários. A Organização do Conhecimento produz representações sistematizadas de conhecimento, que podem ser utilizadas na Organização da Informação, para padronizar as representações dos conteúdos dos recursos informacionais, visando facilitar a recuperação e o uso deles. Entende-se que a organização do conhecimento é condição necessária para a organização da informação, ou mais especificamente dos recursos informacionais, sob o aspecto dos seus conteúdos.

A área da Organização da Informação (OI) compreende todos os estudos relacionados aos processos e instrumentos utilizados na organização de recursos informacionais de qualquer natureza, no intuito de viabilizar o atendimento às necessidades de informação de uma determinada comunidade de usuários. Quando se trata da Organização da Informação, existem definições a partir de diferentes abordagens de acordo com as escolas de pensamento de diversos autores. Em relação à informação, Le Coadic (2004) destaca que a

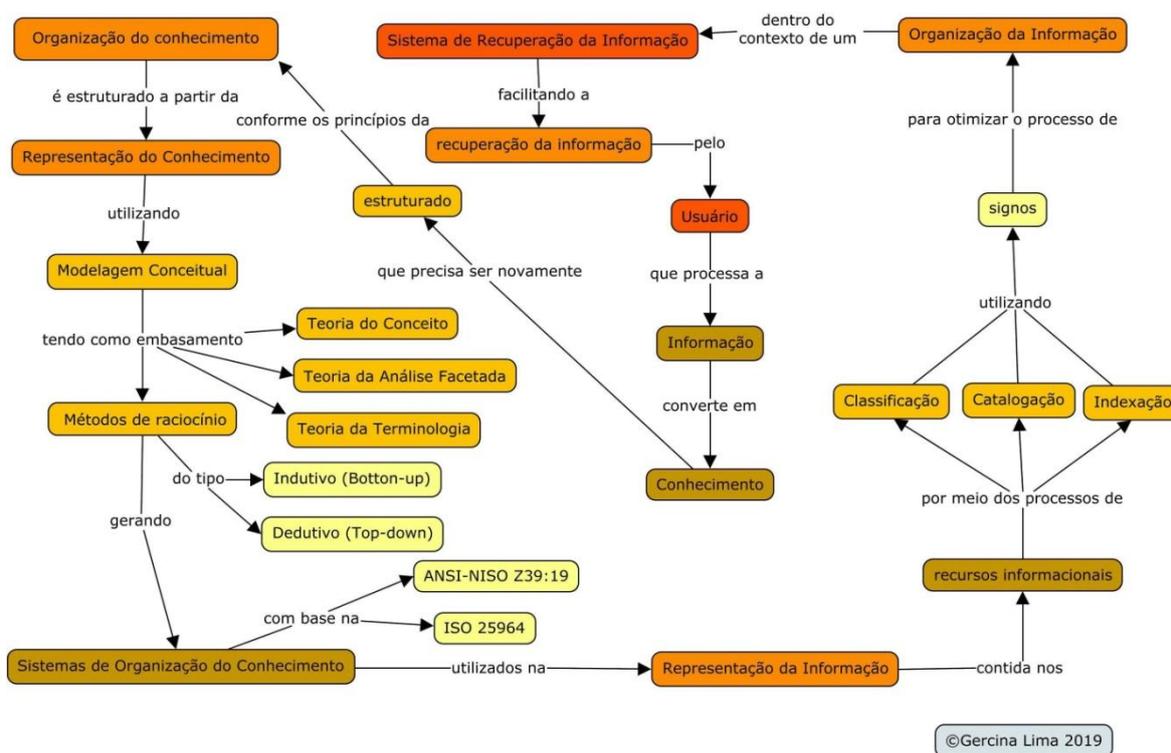
Informação é um conhecimento inscrito (registrado) em forma escrita (impressa ou digital), oral ou audiovisual em um suporte... é um significado transmitido a um ser consciente por meio de uma mensagem inscrita em um suporte espacial-temporal: impresso, sinal elétrico (LE COADIC, 2004, p. 4).

Assim, a informação só é valiosa à medida que é estruturada, representada e recuperada. Dahlberg (2006) define Organização da Informação (OI) como a ordenação de objetos de forma a criar um elo entre o objeto de uma área e sua própria atividade.

Pode-se dizer que essas duas abordagens se complementam, pois, enquanto a área de OC organiza o conhecimento a partir da modelagem de domínios, representando sua estrutura com base na descrição de conceitos e dos relacionamentos semânticos entre eles, na área de OI ocorre o processo de representação do conteúdo de um documento específico, o qual tem o conhecimento registrado, para que ele seja posteriormente recuperado pelo usuário em um Sistema de Recuperação da Informação (SRI). O círculo informacional se completa quando a informação é convertida em conhecimento, promovendo o desenvolvimento e a comunicação, e gerando novos armazenamentos de conhecimentos registrados e retroalimentando o ciclo, pois é esse processo, o qual vai concatenando e assimilando as ideias, que permite a elaboração de um conhecimento novo, e, em muitos casos, buscando aportes teóricos em mais de uma área do conhecimento.

De outra perspectiva, pode-se inferir que se organiza um domínio do conhecimento para representá-lo, fundamentado em teorias e mediante técnicas de representação do conhecimento, gerando instrumentos de representação do conhecimento, do tipo Sistema de Organização do Conhecimento (SOC). Esses instrumentos auxiliam na representação da informação registrada em recursos informacionais em um SRI, por meio dos processos de representação descritiva e temática (catalogação, indexação e classificação), para que seja representado o conhecimento daquele domínio específico. Em síntese, organizar o conhecimento é olhar para um domínio e sistematizá-lo para sua representação; por outro lado, a organização da informação utiliza-se desses instrumentos para representar o conteúdo dos recursos informacionais, no âmbito de um Sistema de Recuperação da Informação (SRI), conforme representado na Figura 2.

Figura 2 – A Organização do Conhecimento e da Informação com vistas à Recuperação da Informação



Fonte: Elaborado pela autora.

Observa-se que os diferentes pontos de vista existentes, juntamente com a falta de entendimento das bases teóricas e a inexistência de uma terminologia unificada que possa mediar esses diferentes olhares dificultam a consolidação da Organização do Conhecimento e da Informação como área do conhecimento.

Na próxima seção, descrevem-se os processos de representação do conhecimento e da informação, nos quais esse procedimento ocorre em momentos distintos.

3. Representação do Conhecimento

A Representação do Conhecimento pode ser vista como uma abordagem multidisciplinar em diversos campos temáticos. Do ponto de vista filosófico, Mora (1996, p. 629), em seu Dicionário de Filosofia, refere-se ao conceito "representação" como "[...] diversos tipos de apreensão de um objeto intencional". O autor acrescenta que Aristóteles referia-se à representação como um fenômeno da fantasia (intelectual ou sensível); os estoicos a viam como impressão; para os escolásticos, era como um objeto de apresentação intencional; Descartes a via como imaginação; Spinoza distinguia-a entre apreensão sensível e conceitual; no sentido de Leibniz, era como a perfeição; Locke e Hume veem a representação como a ideia; e Schopenhauer referia-se à "[...] forma do mundo dos objetos como manifestações da vontade [...]" (MORA, 1996, p. 630). Apesar de tantas interpretações, a representação é uma questão relevante no processo informacional, visto que cabe a ela a função mediadora entre a apreensão do real e a sua materialização mediante formas simbólicas.

A Representação do Conhecimento esbarra na dificuldade consensual de representá-lo devido aos modos de fazê-lo advindos de diferentes abordagens das áreas do conhecimento, tais como Ciência da Informação, Ciência da Computação, Linguística, Ciências Cognitivas, Inteligência Artificial, Administração, dentre outras. No entanto, depara-se com modelos diferenciados para representação do conhecimento, que trazem os propósitos e as características de acordo com o ponto de vista e necessidades de cada área.

Assim, o termo Representação do Conhecimento tem sido definido de forma diferente em vários campos científicos. Uma visão geral sofisticada, com base em grande parte teórica (filosófica e lógica) é dada por Sowa (2000). Pesquisas em Inteligência Artificial (IA) discutem métodos de modelagem de conhecimento ou engenharia do conhecimento (STUDER; BENJAMINS; FENSEL, 1998). Outro ponto de vista diferente é obtido por meio de estudos linguísticos e sua elucidação da Representação do Conhecimento (LÖBNER, 2002). Na Ciência Cognitiva, a Representação do Conhecimento baseia-se no modelo de mundo do indivíduo, ao contexto e à maneira de cada um ver o mundo consistindo no somatório de diferentes estruturas do Conhecimento (MEY, 1982), (GARDNER, 1996). Na perspectiva da Ciência da Informação, os autores Cleveland e Cleveland (2001), Lancaster (2003), Stock e Stock (2008) tratam a Representação do Conhecimento como uma abordagem para solucionar

problemas, tais como: estruturar, armazenar informações e como encontrá-las e recuperá-las com precisão e eficácia.

Nota-se que a Representação do Conhecimento tem sido estudada em diversas ciências, com o intuito de entender como ocorre o processamento do conhecimento na mente humana e como materializar este saber. A Representação do Conhecimento na forma simbólica é uma questão que preocupou o mundo da Documentação desde sua origem. Ela está sempre relacionada com as formas de expressão da informação, conforme destacado por Vickery (1978). O problema é relevante agora em muitas outras situações além dos documentos e índices. A estrutura dos dados nos programas de computadores, a estrutura sintática e semântica da linguagem natural, a representação do conhecimento na inteligência artificial, os modelos da memória humana: em todos esses campos, é necessário decidir como o conhecimento pode ser representado de forma que as representações possam ser manipuladas (VICKERY, 1986, p. 145).

No contexto específico da Organização do Conhecimento, Barité (2015) destaca que Representação do Conhecimento:

é uma parte da organização do conhecimento na qual estão inseridos todos os processos de simbolização notacionais e conceituais do conhecimento humano dentro do campo específico de cada área. Classificação, indexação e todos os aspectos linguísticos e computacionais relacionados com a tradução simbólica do conhecimento estão incluídas na representação do conhecimento (BARITÉ, 2015, p. 136).

Assim sendo, o autor considera o símbolo como uma representação codificada ou conceitual de uma notação (uma notação ou um termo), que é o resultado dos processos de análise e síntese de um documento, imagem, figura, ou sinal que expressa um conceito em um convencional caminho. Para tanto, a Representação do Conhecimento engloba conceitos e estruturas, estabelecendo relações estruturais e sistemáticas de associação e distinção. San Segundo (2004, p. 109) destaca que "se conhecimento é um processo de integração, representação de conhecimento será: conceitos, teorias, modelos, formatos, descrições e estruturas que têm um significado de dimensão da simbolização da informação e, mais atualmente, de informação eletrônica".

O ato de representar é o processo de substituir uma realidade por intermédio de um significado simbólico que remete à ideia daquela realidade novamente. Nesse processo de representação, a modelagem requer o deslocamento do objeto do mundo real para um espaço de representação. Para essa representação, podem-se utilizar dois métodos de raciocínio: o método indutivo e o método dedutivo, denominados, na Ciência da Computação, por "*bottom-up*" e "*top-down*", respectivamente

(CAMPOS, 2001), que se baseiam nos princípios da Garantia Literária¹ e na Garantia do Usuário².

Campos, Souza e Campos (2003, p. 10) destacam que esses métodos contribuem na representação do conhecimento, pois:

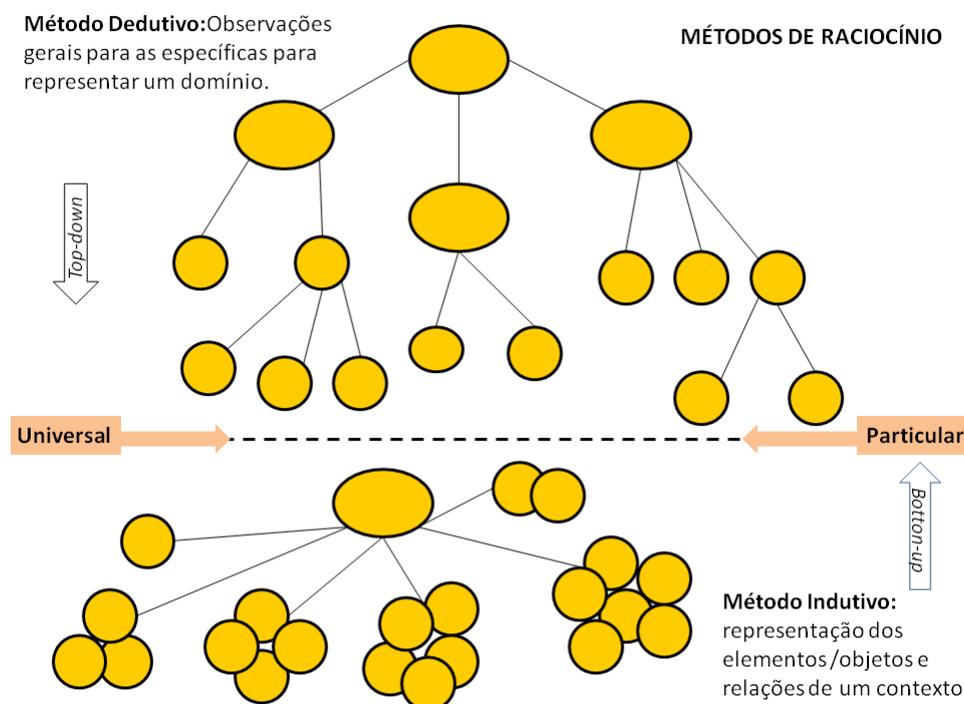
O método indutivo possibilita a elaboração de modelos, partindo, desde o início, da representação dos elementos/objetos e relações de um contexto. Já o método dedutivo propõe que se elaborem mecanismos de abstração para pensar primeiramente o domínio/contexto, independentemente de pensar os elementos e suas relações – esta seria uma etapa posterior.

O método indutivo parte da coleta e da análise detalhada de casos, objetos, ou expressões da linguagem especializada, para obter uma estrutura organizada de conceitos, para serem agrupados de acordo com as classes a que pertencem, conforme seus atributos. No método dedutivo, segue-se um processo de divisão lógica do conhecimento, partindo de categorias gerais com os conceitos mais gerais aos conceitos mais específicos, criando um conjunto de classes hierárquicas, no qual cada subclasse de elementos é um conjunto da classe imediatamente anterior, conforme demonstrado na Figura 3.

¹ Um princípio que se utiliza do tema descoberto no documento para encontrar termos que podem ser utilizados em um sistema de classificação (BARITÉ, 2011, p. 78).

² A garantia de usuário, ou garantia de uso, que “tem como foco a formação de vocabulários e indexação atenta às necessidades de informação dos usuários, verificando tendências de busca em sistemas de informação e coletando demandas informacionais” (LIMA, 2014, p. 20).

Figura 3 – Métodos de raciocínio – dedutivo e indutivo



Fonte: Elaborado pela autora.

A Representação do Conhecimento tem como produtos os instrumentos terminológicos. Assim, no seu processo, é preciso estruturar um domínio e representá-lo mediante expressões abreviadas que formarão os Sistemas de Organização do Conhecimento (SOC), que são definidos por Bräscher e Café (2008) como sistemas conceituais que são criados para sistematizar os conceitos e suas relações, dentro de um determinado domínio, e também para a recuperação de informações.

No âmbito da Biblioteconomia e Ciência da Informação (BCI), os principais fundamentos teórico-metodológicos utilizados pelos classificacionistas para modelagem e representação de um domínio do conhecimento são a Teoria da Análise Facetada (TAF) criada por Ranganathan, em 1967, e a Teoria do Conceito, criada por Dahlberg (1978a), que trazem forte influência da lógica aristotélica para a Representação do Conhecimento.

Shiyali Ramamrita Ranganathan (1892-1972), matemático indiano, tornou-se bibliotecário depois de constatar que a maioria dos assuntos tratados em cinco periódicos diferentes era formada de assuntos compostos e, também, pela sua insatisfação com os sistemas de Classificação Decimal de Dewey (CDD) e Classificação Decimal Universal (CDU). Em seus estudos, Ranganathan (1967) considerou o conhecimento como um elemento multidimensional: as interligações de cada conceito espalham-se em muitas direções e, frequentemente, cada assunto

Ranganathan descreve o assunto como “corpo de ideias organizadas ou sistematizadas, cuja extensão e intensão devem ser coerentemente com o domínio de interesse e confortavelmente ajustadas à competência intelectual e campo especializado de um indivíduo” (RANGANATHAN, 1967, p. 82). Todo assunto é oriundo de um assunto básico somente, e seu componente vem de uma ou mais ideias isoladas, formando, assim, o assunto composto. O assunto básico é um assunto sem nenhuma ideia isolada como componente. A ideia isolada (isolado) é alguma ideia, ou complexo de ideias, moldada(o) para formar um componente de assunto; mas, sem seu contexto, ela não é considerada um assunto. Para o autor, ideia é um produto do pensamento, da reflexão, da imaginação, que passou pelo intelecto, integrando, com a ajuda da lógica, uma seleção de conjuntos de percepções, que são diretamente apreendidas pela intuição e depositadas na memória. A informação seria dada no momento em que uma ideia é comunicada por outros ou obtida a partir do conhecimento pessoal.

De acordo com Ranganathan (1967), o conhecimento é representado pela totalidade das ideias armazenadas pela memória humana, passando por processos sensoriais, permeados pelas experiências cognitivas de cada indivíduo, formando, assim, os conceitos como produto do pensamento, representando a “unidade de pensamento de qualquer nível de complexidade”. Nessa perspectiva, o conhecimento é considerado um elemento multidimensional: as interligações de cada conceito espalham-se em muitas direções e, frequentemente, cada assunto constitui-se de uma síntese de vários conceitos múltiplos ligados, o que caracteriza essa abordagem como analítico-sintética.

Na década de 1970, a autora alemã Ingetraut Dahlberg apresenta a teoria do conceito (1978a), que tem por base os princípios aristotélicos, os princípios analítico-sintéticos ranganathianos, além de buscar aportes teóricos na Teoria Geral da Terminologia (TGT) de Wüster (1998). A teoria do conceito tem como proposição elementar e fundamental a afirmação de que o conceito é uma representação do conhecimento, e apresenta princípios que podem auxiliar na determinação do conceito e de suas relações.

Para Dahlberg (1978a), a ideia de conceito está ligada às características dos objetos. A autora considera conceito como uma série de enunciados (características) verdadeiros sobre um objeto. Para ela, o conceito é uma unidade do conhecimento. Na elaboração de um conceito, são colocados elementos que se articulam numa unidade estruturada. Para que essa unidade seja, de fato, estruturada, a autora coloca a necessidade de análise e síntese dos enunciados verdadeiros de um conceito, que podem apresentar-se sob a forma de característica ou categoria (hierarquia de características) (DAHLBERG, 1978a, p. 102). Para

a construção do conceito, Dahlberg (1978b) propõe um modelo, no qual essa unidade do conhecimento é elaborada atendendo a essas afirmações verdadeiras: (A) sobre um objeto específico (B) representado por signo linguístico (C) com fins de comunicação. Assim, a autora sugere a forma de um triângulo para representar esses componentes, em que: o vértice superior é o objeto, o vértice esquerdo é o significado, e o vértice direito é o signo linguístico (termo).

Enquanto para Ranganathan a ideia/conceito é o produto do pensamento, salientando a característica intuitiva e dedutiva da Teoria da Análise Facetada (TAF), Dahlberg aponta que o conceito é uma unidade do conhecimento, evidenciando os aportes teóricos da Teoria da Terminologia utilizados na construção do conceito a partir de seus atributos e sua representação dentro de um campo do conhecimento.

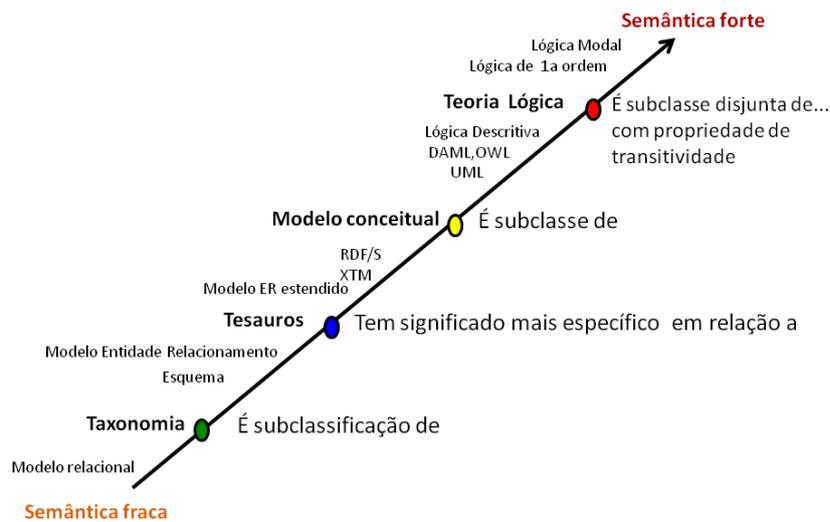
Assim como Aristóteles e Ranganathan, Dahlberg apresenta as categorias que permitem representar o domínio específico, por meio da possibilidade de listar todas as características possíveis de um conceito, tendo por base as categorias aristotélicas, a saber: Matéria (substância), Qualidade, Quantidade (extensão), Relação, Processo (atividade), Modo de ser, Passividade, Posição, Localização (lugar) e Tempo.

Os fundamentos advindos da Linguística e Terminologia também são importantes na Representação do Conhecimento para elaboração dos SOC, porque é preciso considerar esses instrumentos como um sistema de conceitos, cujos elementos são definidos a partir das relações de equivalência, as relações lógicas (gênero-espécie, analítica e de oposição), as relações ontológicas (partitiva, sucessão e material-produto), as relações de efeito (causalidade, instrumental e descendência) entre outras, que formam a "estrutura" de uma linguagem comunicativa. A Terminologia contribui com os princípios teóricos para a criação de sistemas de conceitos bem estruturados. Isso acontece seja com a Teoria Geral da Terminologia (TGT), iniciada por Eugen Wüster (1998), que se baseia nas dicotomias "termo e palavra", seja com os princípios da Teoria Comunicativa da Terminologia (TCT), proposta por M. T. Cabré (1999). A última traz novos subsídios teóricos para identificar as funções (pragmáticas, semânticas e sintáticas) dos termos e os aspectos de forma (variação linguística) e conteúdo, além da relação existente entre o termo (como unidade de conhecimento) e o conceito que representa.

Os SOC são instrumentos com diferenciados níveis de estruturação e formalidade, com diversificadas abordagens, que atendem a aspectos e objetivos variados, do ponto de vista teórico e metodológico. Nesse conjunto de instrumentos, incluem-se as taxonomias, os sistemas de classificação, tesouros, ontologias, entre outros, apontados na literatura e organizados conforme Hodge, 2000; Smith, Welty, 2001; Hill, L. et. al. 2002 ; Daconta, Obrst, Smith, 2003; McGuinness, 2003; Zeng, 2008;

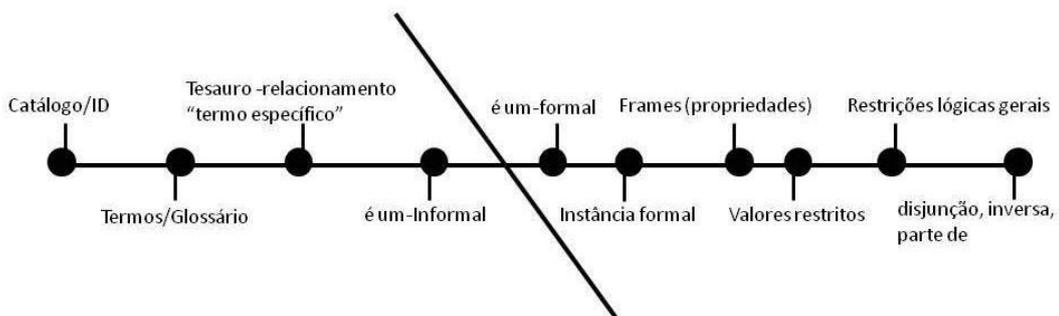
Soergel, 2008; Guarino, Oberle, Staab, 2009; Bratková, Kučerová, H., 2014. Essas estruturas são organizadas de diferentes abordagens por esses autores. As figuras seguintes (6, 7, 8, 9) apresentam alguns espectros de acordo com o grau de formalidade.

Figura 6 – Espectro Ontológico baseado na semântica formal



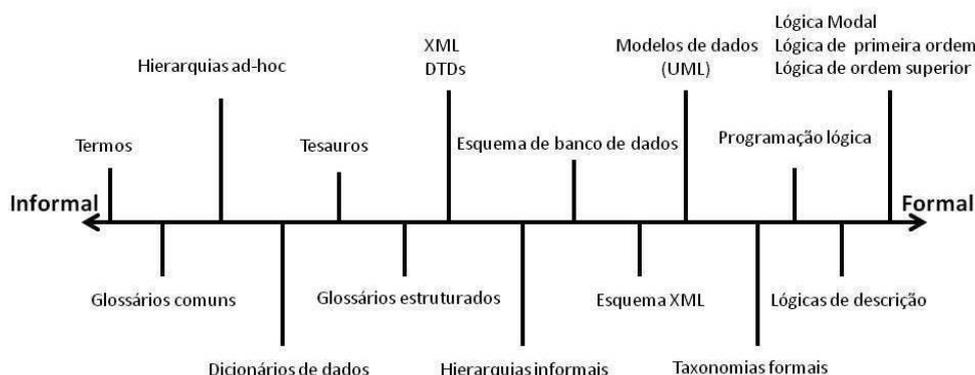
Fonte: Traduzido de DACONTA, OBRST, SMITH (2003) e OBRST (2003).

Figura 7 – Espectro Ontológico baseado na semântica formal



Fonte: Traduzido de McGUINNES (2003).

Figura 8 – Espectro Ontológico baseado na formalidade



Fonte: Traduzido de GUARINO, OBERLE, STAAB (2009).

FIGURA 9 – Tipologia dos sistemas de organização do conhecimento de acordo com a sua força semântica

Força semântica	Categoria		Tipos		A	Lista de palavras		
↓	Vocabulários estrutura linear	Organização de termos	1	lista de termos	A	Lista de palavras		
			2	stopwords/lista de exclusão				
			3	vocabulário controlado/dicionário				
			Classificação Estrutura hierárquica	Organização de termos	4	lista, lista de códigos	B	A + definição
					5	glossário		
					6	terminologia		
					7	anel de sinônimos	C	B + equivalência
Redes de Conceitos estrutura em rede	Organização de conceitos	8	gazetters	D	C + termos referidos			
		9	lista/arquivo de autoridades					
		10	lista de cabeçalho de assunto					
		Redes de Conceitos estrutura em rede	Organização de conceitos	11	esquema de categorização	E	D + hierarquias	
				12	taxonomia			
				13	esquema de classificação			
				14	tesauro	F	E + associações	
				15	rede semântica			
				16	ontologia			G

Fonte: Adaptado de BRATKOVÁ, E.; KUČEROVÁ, H (2014, p.3).

Não é objetivo, neste artigo, descrever ou avaliar as abordagens apresentadas por esses autores, ou aprofundar em discussões nessa perspectiva. Optou-se por apresentar, brevemente, neste estudo, quatro desses tipos de SOC por serem os instrumentos que são construídos como prática pedagógica, no âmbito da disciplina Organização do Conhecimento e da Informação, do PPGOC-ECI-UFMG e por serem os mais utilizados na

Biblioteconomia e Ciência da Informação (BCI), a saber: Taxonomia, Sistema de Classificação Bibliográfico, Tesouro e Ontologia.

A Taxonomia é um conjunto de termos estruturados de forma hierárquica e que representam as áreas nas quais são aplicados. Ela surgiu na área da Biologia, com Linnaeus (Taxonomia de Lineu), no século XVIII, que criou um sistema hierárquico dos seres vivos. Seus fundamentos têm por base as Teorias da Classificação, sobretudo da Classificação Facetada, que dão subsídios para a classificação sistemática e ordenação de termos, apoiada nos princípios da Lógica. Sua estrutura é composta de um conjunto de termos arranjados em hierarquias representando o domínio modelado, que organiza conceitos, segundo as suas semelhanças e diferenças. É um instrumento pré-coordenado, pois estabelece a formação dos vários rótulos (assuntos simples ou compostos) no momento de sua construção, fazendo o controle da ambiguidade, da sinonímia e dos relacionamentos hierárquicos. Entre suas funções, está o mapeamento do conhecimento de um domínio e o estabelecimento de um rótulo para as informações disponibilizadas, sendo um mecanismo de acesso em ambiente digital (LIMA; MACULAN, 2017).

No âmbito da Ciência da Informação, Vickery (1975, p. 10) definiu a taxonomia como um instrumento utilizado para "organizar todas as entidades de um universo em uma simples hierarquia: uma grande árvore na qual cada entidade pertence apenas a uma única classe, cada espécie a um único gênero, e assim por diante".

O conceito de Taxonomia sofreu transformações, e, na contemporaneidade, o termo é aplicado em sentido mais abrangente, sendo utilizado em meio digital para criar estruturas de um domínio e rótulos para metadados, que permitem organizar sistematicamente itens de informação, auxiliando na recuperação de informação. Tem sido utilizado como mapas conceituais de um domínio, orientação permitida pela estrutura de termos que dão acesso aos tópicos, para exploração (navegação) em um serviço de recuperação e gestão de informações corporativas, como, por exemplo, no e-comércio, em bibliotecas digitais. Nesses serviços, por vezes encontramos uma Taxonomia Facetada, na qual um conteúdo não está restrito a uma única dimensão, oferecendo diferentes opções de busca ao usuário, possibilitando a navegação pela estrutura facetada.

Por sua vez, o Sistema de Classificação Bibliográfico é definido por Langridge (1977) como um mapa completo de qualquer área do conhecimento, mostrando todos seus conceitos e suas relações. Os sistemas de classificações bibliográficas são elaborados tendo como base as classificações filosóficas; porém, são utilizados na organização de documentos, no arranjo para sua disposição física e na sua função de possibilitar a recuperação. Suas finalidades principais são (1) a ordenação

dos documentos nas estantes ou nos arquivos; (2) a ordenação das referências bibliográficas ou das fichas em catálogos (PIEDADE, p. 65, 1983).

Esses sistemas são caracterizados como sistemas verbais-simbólicos, geralmente constituídos por três elementos: (1) um esquema de classificação, agrupando elementos sistematicamente, por suas semelhanças; (2) uma notação da classificação (número de chamada); (3) um índice alfabético, que facilita a busca do usuário. Sua estrutura, por divisão de classes, baseia-se em dois princípios básicos: (a) de divisão, que acata critérios pré-estabelecidos, obedecendo a uma característica específica de divisão; (b) de classes simples e classes compostas, em que a classe simples identifica uma única característica, na qual não cabem mais subdivisões; e a classe composta identifica mais de uma característica, em que podem caber mais subdivisões.

Existem diversos sistemas de classificação bibliográfica; citam-se alguns, a saber: (1) Classificação Decimal de Dewey (CDD), de 1876 – quase-enumerativos, com tabelas de assunto e subdivisões comuns; (2) Classificação da Biblioteca do Congresso, de 1902 – enumerativos, com tabela única; (3) Classificação Decimal Universal (CDU), de 1905 – quase facetados, com tabelas enumerativas, tabelas de subdivisões comuns e de subdivisões especiais; (4) Classificação dos Dois Pontos, de 1933 – analítico-sintéticos: sistemas facetados com assuntos básicos, tabelas de subdivisões comuns, tabelas auxiliares especiais, e sequência de combinação entre os conceitos, regida pelos princípios a serem deliberados.

Há ainda o tesauro. A origem etimológica é do latim *thesaurus*, que, por sua vez, originou-se do grego *thesaurós*. Em ambas as origens, o sentido é idêntico ao tesauro em português. Gomes (1990) afirma que o tesauro é uma “linguagem documentária dinâmica que contém termos relacionados semântica e logicamente, cobrindo de modo compreensivo um domínio do conhecimento” (GOMES, 1990, p. 16).

No âmbito da Biblioteconomia e Ciência da Informação (BCI), a Norma ISO 25964-1 (2011) define tesauro como um vocabulário controlado e estruturado no qual os conceitos são representados por termos descritores, preferidos e não preferidos, formando um sistema de conceitos inter-relacionados (ISO 25964-1, 2011, p. 12). Como função, a Norma ISO 25964-1 (2011) destaca que a

proposta de um tesauro é guiar tanto o indexador como o pesquisador para selecionar o mesmo termo preferido ou uma combinação de termos preferidos para representar determinado objeto. Por essa razão, um tesauro é otimizado para a navegabilidade humana e para a cobertura terminológica de um domínio (ISO 25964-1, 2011, p. 12).

Essa estrutura normaliza a terminologia, mediante a diferenciação dos termos preferidos (descritores) dos não preferidos (não descritores), o que possibilita a explicitação das relações semânticas entre conceitos, formando uma estrutura pós-coordenada a partir de relações conceituais, que são definidas pelas relações: (a) hierárquicas (gênero-espécie; todo-parte); (b) de equivalência (sinônimos); (c) associativas (ligações semânticas). De forma geral, esses relacionamentos são explicitados a partir de símbolos, tais como: (1) USE: indica sinonímia e que há outro termo preferencial; (2) UF (usado para): indica sinonímia e que é o termo preferencial; (3) BT: indica a classe mais geral do termo; (4) NT: indica termo específico; (5) RT: indica termo em relações associativas diferentes.

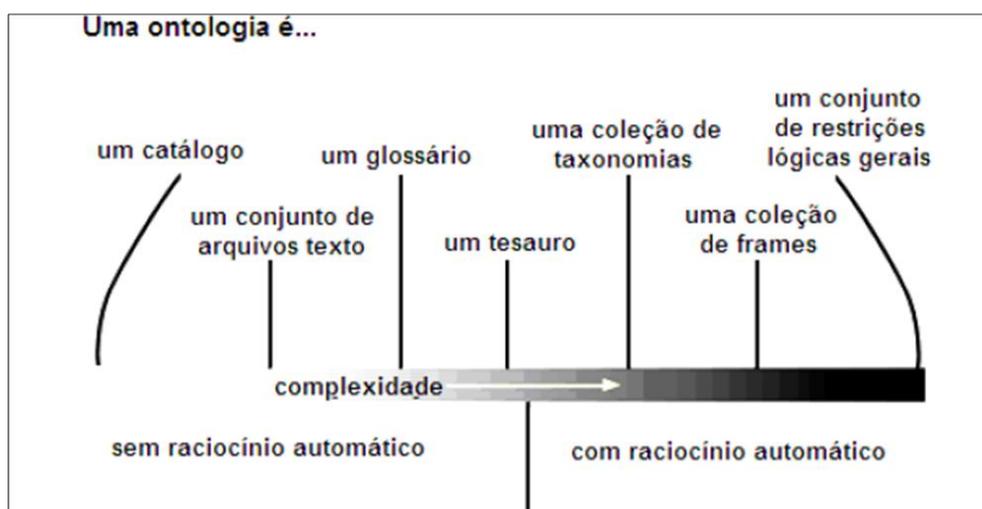
A ontologia é estudada em diversas áreas do conhecimento e tem sua origem na filosofia; e a definição de ontologia como uma ciência se deve a Aristóteles (1978). Na Ciência da Informação e na Ciência da Computação, a ontologia é estudada como instrumento de representação do conhecimento, em geral, como um artefato computacional. Na BCI, segundo Vickery (1997) e Soergel (1999), o termo ontologia começou a ser utilizado no final da década de 1990, principalmente por pesquisadores da área de Organização do Conhecimento, no sentido de ser um mecanismo de representação do conhecimento, com uso expresso da semântica. Vickery (1997, p. 284) foi o primeiro autor na área de Ciência da Informação a estudar sobre ontologia; ele a define como “uma formalização sistemática de definições de conceitos, relacionamentos e regras que capturam o conteúdo semântico de um domínio em um formato legível por máquina”. Campos (2010) reafirma o interesse atual da Biblioteconomia e Ciência da Informação (BCI) pela ontologia, explicando que esse instrumento define um vocabulário para ser compartilhado em um domínio em particular, incluindo a definição dos conceitos básicos e as relações que há entre eles, de forma inteligível pela máquina.

Do ponto de vista da Ciência da Computação, Tom Gruber, em 1993, traz uma das primeiras contribuições em relação à definição de uma ontologia considerando-a como “uma especificação explícita de uma conceitualização”. Essa definição, porém, foi considerada muito geral, deixando grande espaço para outras interpretações e, mesmo posteriormente, com a definição de Guarino (1998)⁵, mais esclarecedora e formalizada, continuaram a proliferar novas definições sobre ontologia; tanto que, em 1999, os autores Welty, Lehmann, Gruninger e Uschold (1999) apresentam um amplo espectro de artefatos de informação que já

⁵ Uma teoria lógica correspondente ao significado pretendido de um vocabulário formal, ou seja, seu compromisso ontológico com uma conceitualização particular de mundo (GUARINO, 1998).

foram classificados como ontologias, conforme apresentados na Figura 10. Smith e Welty (2001) esclarecem que esses sistemas de informação tão simples quanto um catálogo, com um código único, ou tão complexos quanto sistemas de informações, que usam os axiomas de primeira ordem, satisfazem à definição de Gruber (1993), pois todos estão sob a mesma aceção de ontologia.

Figura – 10 Diversas acepções de ontologia



Fonte: Traduzido de Welty, Lehmann, Gruninger e Uschold (1999).

Pode-se afirmar que a ontologia é uma lista de conceitos ou entidades dentro de um domínio específico, os quais podem ser estruturados de forma hierárquica, por meio de relações semânticas explicitadas formalmente, em meio informatizado. A estrutura da ontologia nem sempre é a mesma, e, para Gruber (1992), os componentes básicos são: (1) termos e definições, que explicitam e formalizam o sentido dos conceitos; (2) classes, as quais são organizadas em estrutura conceitual hierárquica; (3) relações, que representam a relação entre os conceitos e classes; (4) axiomas, que representam sentenças verdadeiras; (5) instâncias, que representam a especificidade dos dados; (6) atributos, que descrevem características de conceitos e instâncias. Todos esses componentes devem possuir uma representação explicitamente formal para que a ontologia seja processável por máquina, para controlar a ambiguidade, a sinonímia, os relacionamentos hierárquicos e associativos, os relacionamentos complexos (regras e axiomas) e são exclusivamente para uso tecnológico.

Entre as tecnologias utilizadas na construção de ontologias, estão as linguagens *Resource Description Framework (RDF)*, *Resource Description Framework Schema (RDF-S)*, *Web Ontology Language (OWL)*, *Linguagem eXtensible Markup Language (XML)*, e o *Protocol and RDF Query Language (SPARQL)*.

Os Sistemas de Organização do Conhecimento são utilizados no contexto da Organização da Informação com o intuito de organizar e representar para o armazenamento dos recursos informacionais (documentos), com o objetivo de sua recuperação e comunicação, seja por meio digital ou não.

Nesse prisma, Zhonghong, Chaudhry, Khoo (2006) apontam as diferenças entre esses instrumentos quanto a suas formas, escopo, objetos tratados, funções, estruturas, termos e foco, conforme pode ser observado no Quadro 1.

Quadro 1 – Diferenças entre classificação, tesauros, taxonomias e ontologias

Formas	Escopo	Objetos tratados	Funções	Estrutura hierárquica	Termos	Foco
Classificação	Comunidade de bibliotecários e disciplinas acadêmicas	Coleções	Classificação e arquivamento	Unidimensional e uso combinado de notações	Classes	Em conteúdo
Tesauros	Ambiente online e comunidade acadêmica	Documentos	Indexação e pesquisa	Relações com termos da web	Termos	Em conteúdo
Taxonomias	Ambiente Web e organizacional	Recursos digitais	Categorização, busca e navegação	Estrutura dinâmica	Categorias	Em usuários
Ontologias	Interoperabilidade entre sistemas computacionais/ Inferência através do computador	Recursos digitais	Categorização para troca de informação entre sistemas	Estrutura dinâmica	Termos, categorias e classes	Sistemas/Em conteúdo

Fonte: Adaptação de Zhonghong, Chaudhry, Khoo (2006), citado em PONTES JUNIOR, J.; CARVALHO, R. A.; AZEVEDO, A. W. (2013)⁶.

3.1 Representação da Informação

Nas sociedades orais, não existia nenhum modo sistematizado de armazenar as representações para futura reutilização nesse tipo de comunicação, pois não havia a escrita. A escrita foi um dos mais importantes desenvolvimentos técnicos do ser humano. Diversos tipos de sistemas de sinais gráficos, incluindo o alfabeto, tornaram-se os instrumentos principais da escrita. A comunicação escrita e o modo de

⁶ PONTES JUNIOR, João de; CARVALHO, Rodrigo Aquino de; AZEVEDO, Alexander William. Da recuperação da informação à recuperação do conhecimento: reflexões e propostas. *Perspectivas em Ciência da Informação*, v.18, n.4, p.2-17, out./dez. 2013.

transmissão dos textos sofreram profundas mudanças com o advento da imprensa. Para McGarry (1999, p. 73), a escrita em suportes artesanais foi a tecnologia de comunicação mais avançada, desde o quarto milênio a.C. até a invenção da imprensa e tipografia na Mogúncia (c.1450), quando Johann Gutenberg compôs, com tipos móveis, o texto da Bíblia, o primeiro livro a passar do manuscrito para o papel impresso. A representação documentária ocorre desde as primeiras tentativas de registro do conhecimento em suportes que pudessem ser armazenados e, posteriormente, recuperados para utilização.

Observa-se que os avanços que ocorreram desde a década de 1950 até os dias atuais foram relevantes e marcaram o desenvolvimento tanto na forma de armazenamento e representação quanto na recuperação da informação. De acordo com Bräscher e Café (2008), o produto desse processo descritivo é a Representação da Informação (RI), entendida como um conjunto de elementos que representam os atributos de um objeto informacional específico. Hjørland (2008) afirma que a Representação da Informação compreende a descrição (de conteúdos, características e objetivos) de documentos e a organização dessas descrições de forma que possam ser acessados e recuperados para atender a necessidades e utilização dos usuários.

Para a representação da informação, é preciso traduzir a linguagem natural do documento em condensações do seu conteúdo temático. Na linguagem natural, há muita ambiguidade, pois uma mesma palavra pode ter mais de um significado (polissemia) e também um mesmo significado pode ter diferentes palavras que a representam (sinonímia). Essa vasta possibilidade de representações que um conceito pode ter em uma LN interfere no momento da representação da informação e, assim, na recuperação dos documentos, tendo em vista que restringe a coincidência dos termos utilizados pelo usuário com os termos utilizados na representação dos conceitos pelo profissional da informação.

A Organização da Informação supõe um processo desdobrado em dois subprocessos solidários: o armazenamento e a recuperação. Nessa medida, compreende-se que a entrada de dados inconsistentes implicará a saída de dados também inconsistentes. Nesse processo, são ativados mecanismos cognitivos que influem tanto na entrada como na saída do sistema de recuperação da informação (SRI), porque são dependentes da maneira como usamos nossa mente para realizar abstrações. A representação da informação dentro do contexto da BCI é realizada por meio dos processos de catalogação, de indexação e de classificação em um Sistema de Recuperação da Informação (SRI).

No processo de catalogação, também conhecido como representação descritiva, descreve-se um item bibliográfico com intuito de torná-lo único

entre os demais em um determinado acervo, permitindo identificá-lo, localizá-lo e representá-lo nos catálogos. E, segundo Mey (1995, p. 5)

Catálogo é o estudo, preparação e organização de mensagens codificadas, com base em itens existentes ou passíveis de inclusão em um ou vários acervos, de forma a permitir interseção entre as mensagens contidas nos itens e as mensagens internas dos usuários.

Na Antiguidade, o catálogo elaborado na biblioteca de Alexandria foi considerado um marco na história do controle bibliográfico. Segundo Mey (1995, p. 9): “[...] catálogo é um canal de comunicação estruturado, que veicula mensagens contidas nos itens, e sobre os itens, de um ou vários acervos, apresentando-as sob forma codificada e organizada, agrupadas por semelhanças, aos usuários desse(s) acervo(s)”. As bases teóricas da catalogação remontam aos princípios de diferentes autores, tais como Anthony Panizzi, (1839, 1841), Charles C. Jewett (1853), Charles Ami Cutter (1876,1904).

Os elementos que compõem a representação descritiva são: (a) a estrutura dos campos; (b) a descrição do item; (c) os pontos de acesso descritivos. Esses elementos constituem a ficha catalográfica com a descrição física do item, indicando as relações bibliográficas com elementos descritivos. Na representação da informação, os metadados atuam como referenciais ao item informacional representado, e como intermediários entre o objeto representado e o usuário. Essa representação descritiva é orientada por códigos de catalogação, normas, formatos e modelos e padrões, entre eles (1) o código Anglo American Cataloging Rules – AACR2, (2) o formato MARC – (Machine Readable Cataloging), (3) o ISBD “*International Standard Bibliographic Description*”, (4) o modelo *Functional Requirements for Bibliographic Records* (FRBR), criado em 1997 pela IFLA com uma nova perspectiva de estrutura e relacionamentos dos registros bibliográficos, visando responder às necessidades dos usuários. Além do modelo FRBR, foram desenvolvidos outros dois modelos conceituais *Functional Requirements for Authority Data* (FRAD) e o *Functional Requirements for Subject Authority Data* (FRSAD). Esses modelos foram desenvolvidos como um modelo de referência e, portanto, operam no nível conceitual, não levando a análise dos requisitos ao nível de um modelo físico de dados. O Recurso Descrição e Acesso (RDA) foi desenvolvido para substituir o AACR2, a partir dos modelos conceituais apresentados anteriormente: o FRBR e o FRAD. Contudo, diferentemente do AACR2, o RDA tem por base uma estrutura teórica, pensando no ambiente digital, e, conseqüentemente, seu escopo é mais abrangente do que o do AACR2. O propósito do RDA é que possa ser utilizado para a descrição de recursos tradicionais ou não (analógicos e

digitais), com foco no usuário que busca esse registro e não no catalogador que cria o registro.

A indexação é outro importante processo de representação que ocorre dentro de um SRI, no qual se espera que o indexador leia o documento e faça a distinção entre a informação relevante e a periférica para melhor representá-lo, para a posterior recuperação. De acordo com a ISO 5963 (1985), a indexação é vista como “[...] a representação do conteúdo dos documentos por meio de símbolos especiais, quer retirados do texto original, quer escolhidos numa linguagem de informação ou de indexação”. Esse processo é realizado em duas etapas: a primeira é a análise do documento para identificação de seu conteúdo informacional; a segunda, a de tradução dos conceitos nos termos de uma linguagem de indexação. As duas etapas envolvem processos interpretativos complexos, o que justifica, de certo modo, a ausência de metodologias que possam dar consistência ao processo. A primeira etapa, análise de assunto, de acordo com os Princípios do UNISIT (1981, p. 8), possui três estágios, que são: (1) compreensão do conteúdo do documento como um todo; (2) identificação dos conceitos que representam esse conteúdo e (3) seleção dos conceitos válidos para recuperação, sendo que se destaca que “na prática, esses três estágios se superpõem”. Na segunda etapa, a tradução, usam-se os sistemas de organização do conhecimento, do tipo tesouros e sistema de classificação bibliográfico. Embora algumas vezes a complexidade desse processo possa ser minimizada, essa é, sem dúvida, umas das atividades mais subjetivas do cotidiano do bibliotecário. A capacidade intelectual de reconhecer o conteúdo de um documento em análise é a questão central do procedimento de indexação.

A Classificação é uma atividade que envolve agrupamentos de coisas semelhantes e a separação daquilo que é diferente, em uma sequência lógica e útil. A palavra classificar vem do latim *classis*, que designava os grupos nos quais se dividia o povo romano. Trata-se de uma atividade fundamental da mente humana que processa as ideias e as distingue a partir das características em comum. Segundo Rodríguez Bravo (2011, p. 156) “classificar se refere ao ato de organizar o universo do conhecimento em alguma ordem determinada”. Para a autora, o termo classificação é polissêmico, pois, além de definir um determinado sistema de classificação, define o processo de classificar, tendo dois objetivos: (a) facilitar a descrição do conteúdo de um documento, colocando-o em uma classe, e (b) ordenar os documentos de acordo com um determinado sistema de classificação, para sua posterior recuperação.

De acordo com Urdiciain (2004), a classificação bibliográfica se inspirou nos princípios da classificação da lógica e dos sistemas filosóficos, sendo uma das linguagens de indexação mais antigas a ser

usadas para a organização da informação em um Sistema de Recuperação da Informação. A base de toda classificação é o princípio da divisão dicotômica (divisão lógica de um conceito em dois outros conceitos, em geral contrários, que lhe esgotam a extensão) conforme exemplificado pela Árvore de Porfírio (Figura 3), baseada na lógica de tipo aristotélica, já mencionada neste trabalho.

As etapas da classificação consistem em examinar o documento para determinar seu assunto principal e, por intermédio dos procedimentos da análise de assunto, posteriormente, consultar em qual classe do conhecimento esse documento poderá ser inserido para sua recuperação. Deve-se levar em consideração os objetivos da instituição e as necessidades dos usuários. Para auxiliar na construção da notação⁷, o classificador conta com as tabelas auxiliares, que acompanham a maioria dos sistemas de classificação, tais como Classificação Decimal de Dewey (CDD) e Classificação Decimal Universal (CDU), que permitem a representação dos aspectos secundários do documento: lugar, tempo, forma e língua, nesse processo de organização da informação.

Conforme apresentado anteriormente, conclui-se que a Organização do Conhecimento e da Informação perpassa a representação do conhecimento e da informação no momento em que o/a classificcionista modela um domínio do conhecimento e o representa em forma de um sistema de organização do conhecimento. Esses sistemas auxiliam os classificadores na representação dos conteúdos dos documentos, nos processos de indexação e classificação, além da representação descritiva, no processo de catalogação.

Na próxima seção, discorre-se sobre as contribuições que essas teorias e técnicas podem propiciar para a representação do conhecimento e da informação no âmbito da Web. As tecnologias da Web Semântica trouxeram, também, uma grande contribuição para os processos de organização, representação e recuperação da informação.

4. Representação do Conhecimento e da Informação na Web: contribuições mútuas

Com as constantes inovações tecnológicas da informação e da comunicação, as informações digitais fazem parte de nosso universo. Nesse contexto, é preciso estruturas de representação e mecanismos de acesso que deem conta do tratamento desse novo ambiente e tipo de recurso informacional.

⁷ Conjunto de símbolos e regras de aplicação que permitem a representação de conceitos, e, eventualmente, de suas relações [...] notação de um esquema de classificação [...] é a representação simbólica de suas classes, divisões e subdivisões (CUNHA; CAVALCANTI, 2008, p. 261).

Anteriormente, o armazenamento da informação era feito somente nos discos rígidos dos computadores com grande capacidade de processamento. Atualmente, pode-se realizar esse armazenamento de maneira diversa, de modo que há a possibilidade de arquivos, dados e aplicativos serem acessados em qualquer lugar e a qualquer hora, utilizando-se tanto um computador como dispositivos móveis, desde que haja conexão com a Internet, por meio do armazenamento nas chamadas “nuvens” (LIMA, 2018).

Em 1989, o físico inglês Sir Timothy John Berners-Lee, do Centro Europeu de Pesquisas Nucleares, apresentou à comunidade a WWW (*World Wide Web*), que tinha como base três tecnologias fundamentais: o HTML (*Hypertext Markup Language*); o servidor HTTP (*Hypertext Transfer Protocol*) e o URI (*Unified Resource Identifier*). Essas tecnologias possibilitaram um avanço na Internet, pois facilitaram o compartilhamento de documentos entre pesquisadores, permitindo que eles fossem acessados e descobertos tanto por seres humanos quanto por máquinas. Sendo uma rede interativa, a Web possibilita, também, a cooperação humana, com acesso remoto simultâneo aos documentos, utilizando tecnologias como o *linked data* e modelos estruturados para melhorar a representação semântica e a recuperação da informação.

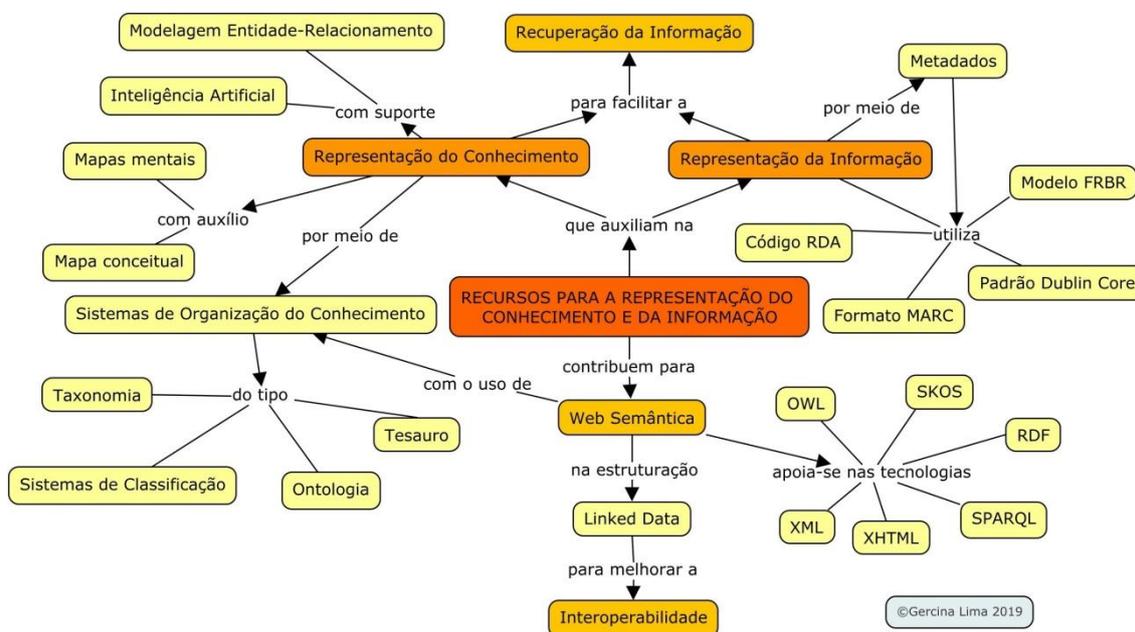
Porém, desde seu surgimento nos anos 1990, a Web continua passando por transformações, tanto no âmbito tecnológico quanto nas suas relações com o usuário. Historicamente, saímos de uma rede estática e marcada por interesses comerciais, a Web 1.0 (por volta do ano 1990), e adentramos na nova era, que foi denominada Web 2.0 (por volta do ano 2000), que prioriza a criação de sites, programas e serviços web centrados na interação dos usuários. As necessidades dos usuários passam a ser o ponto central de estudos na área, já que a recuperação da informação é o principal foco de toda a área da Organização da Informação e do Conhecimento. Muitas dessas pesquisas têm sido relacionadas com a interação entre o sistema de RI e o usuário, com o intuito de melhorar a eficiência dos sistemas de recuperação de informação e dos processos de catalogação, de classificação, e de indexação de assunto de um documento. Com o crescimento e o desenvolvimento da Internet cada vez mais acelerados, surgem novos interesses de pesquisa em tecnologias da Web com uma participação cada vez maior dos usuários, que fornecem metadados e contribuem para a marcação social, mas exigem criações de interfaces mais intuitivas.

Para Soergel (1997), o melhor Sistema de Recuperação da Informação é aquele que suporta qualquer tipo de busca com facilidade dentro do espaço multidimensional da informação. Para isso, precisam resolver as dificuldades que se referem à falta de controle de sinonímia, erros de grafia, flexões de gênero, polissemia, entre outros, que dificultam

a recuperação das informações. Isso geralmente ocorre por causa do uso da linguagem natural, falta de contextualização semântica dos recursos informacionais, porque as máquinas não podem interpretar o significado do conteúdo inserido. Dessa forma, é necessário cada vez mais que novos recursos de informação sejam conectados, tornando-se indispensável criar ferramentas tecnológicas para estruturar e organizar o conhecimento contido nos recursos digitais, de maneira a facilitar a compreensão pela máquina, o gerenciamento e a troca de informações com outros programas. Segundo Zeng (2018), a interoperabilidade semântica tem sido um tópico discutido nas comunidades da área muito antes da chegada da Internet; no entanto, nunca foi tão urgente quanto agora no atual cenário de informações digitais.

É nessa conjuntura que as teorias, técnicas e instrumentos da Organização do Conhecimento e da Informação descritas nas seções anteriores se conectam para construir um espaço integrado de informação, em conjunto com os padrões, modelos e linguagens da Web, nos diferentes níveis da representação da informação e do conhecimento, conforme ilustra a Figura 11, que esquematiza os principais recursos para representação do conhecimento e da informação do ponto de vista dos modelos conceituais e das tecnologias para suas representações.

Figura 11 – Recursos para representação do conhecimento e da informação na Web

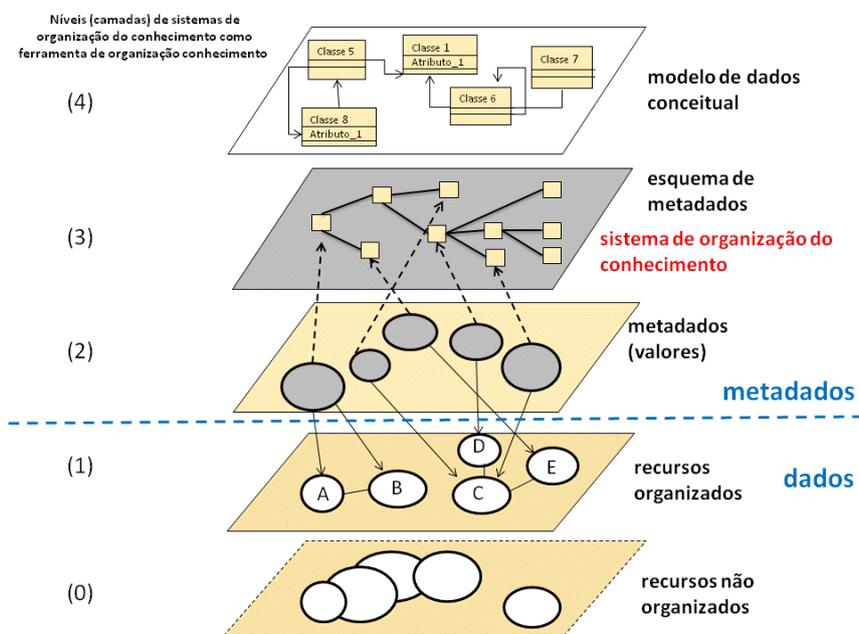


Fonte: Elaborado pela autora.

Nessa perspectiva, os Sistemas de Organização do Conhecimento (SOC) podem contribuir como (1) uma ferramenta de organização do conhecimento e (2) e como um recurso de informação. Bratková,

e Kučerová (2014, p. 5) separam esquematicamente os componentes que são utilizados como recursos para a organização do conhecimento, conforme está ilustrado na Figura 12.

Figura 12 – Componentes utilizados como recursos para a organização do conhecimento



Fonte: Adaptado de BRATKOVÁ, E.; KUČEROVÁ, H., 2014, p. 5.

Para os autores, o nível (0) representa os dados não organizados e não estruturados; o nível (1) representa um agrupamento de recursos de acordo com um critério específico que forma as estruturas organizadas; o nível (2) apresenta a organização do conhecimento por intermédio de metadados que podem descrever tanto as características de conteúdo quanto as características físicas, ou fazer uma referência a uma fonte de informação; agregando valor ao metadado; no nível (3) já existe um esquema de metadado mais estável, que possui não somente os valores, mas permite os relacionamentos de seus elementos e possíveis atributos, como ocorre em uma representação estruturada, sendo uma função dos KOS; o último nível (4) é a representação da “superestrutura” de metadados, a qual permite que esses modelos conceituais possibilitem pesquisas teóricas e sejam implementados em linguagens formais no ambiente da Web. Dessa forma, os SOC auxiliam na representação de um domínio em estruturas que são decodificadas por máquinas, a partir de modelos conceituais.

Os metadados, definidos como “dados sobre os dados” têm um papel fundamental na estruturação da Web, principalmente a semântica. A finalidade principal dos metadados é documentar e organizar, de forma estruturada, os dados das organizações, com o objetivo de minimizar

duplicação de esforços, facilitar a manutenção dos dados e permitir o relacionamento entre os recursos de informação. A Web Semântica baseia-se essencialmente nos metadados e em suas representações, que são estruturados com elementos de descrição de conteúdo dos dados. Existem diversos padrões de metadados desenvolvidos para diferentes finalidades (SOUZA; CATARINO; SANTOS, 1997, p. 95).

O padrão *Dublin Core* (DC) é um dos padrões de metadados utilizados para a descrição dos documentos disponíveis na Web. Inicialmente, os metadados utilizados para organizar a Web somente descreviam os dados, tais como autor, título, descrição do conteúdo, palavras-chaves etc. expressos pelo uso das *tags* <meta> na linguagem HTML para registrá-los na web. Essas tecnologias, também, contribuíram para melhorar os considerados serviços tradicionais dos centros de informação, possibilitando oferecer um serviço unidirecional que auxiliava o usuário a ter acesso à informação disponível na rede, a partir de navegadores em ambiente *on-line*, mas sem interatividade, com sites estáticos, como acesso às bases de dados bibliográficos, podendo ser também *off-line*.

As tecnologias da Web Semântica trouxeram uma grande contribuição para os processos de organização, representação e recuperação da informação. Em relação à representação descritiva, teve-se a possibilidade de mudança do catálogo tradicional para os catálogos automatizados, inserindo as informações contidas em uma ficha catalográfica, que não podem ser simplesmente digitadas no computador, nos campos do Formato MARC – *Machine Readable Cataloging*, para produzir um catálogo automatizado. Possibilitaram a criação dos *Online Public Access Catalog* (Catálogo On-line de Acesso Público – OPAC), que permitiu que várias bibliotecas formassem uma rede, disponível por meio da intranet ou da internet, sendo acessada em uma mesma interface, em forma de um “portal”, permitindo, assim, que tanto o profissional da informação quanto o usuário possam, simultaneamente, organizar e ter acesso à informação. Bhattacharya (2016) aponta que uma das vantagens do OPAC Web é permitir que o usuário possa pesquisar informações, conectando-se ao *Uniform Resource Localizador* (URL) da Web OPAC a qualquer momento e de qualquer lugar do mundo. Além disso, facilita para os usuários a solicitação das informações sobre empréstimos, reservas, relacionadas a seu próprio perfil da biblioteca, bem como reservas automáticas. Além disso, o catálogo OPAC integra, além dos resultados do próprio catálogo, os dados advindos dos repositórios

institucionais, das bases de dados, das editoras, de fontes externas e contribuições dos próprios usuários com o auxílio do protocolo Z39.50⁸.

Hoje o que mais se precisa na Web são recursos para que a padronização de diferentes tipos de dados seja comunicada e entendida, possibilitando a interoperabilidade semântica. Para alcançar tal objetivo, de conectividade e interoperabilidade de dados na web, Berners-Lee (2006) estabeleceu o *Linked Data*, que foi adotado pelo grupo *Semantic Web Education and Outreach* (SWEO), ligado à W3C. Resumidamente, o *Linked Data* é um conjunto de boas práticas a serem seguidas para a implementação e a publicação de dados na web, de modo que permita o compartilhamento e o acesso de conteúdos entre pessoas e máquinas (BERNERS-LEE, 2006). Para tanto, a integração de dados é feita a partir de regras, normas de formatos (extensões) e metadados (para a descrição dos aspectos semânticos dos dados), facilitando o compartilhamento e o reúso de informações advindas de fontes diversas.

Para esse fim, existem tecnologias de representação do conhecimento na Web que permitem essa padronização, algumas já citadas anteriormente, mas entre elas destacam-se: Padrão Dublin Core, *Resource Description Framework* (RDF), *eXtensible Markup Language* (XML), *Protocol and RDF Query Language* [SPARQL], *Simple Knowledge Organization System* (SKOS) e *Web Ontology Language* (OWL). Essas tecnologias possibilitam a estruturação e a organização do conhecimento contido nos recursos digitais, facilitando a sua compreensão pela máquina e a troca de informações. Pode-se dizer que essas tecnologias de representação de conhecimento são o núcleo central da Web Semântica, pois possibilitam maior atribuição de semântica de modo formal que culmina no aumento da possibilidade de interpretação pela máquina (LIMA, 2018). A Figura 13 ilustra as melhores práticas para atingir a interoperabilidade de acordo com normas e recomendações aceitas pelas comunidades globalmente.

De acordo com Zeng (2018), a partir da base, a interoperabilidade na camada do sistema aborda questões sobre incompatibilidades entre *hardware* e sistemas operacionais, para o intercâmbio técnico de dados por intermédio de redes, computadores, aplicativos e serviços da web. Problemas sintáticos que afetam diretamente qualquer esforço de interoperabilidade são as diferenças na codificação, decodificação e representação de dados. Os padrões de linguagem de dados mais importantes que permitem a troca de dados por intermédio de formatos

⁸ Z39. 50 é um protocolo de comunicação entre computadores, desenhado para permitir pesquisa e recuperação de informação – documentos com textos completos, dados bibliográficos, imagens, multimeios – em redes de computadores distribuídos. Baseado em arquitetura cliente/servidor e operando sobre a rede Internet, o protocolo permite um número crescente de aplicações (MOEN, 1995).

de dados comuns são as recomendações oficiais do W3C (*World Wide Web Consortium*) desenvolvidas para a Web Semântica. As variações arquitetônicas de informações em estruturas de dados, modelos de dados e esquemas adicionam outra camada de desafios de interoperabilidade.

Figura 13 – Normas e recomendações para solucionar problemas de interoperabilidade



Fonte: Adaptado de ZENG (2018, p. 6).

Nos esforços para permitir a troca de dados por intermédio de estruturas predefinidas, modelos conceituais foram estabelecidos pelas comunidades da LAM (biblioteca, arquivo e museu) na era digital. Os modelos conceituais são independentes de qualquer sistema de sintaxe e aplicativo de codificação específico. Esses são mais bem visualizados a partir de padrões comunitários desenvolvidos para criar dados estruturados e fornecer acesso a recursos de informação em várias comunidades LAM, entre eles: (1) IFLA LRM (Modelo de Referência de Biblioteca), um modelo formalmente adotado pelo Comitê Profissional da IFLA em agosto de 2017, que consolidou três modelos de família IFLA FRBR e abrange todos os aspectos dos dados bibliográficos; (2) DCMI *Abstract Model*, uma recomendação da *Dublin Core Metadata Initiative* (DCMI) (2007) especificando os componentes e construtos usados nos metadados Dublin Core; (3) BibFrame (Framework Bibliográfico), um novo modelo (versão 2.0, 2016) iniciado pela Biblioteca do Congresso para descrever dados bibliográficos; CIDOC-CRM, (4) o Modelo de Referência Conceitual (CRM) produzido pelo Comitê Internacional para Documentação (CIDOC) do Conselho Internacional de Museus (ICOM) para descrever os

conceitos implícitos e explícitos e as relações usadas na documentação do patrimônio cultural; (5) RiC-CM (modelo conceitual Records in Context), primeiro rascunho lançado em setembro de 2016 pelo Grupo de Especialistas do Conselho Internacional de Arquivos sobre Descrição Arquivística. Muitos modelos e perfis de domínio foram desenvolvidos sob o guarda-chuva de um modelo conceitual para garantir consistência e compreensão, bem como a interoperabilidade entre domínios de LAMs, e para ir além dos ambientes restritos. Um esforço significativo é um perfil simplificado do CIDOC-CRM, chamado *Linked Art Profile* do CIDOC-CRM (ZENG, 2018).

A interoperabilidade é basicamente impulsionada pela comunicação com um propósito coerente. Na prática da integração e da interoperabilidade, vários contextos (incluindo, mas não se limitando a, tempo, estrutura espacial, e terminologia) devem ser abordados. De acordo com Barros (2004):

Por contexto compreende-se o enunciado que exprime uma ideia completa, no qual o termo estudado se encontra atualizado. A identificação das características de um conceito num contexto é possível graças aos descitores (BARROS, 2004, p. 109).

Em geral, um contexto é comumente entendido como sendo as circunstâncias que formam o cenário para um evento, declaração, processo ou ideia, e em termos dos quais o evento, a declaração, o processo ou a ideia podem ser compreendidos e avaliados.

Portanto, considera-se que o cerne da Web Semântica é composto pela integração de dados vinculados, em grande escala e em vários níveis de complexidade, pela acessibilidade e pela possibilidade de raciocínio sobre os dados disponibilizados na Web. Vislumbra-se como um grande desafio para o desenvolvimento da Representação do Conhecimento e da Informação na Web a busca por soluções para a interoperabilidade, visando ao compartilhamento de recursos informacionais, dando ênfase ao desenvolvimento de esquemas e padrões de metadados, para facilitar o acesso a informações nos diversos repositórios digitais publicados na Web. Além disso, é notória a necessidade de realizar pesquisas que possibilitem a compatibilização entre Sistemas de Organização do Conhecimento e a identificação de metodologias para que eles possam ser acessados e manipulados on-line.

Para isso, o modelo de dados *Simple Knowledge Organization System* (SKOS) tem potencialidade para representar a estrutura desses diferentes tipos de (SOCs), permitindo suas publicações com seus dados interligados por intermédio da proposta *Linked Data* (dados ligados entre si). No entanto, o mundo continua a crescer com a tendência digital, e as tecnologias e as linguagens estão mudando rapidamente. Com isso, os Sistemas de Recuperação de Informação, desde os mais simples até os

mais complexos, podem fazer uso da tecnologia digital para aperfeiçoar as suas atividades em vários contextos. Nesse caso, irão sempre depender dos avanços tecnológicos da Web Semântica para dar acesso à informação com qualidade.

5. Considerações finais

A partir das reflexões, conclui-se que muitos estudos têm sido realizados, principalmente no que tange a procurar soluções para aumentar a estruturação dos dados para integrar de forma significativa as várias fontes heterogêneas. Para que a Web Semântica funcione, é necessário construir relacionamentos entre os aplicativos, utilizando uma linguagem comum para que os aplicativos possam trocar informações entre si. A estrutura da Web Semântica permitiu que máquinas conseguissem decodificar o significado do conteúdo de cada recurso, adicionando um maior significado semântico a esses recursos. Assim, se por um lado a tecnologia tem participado ativamente na resolução de problemas da recuperação da informação de forma a minimizar seus impactos, por outro, resta um caminho a ser percorrido em relação à representação do conhecimento e da informação, objetivando sua interoperabilidade.

É sabido que, para progredir da Web Sintática para a Web Semântica, tem-se que passar do nível de informação para o nível de conhecimento. Nesse sentido, nota-se que existem gargalos a serem resolvidos, tais como: (1) as dificuldades relativas ao aumento das publicações na Internet; (2) a grande quantidade de dados não estruturados; (3) a dificuldade cada vez maior de acesso a esse volume de informação e sua recuperação; (4) as informações distribuídas em diversas plataformas e (5) a dificuldade de padronização das tecnologias da web para a interoperabilidade semântica.

As soluções vislumbradas perpassam (1) pela criação de ontologias para auxiliar na disponibilização de conteúdos que necessitam ser mais bem explicitados na web para serem recuperados; (2) pela disponibilização de informações em diversos formatos e acessos; (3) pelo serviço de referência virtual para ajudar os usuários a localizar as melhores fontes de informação; (4) pelo desenvolvimento de repositórios com recursos de aprendizagem disponíveis em diferentes formatos; (5) pelo armazenamento da informação em nuvens (*Cloud Computing*); (5) pela melhora da padronização para tornar o conteúdo acessível para usuários de necessidade especial.

Acredita-se que as teorias e as técnicas da organização e representação do conhecimento e da informação oriundas da Biblioteconomia e Ciência da Informação (BCI) conjuntamente com a adoção de um conjunto de tecnologias inteligentes podem contribuir para

que o usuário ganhe, cada vez mais, autonomia para acessar a informação, utilizando diferentes ferramentas e alcançando um melhor nível de relevância nos resultados da busca na Web. Supõe-se que a inteligência artificial terá um papel fundamental na criação de algoritmos, que auxiliará a máquina a nos dar respostas mais eficientes e eficazes no processamento e na recuperação da informação.

Dentro da perspectiva dos estudos e ensino nessa área, pode-se constatar que os avanços que ocorreram desde a década de 1950 até os dias atuais foram relevantes e marcaram o desenvolvimento da área de Organização do Conhecimento, tanto no armazenamento e na representação quanto na recuperação da informação, porém ainda carece de mais pesquisas. As necessidades dos usuários deveriam se tornar o ponto central de estudos na área, já que a recuperação da informação é o principal foco de toda a área da Organização do Conhecimento e da informação. Defende-se, aqui, uma maior integração entre os profissionais da área de Organização do Conhecimento com outros profissionais das áreas interdisciplinares, para que os avanços nas pesquisas se tornem uma meta comum.

Vale destacar que este trabalho não teve a pretensão de aprofundar na temática, mas discorrer sobre as teorias e técnicas da Organização e Representação do Conhecimento e da Informação e suas possíveis contribuições para Web, bem como os avanços tecnológicos que podem subsidiar essas realizações e, finalmente, fica registrada uma visão geral do conteúdo que subsidia a prática pedagógica no contexto da disciplina Organização do Conhecimento e da Informação, ministrada por mim, no Programa de Pós-graduação Gestão e Organização do Conhecimento (PPGGOC), da Escola de Ciência da Informação da UFMG.

Agradecimento

Agradeço ao apoio do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq/Brasil), pela concessão da bolsa de Produtividade de Pesquisa - PQ.

Referências

ARISTÓTELES. *Metafísica*: livro alfa. São Paulo: Abril Cultural, 1978.

ARISTÓTELES. *Órganon*. São Paulo: Edipro, 2010.

BARROS, L. A. *Curso básico de terminologia*. São Paulo: Editora da Universidade de São Paulo, 2004.

BERNERS-LEE, T. *Linked Data*. Design Issues [site]. 2006. Disponível em: <http://www.w3.org/DesignIssues/LinkedData.html>. Acesso em: 01 jul. 2019.

BHATTACHARYA, A. Library 3.0 and its impact on modern library services. *International Journal of Next Generation Library and Technologies*, v. 1, n. 1, 2016. Disponível em: <http://www.ijnglt.com/files/Vol%202%20Issue%201/Anindiya.pdf> . Acesso em: 14 nov. 2019

BRÄSCHER, M.; CAFÉ, L. Organização da Informação ou Organização do Conhecimento? *In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO*, 9., 2008, São Paulo. *Anais [...]*. São Paulo: ANCIB, 2008. Disponível em: <http://enancib.ibict.br/index.php/enancib/ixenancib/schedConf/presentations>. Acesso em: 17 jun. 2019.

BRATKOVÁ, E.; KUČEROVÁ, H. Knowledge Organization Systems and Their Typology. *Knihovna*, República Tcheca, v. 25, supplementum 2, p. 1-25. 2014. Disponível em: http://knihovna.nkp.cz/knihovna142_suppl/1402suppl01.htm. Acesso em: 19 jun. 2019.

CABRÉ, M. T. *La terminología: representación y comunicación: elementos para una teoría de base comunicativa y outros artículos*. Barcelona: Instituto Universitario de Lingüística Aplicada, 1999.

CAMPOS, M. L. A. *A organização de unidades de conhecimento: o modelo conceitual como espaço comunicacional para a realização da autoria*. 2001. Tese (Doutorado em Ciência da Informação) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2001.

CAMPOS, M. L. A. O papel das definições na pesquisa em ontologia. *Perspectivas em Ciência da Informação*, Belo Horizonte, v. 15, n. 1, p. 220-38, 2010. Disponível em: <http://www.scielo.br/pdf/pci/v15n1/13.pdf>. Acesso em: 19 set. 2019.

CAMPOS, M. L. A; SOUZA, R. F. de; CAMPOS, M. L. M. de. Organização de unidades de conhecimento em hiperdocumentos: o modelo conceitual como espaço comunicacional para a realização da autoria. *Ciência da Informação*, Brasília, v. 32, n. 2, p. 7-16, 2003.

CLEVELAND, D.; CLEVELAND, A. *Introduction to indexing and abstracting*. 3. ed. Colorado: A division of Greenwood Publishing Group, 2001.

CUNHA, M. B. da; CAVALCANTI, C. R. de O. *Dicionário de Biblioteconomia e Arquivologia*. Brasília: Briquet de Lemos, 2008. XVI, 451 p.

CUTTER, C. A. *Rules for a dictionary catalog*. Washington: UNT Libraries Digital Collections, 1876. 274 p.

DACONTA, M. C.; OBRST, L. J.; SMITH, K. T. *The Semantic Web: A Guide to the Future of XML, Web Services, and Knowledge Management*. Indiana: Wiley, 2003. 281 p.

DAHLBERG, I. Teoria do conceito. Tradução Astério Tavares Campos. *Ciência da Informação*, Rio de Janeiro, v. 7, n. 2, p. 101-107, 1978a.

DAHLBERG, I. A referent-oriented, analytical concept theory of Interconcept. *International Classification*, Frankfurt, Alemanha, v. 5, n. 3, p. 122-151, 1978b. Disponível em: <http://revista.ibict.br/ciinf/article/view/115/115>. Acesso em: 18 ago. 2019.

DAHLBERG, I. Knowledge organization: its scopes and possibilities. *Knowledge Organization*, Frankfurt, v. 20, n. 4, p. 211-222, 1993.

DAHLBERG, I. Knowledge organization: a new science? *Knowledge Organization*, Frankfurt, v. 33, n. 1, p. 11-19, 2006.

GARDNER, H. *A nova ciência da mente: uma história da revolução cognitiva*. São Paulo: EDUSP, 1996.

GOMES, H. E. Manual de elaboração de tesouros monolíngues. Brasília: o Programa, 1990. 78 p.

GRUBER, T. R. *A Translation Approach to Portable Ontology Specifications*. Knowledge Systems Laboratory, Computer Science Department Stanford University Stanford, California September 1992 Technical Report KSL 92-71, 1992.

GRUBER, T. R. A translation approach to portable ontologies. *Knowledge Acquisition*, Reino Unido, v. 5, p. 199-220, 1993. Disponível em: <http://www.sciencedirect.com/science/article/pii/S1042814383710083?np=y>. Acesso: 10 ago. 2019.

GUARINO, N. Formal Ontology and Information System. In: Guarino, N. (Ed.) *Formal Ontology in Information System*. Amsterdam: IOS Press, 1998. p 3-15.

GUARINO, N.; OBERLE, D.; STAAB, S. What Is an ontology? In: *Handbook on ontologies*. 2nd ed. STAAB, S., STUDER, R. BerlinSpringer, 2009. p. 1-17.

HILL, L., et al. Integration of Knowledge Organization Systems into Digital Library Architectures. In: ASIS&T, 13., 2002; SIG/CR WORKSHOP, RECONCEPTUALIZING, CLASSIFICATION RESEARCH, 2., 2002. *Anais [...]*. Silver Sping: ASIS&T, 2002.

HJORLAND, B. What is Knowledge Organization (KO)? *Knowledge Organization*, Frankfurt, Alemanha, v. 35, n. 2-3, p. 86-101, 2008.

HODGE, G. *Systems of Knowledge Organization for Digital Libraries: beyond traditional authority files*. Washington: The Council on Library and Information Resources, 2000. Disponível em: <https://www.clir.org/wp-content/uploads/sites/6/pub91.pdf>. Acesso em: 21 maio 2019.

INTERNATIONAL STANDARD ORGANIZATION. *ISO 5963-1985 – Documentation: methods for examining documents, determining their subjects, and selecting indexing terms*. Suíça: ISSO, 1985.

INTERNATIONAL STANDARD ORGANIZATION. *ISO 25964: thesauri and interoperability with other vocabularies. Part 1: Thesauri for information retrieval*. Geneve: International Standard Organization, 2011.

JEWETT, C. C. *Smithsonian report on the construction of catalogs of libraries, and their publication by means of separate, stereotyped titles, with rules and examples*. 2nd ed. Washington: Smithsonian Institution, 1853. Disponível em: <https://archive.org/details/constcatlib00jewerich/page/n4>. Acesso em: 21 jun. 2019.

LANCASTER, F. W. *Indexação e resumos: teoria e prática*. 2. ed. Brasília: Briquet de Lemos, 2003.

LANGRIDGE, D. W. *Classificação: abordagem para estudantes de biblioteconomia*. Tradução de Rosali P. Fernandez. Rio de Janeiro: Interciência, 1977.

LARA, M. L. G. de. Conceitos de organização e representação do conhecimento na ótica das reflexões do grupo tema. *Informação & Informação*, Londrina, v. 16. n. 3. p. 92-121, 2011.

LE COADIC, Yves-François. *A Ciência da Informação*. 2. ed. Brasília: Briquet Lemos Livros, 2004.

LIMA, G. A.; MACULAN, B. C. M. S. Estudo comparativo das estruturas semânticas em diferentes sistemas de organização do conhecimento. *Ciência da Informação*, Brasília, DF, v. 46, n.1, p. 60-72, 2017.

LIMA, G. A. Da organização do conhecimento à recuperação da informação: teorias e técnicas. In: LIMA, G.A. et. al. (Org.) *Pesquisa Integrada em Organização do Conhecimento e Recuperação da Informação*. Belo Horizonte: ECI/UFMG, 2016.

LIMA, G. A. *Representação, recuperação e acesso da informação: a evolução da Biblioteca 1.0 à Biblioteca 3.0*. In: CAMPOS, M. L. A. et al. (Org.). *Produção, tratamento, disseminação e uso de recursos informacionais heterogêneos: diálogos interdisciplinares*. Niteroi: IACS/UFF, 2018. p. 80-88.

LIMA, R. L. L.de. *Garantia literária: revisão de literatura – organização e representação da informação*. 2014. 34 f. Monografia (Graduação em Biblioteconomia) – Departamento de Biblioteconomia, Universidade Federal do Rio Grande do Norte, Natal, 2014. Disponível em: https://monografias.ufrn.br/jspui/bitstream/123456789/922/1/RayssaLLL_Monografia.pdf. Acesso em: 10 ago. 2019.

LÖBNER, S. *Understanding semantics*. London: Arnold, 2002.

McGARRY, K. *O contexto dinâmico da informação: uma análise introdutória*. Trad. Helena Vilar de Lemos. Brasília: Briquet de Lemos/Livros, 1999. 206 p.

MCGUINNESS, D. L. Ontologies come of age. In: FENSEL, Dieter. et. al (Ed.). *Spinning the Semantic Web: bringing the World Wide Web to its full potential*. Cambridge, Mass: The MIT Press, 2003. p. 171–196. Disponível em: <http://www-ksl.stanford.edu/people/dlm/papers/ontologies-come-of-age-mit-press-%28with-citation%29.htm>. Acesso em: 01 out. 2019.

MEY, E.S.A. *Introdução à catalogação*. Brasília: Briquet de Lemos, 1995.

MEY, M. de. *The cognitive paradigm: an integrated understanding of scientific development* Chicago: University of Chicago, 1982.

MOEN, W. *A guide to the ANSI/NISO Z39.50 protocol: Information retrieval in the information infrastructure*. Maryland: National Information Standards Organization, 1995.

MORA, J. F. *Dicionário de filosofia*. 2. ed. Trad. Roberto Leal Ferreira. São Paulo: Martins Fontes, 1996.

MORUJÃO, C. A lógica modernorum: lógica e filosofia da linguagem na escolástica dos séculos XIII E XIV. *Revista Filosofia de Coimbra*, Coimbra, n. 28, p. 301-322, 2006. Disponível em: http://www.uc.pt/fluc/dfci/publicacoes/a_logica_modernorum . Acesso em: 18 jun. 2019.

OBRST, L. Ontologies for Semantically Interoperable Systems. In: ACM CIKM, 3., 2003. *Anais [..]*. Louisiana: ACM, 2003. Disponível em: https://www.researchgate.net/publication/221613357_Ontologies_for_semantically_interoperable_systems. Acesso em: 25 ago. 2019.

PANIZZZI, A. *Rules for the Compilation of the Catalogue*. London: Printed by order of the Trustees, 1841.

PIEDADE, M. A. *Introdução à teoria da classificação*. 2. ed. Rio de Janeiro: Interciência, 1983.

PONTES JUNIOR, J. de; CARVALHO, R. A. de; AZEVEDO A. W. Da recuperação da informação à recuperação do conhecimento: reflexões e propostas. *Perspectivas em Ciência da Informação*, Belo Horizonte, v. 18, n. 4, p. 2-17, 2013.

RANGANATHAN, S. R. *Prolegomena to library classification*. Bombay: Asia Publishing House, 1967.

RODRÍGUEZ BRAVO, B. Apuntes sobre representación y organización de la información. Gijón: Trea, 2011.

SAN SEGUNDO, R. A New Conception of Representation of Knowledge. *Knowledge Organization*, Frankfurt, Alemanha, v. 31, n. 2, 2004.

SMITH, B.; WELTY, C. Ontology: towards a new synthesis. In: INTERNATIONAL CONFERENCE ON FORMAL ONTOLOGY IN INFORMATION SYSTEMS, 2., 2001. *Anais [..]*. Maine: ACM, 2001.

SOERGEL, D. Digital Libraries and Knowledge Organization Systems. In: KRUK, Sebastian (Ed.) *Semantic Digital Libraries*. Berlin, Heidelberg: Springer, 2008.

SOERGEL, D. The rise of ontologies or the reinvention of classification. *Journal of the American Society for Information Science*, v. 50, n. 12, p. 1119-1120, 1999.

SOERGEL, D. Functions of a Thesaurus / Classification / Ontological Knowledge Base. Maryland: University of Maryland, 1997.

SOUZA, T. B. de; CATARINO, M. E.; SANTOS, P. C. dos. Metadados: catalogando dados na Internet. *Transinformação*, Campinas, v. 9, n. 2, 1997. Disponível em: <http://periodicos.puc-campinas.edu.br/seer/index.php/transinfo/article/view/1586>. Acesso em: 05 maio 2019.

SOWA, J. F. *Knowledge representation: logical, philosophical and computational foundations*. Pacific Grove: Brooks/Cole, 2000.

STOCK, W. G., STOCK, Mechtild. *Wissensrepräsentation: Informationen auswerten und bereitstellen*. Oldenburg: GmbH & Co KG, 2008.

STUDER, R.; BENJAMINS, R.; FENSEL, D. Knowledge Engineering: Principles and Methods. *Data & Knowledge Engineering*, Países Baixos, v. 25, p. 161-197, 1998.

UNISIST. Princípios de indexação. *Revista da Escola de Biblioteconomia da UFMG*, Belo Horizonte, v. 10, p. 83-94, 1981.

URDICIAIN, Blanca GIL. *Manual de lenguajes documentales*. Espanha: Trea, 2004.

VICKERY, B. C. *Classification and Indexing in Science*. 3. ed. London: Butterworth, 1975.

VICKERY, B. C. Concepts of Documentation. *Journal of Documentation*, Reino Unido, v. 34, n. 4, p. 279-87, 1978.

VICKERY, B. C. Ontologies. *Journal of Information Science*, Estados Unidos, v. 23, n. 4, p. 277-286, 1997.

VICKERY, B.C., Knowledge representation: a brief review, *Journal of Documentation*, Reino Unido, v. 42, n. 3, p. 145-159, 1986.

WELTY, C.; LEHMANN, F.; GRUNINGER, G.; USCHOLD, M. Ontology: Expert Systems All Over Again? *In: THE NATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE*, 13. 1999. *Anais [...]*. Austin: AAAI, 1999.

WÜSTER, E. Introducción a la teoría general de la terminología y a la lexicografía terminológica. Tradução de Anne-Cécile Nokerman. Barcelona: IULA, 1998.

ZENG, M. L. *Interoperability*. Encyclopedia of Knowledge Organization. ISKO. 2018. Disponível em: <https://www.isko.org/cyclo/interoperability>. Acesso em: 1.º out. 2019.

ZENG, M.L. Knowledge Organization Systems (KOS). *Knowledge organization*. Frankfurt, Alemanha, v. 35, n. 2-3, p. 160-182, 2008. Zhonghong, Chaudhry, Khoo (2006)

ZHONGHONG, W.; CHAUDHRY, A. S.; KHOO, C. Potential and prospects of taxonomies for content organization. *Knowledge Organization*, Alemanha, v. 33, n. 3, p.