

**CAPÍTULO 4**  
**PERSPECTIVAS DE MODELAGEM CONCEITUAL**  
**NA ELABORAÇÃO DE SISTEMAS**  
**DE ORGANIZAÇÃO DO CONHECIMENTO:**  
**SUBSÍDIOS PARA O ENSINO**  
**PERSPECTIVES FOR CONCEPTUAL MODELING**  
**IN THE DEVELOPMENT OF KNOWLEDGE**  
**ORGANIZATION SYSTEMS:**  
**SUBSIDIES FOR TEACHING**

Gercina Ângela de Lima

Universidade Federal de Minas Gerais

limagercina@gmail.com

ORCID: 0000-0003-0735-3856

Maria Luiza de Almeida Campos

Universidade Federal Fluminense

Programa de Pós-Graduação em Ciência da Informação

maria.almeida@pq.cnpq.br

ORCID: 0000-0002-9253-3706

Nair Kobashi

Universidade de São Paulo

nairkobashi@gmail.com

ORCID: 0000-0002-5282-7363

Resumo: Toda representação de conhecimento está sujeita às tecnologias disponíveis em um dado momento/espço temporal. Essa afirmativa se insere perfeitamente no âmbito dos estudos de organização do conhecimento e da recuperação da informação. A afirmativa é também verdadeira no que tange à elaboração de “Modelos de domínios”. Nesse sentido, este artigo apresenta uma sistematização dos conhecimentos que advêm da Biblioteconomia e da Ciência da Informação para a elaboração de modelos de domínios. Estes últimos podem subsidiar o ensino de teorias e métodos de construção de ferramentas de tratamento e recuperação de informação, tais como os sistemas de classificação, os tesouros, as taxonomias e as ontologias.

Palavras-Chave: Organização do Conhecimento. Modelagem conceitual; Instrumentos de representação de informação: Instrumentos de recuperação de informação.

Abstract: All knowledge representation is sensitive to the technologies available at a given time and space. This statement is perfectly within the scope of knowledge organization and information retrieval studies. It is also true in regard to the development of Domain Models. In this sense, we present a systematization of knowledge that comes from Library and Information Science for teaching of theories and methods of constructing instruments for information processing and retrieval, such as classification systems, thesauri, taxonomies, and ontologies.

Keywords: Knowledge organization; Conceptual modeling; Information representation instruments. Information retrieval instruments.

## **1. Considerações iniciais sobre o espaço informacional e a representação de domínios de conhecimento**

Modelos conceituais de domínios são produtos obtidos a partir do reconhecimento de que os campos do saber evidenciam sistemas de conceitos que expressam de forma rigorosa as suas teorias e

métodos. Os modelos conceituais, *lato sensu*, têm por objetivo construir representações de alta qualidade de fenômenos selecionados de algum domínio do conhecimento. No âmbito da Organização do Conhecimento, os modelos resultantes pretendem facilitar a representação e a recuperação de informações em contextos diversos.

Essa questão vem preocupando o campo dos estudos informacionais desde a sua gênese, fato que se torna mais agudo com o desenvolvimento contínuo das tecnologias da informação e comunicação. Esses desafios vêm exigindo pesquisas renovadas com a consequente elaboração de novas teorias e métodos. Nesse sentido, propomos discutir, neste trabalho, diferentes perspectivas teóricas e metodológicas sobre a modelagem de domínios aplicados à elaboração de ferramentas de representação e recuperação de informação, no contexto das redes eletrônicas de informação e comunicação.

Trata-se, portanto, de revisitar teorias e métodos consagrados e também as propostas contemporâneas sobre a construção de modelos conceituais de domínios. São bases fundantes das reflexões os aspectos ontológicos, as questões terminológicas e as questões lógicas e semânticas subjacentes à classificação e à representação de conhecimentos. Esses conhecimentos, como já foi sinalizado por diferentes pesquisadores da área, são indispensáveis para construir ferramentas robustas de representação do conhecimento para viabilizar a difusão, a circulação e a apropriação social da informação. Não por acaso, dedicam-se a essas temáticas os pesquisadores da Ciência da Informação, da Inteligência Artificial, da Psicologia e das Ciências da Linguagem.

No domínio da Biblioteconomia e da Ciência da Informação, teóricos como Ranganathan (1967), Bliss (1929), Feibleman (1954), Vickery (1997) e o próprio Classification Research Group – CRG (Wilson, 1972) construíram teorias com a finalidade de representação de domínios, visando à organização de documentos, mas que, de alguma forma, apresentavam princípios que possibilitavam a representação independentemente dos próprios domínios de saber.

Nesse contexto, dos estudos relacionados à representação e à recuperação de informação, especificamente daqueles ligados às ferramentas atualmente existentes no cenário informacional para a realização de tais processos, existem dúvidas sobre a aplicabilidade dessas ferramentas em ambientes onde questões semânticas se impõem. Até os últimos anos, um dos papéis exercidos pelo profissional de informação se restringia ao desenvolvimento de ferramentas semânticas capazes de representar um dado domínio de conhecimento (por meio de linguagens documentárias como tesouros, taxonomias e esquemas de classificação), possibilitando certa precisão e compatibilidade entre a linguagem utilizada para o tratamento informacional e a linguagem utilizada pelo usuário para a recuperação da informação. Nesse cenário, impera a Web Sintática.

Como apresenta Karin Breitman, a “Internet atual pode ser definida como a Web Sintática” (Breitman, 2005). Na Web Sintática, os computadores fazem apenas a apresentação da informação, porém, o processo de interpretação fica a cargo dos seres humanos. Dessa forma, a Web Sintática é uma rede desenvolvida para a produção de sentido realizada por humanos. A questão que começou a ser colocada há alguns anos no âmbito dos estudos computacionais e que nos levou a uma nova geração da Web, ou seja, à Web Semântica, foi: por que os computadores não podem realizar o trabalho de interpretação de dados?

No cenário da Web Semântica, em que as ontologias formais se inserem, já é possível, de algum modo, não só a representação do conhecimento de um dado domínio, mas também a derivação do conhecimento sobre esse domínio por meio de uma semântica expressa em linguagens formais, possível de ser explorada computacionalmente para fins de inferência (Campos, 2012).

Dessa forma, nossa intenção neste artigo é apresentar uma sistematização inicial sobre os conhecimentos para a elaboração de

modelos de domínios que possam subsidiar o ensino de teorias e métodos de construção de diferentes ferramentas de tratamento e recuperação de informação, tais como os sistemas de classificação, os tesouros, as taxonomias e as ontologias.

## **2. Modelos de representação de domínio no âmbito da Web Sintática**

O conhecimento é algo dinâmico, que está presente na vida do ser humano desde o nascimento e o acompanha até o último dia de sua vida. No percorrer desse caminho, vão sendo adquiridos novos conhecimentos, que sofrem alterações, sendo necessário serem lapidados na sua concepção para que sejam compreendidos, armazenados e, posteriormente, registrados e disseminados dentro de uma comunidade.

Esse processo é realizado pela mente humana, a partir do ponto de vista de como o indivíduo vê seu mundo e o representa por meio de informações sobre objetos pertencentes à sua realidade. Porém, a realidade não pode ser vista sob um único aspecto, restrito às percepções e sensações de um indivíduo. Surge aí a necessidade de se criar mecanismos para a padronização desse conhecimento, de modo a tornar possível o seu compartilhamento social

Cada estrutura de conhecimento pode ser vista como um objeto, ideia ou evento e, também, como um grupo de atributos, ligado a outra estrutura de conhecimento. De acordo com Lima (2018), à medida que aprendemos, apreendemos novas estruturas e ligações, adicionando informações às estruturas existentes, ou alterando essas estruturas por meio de processos de reestruturação. Portanto, a influência do contexto é fundamentalmente importante na categorização. Assim, qualquer classificação individual que desconsidera o contexto pode ser muito limitada e, portanto, de difícil compartilhamento.

Para organizar o conhecimento de qualquer área, desde sua representação à sua recuperação, estudam-se primeiramente os conceitos que compõem esse campo do conhecimento e as relações entre eles. Categorias e estruturas hierárquicas são a melhor maneira de se organizar o conhecimento para recuperação, pelo óbvio motivo de que a informação estruturada é mais fácil de ser recuperada do que a informação desorganizada. Portanto, pode-se dizer que o ser humano, desde os primeiros momentos de vida, percebe o seu mundo, registra, classifica e dá forma categorial às estruturas que espelham o ambiente externo. Nota-se que essa classificação vem de nossa interação com nosso o ambiente. Se não interagimos com o ambiente, não temos o que classificar; o ambiente, por sua vez, influencia o modo como categorizamos o conhecimento.

A área de Organização do Conhecimento (OC) tem suas raízes históricas nas teorias da Classificação. A Classificação pode ser vista como uma área de conhecimento que procura classificar o conhecimento filosófico; no caso da Organização do conhecimento registrado, procura-se criar sistemas de classificação bibliográfica para a representação do conteúdo temático dos documentos, de modo a aumentar a eficácia da recuperação de informações. As duas abordagens da classificação procuram ligar teoria e prática nos estudos e pesquisas realizadas dentro da área da Biblioteconomia e Ciência da Informação (BCI).

As bases da classificação remontam a Aristóteles, em que um todo pode ser dividido em partes, sucessivamente, formando classes e subclasses. Essa lógica de divisão para representar o conhecimento parte do princípio classificatório de gênero e espécie para formar uma estrutura conceitual hierárquica.

Aristóteles foi um dos primeiros estudiosos a pensar o mundo e categorizá-lo (e mesmo antes, com Platão), há mais de dois milênios, com sua inquietude em 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. Destarte, veio da Filosofia, por meio dos estudos de Aristóteles, a primeira preocupação em sistematizar o conhecimento com o intuito de torná-lo padronizado, com a utilização de símbolos que representem a realidade, tornando possível a manifestação dessas ideias com um significado convencional. Portanto, ideias e coisas estariam integradas de forma a materializar o pensamento humano:

Destarte, Aristóteles, em seu *Órganon* (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 Aristóteles 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 conhecimento se divide em ciências teóricas, ciências práticas e ciências poéticas, representando o pensar, o agir e o produzir (Piedade, 1983). Esses estudos de Aristóteles influenciaram o desenvolvimento das teorias de classificação do conhecimento, sendo considerados o embasamento teórico para as classificações bibliográficas que começaram a surgir posteriormente, no século IV. Deve-se mencionar aqui os estudos de Porfírio, que, em sua obra *Isagoge* (Introdução às Categorias), incorporou a lógica aristotélica na categoria substância, considerada como gênero supremo, resultando em cinco predicáveis<sup>1</sup> – gênero, espécie, diferença, propriedade e acidente, que, quando interligados, formam uma lógica sucessiva

---

<sup>1</sup> Porfírio faz distinção entre predicados e predicáveis. Por exemplo: na expressão «todos os homens são mortais», «mortais» é um predicado, mas é atribuído a «todos os homens» de modo diferente do que o seria, por exemplo, o predicado «bípedes sem penas». O predicável será, então, o modo como certo predicado é atribuído a um sujeito (Morujão, 2006, p. 304).

de subordinação, resultando em uma classificação dicotômica, que tornou-se conhecida como “Árvore de Porfírio”.

Para organizar o conhecimento de qualquer área, da sua representação à sua recuperação, estudam-se primeiramente os conceitos que compõem esse campo do conhecimento e as relações entre eles. O núcleo principal desse processo é a escolha dos conceitos e seus relacionamentos e sob quais aspectos se quer representar determinado domínio. Pode-se dizer, então, que os conceitos são a base para a construção da estrutura conceitual, ou modelagem conceitual, e são as relações entre eles que os mantêm reunidos dentro de um determinado domínio e princípio de representação. Para Green (2008), os relacionamentos são o coração da organização do conhecimento:

Tentamos localizar informações relacionadas para a necessidade de um usuário; tentamos localizar documentos contendo palavras que possam ser usadas para transportar a informação procurada; e/ou recuperamos documentos indexados por um descritor de assunto relevante ou classificado em uma categoria de assunto relevante. Usamos relacionamentos para navegar entre assuntos. Ajudamos a localizar documentos que juntos dizem mais do que a soma de suas partes. (p. 158, tradução nossa)<sup>2</sup>

A modelagem conceitual de uma área do conhecimento pode ser influenciada pelo contexto social, cultural, estético, político. A organização de conceitos está diretamente relacionada à capacidade de aprender, supõe a assimilação de novas informações, sua estocagem e sua acomodação dentro de um determinado contexto. Quando

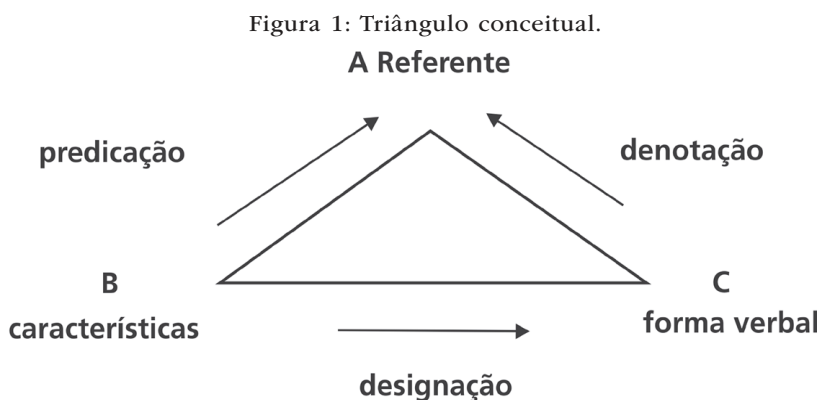
---

<sup>2</sup> We attempt to locate information that relates to a user's need; we attempt to locate documents containing words that might be used in conveying the sought-for information; and/or we retrieve documents indexed by a relevant subject descriptor or classed in a relevant subject category. We use relationships to navigate among subjects. We help to locate documents that together say more than the sum of their parts.



se estabelecem as categorias de um domínio do conhecimento, de certa forma, elas são determinadas pela perspectiva do contexto que foi considerado. Segundo Barros (2004, p. 109): “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 descritores”.

A definição de um conceito, por meio de suas características e predicados, representado pelo seu símbolo verbal, ou seja, o termo, faz com que ele seja contextualizado dentro de uma área específica do conhecimento. O símbolo isolado de seu contexto, mesmo possuindo propriedades semânticas relevantes – se não forem evidenciadas através de sua definição –, não poderá ser interpretado e distinguido de outros conceitos que formam o sistema conceitual daquele domínio. É necessário, portanto, conforme Dahlberg (1978a), que cada objeto seja definido para fixar os seus limites enquanto conceito pertencente àquele domínio específico. Isso pode ser melhor visualizado e simbolizado, por meio do triângulo conceitual proposto por Dahlberg (1978a), conforme Figura 1:

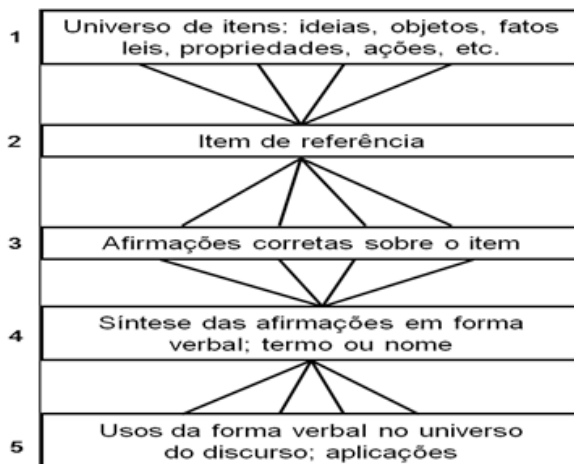


Fonte: Dahlberg, 1978a, p. 144.

O Triângulo Conceitual de Dahlberg tem no ápice o referente, ou seja, o objeto que se quer conceituar; na parte inferior do canto esquerdo, um conjunto de predicados ou as características que formam aquele conceito, e, na parte inferior do canto direito, a forma verbal, que simboliza aquele conceito, representado por um termo.

Dahlberg (1978a) define conceito como “[...] uma unidade de conhecimento, compreendendo afirmativas verificáveis sobre um item selecionado de referência, representado por uma forma verbal”. A autora apresenta um modelo para construção de conceitos (Figura 2), onde estão representadas as relações entre o objeto, o conceito e o termo. O processo mental da formação do conceito se dá por meio de uma linha de pensamento que leva à elaboração do conhecimento, passando por um processo de assimilação da informação pelo cérebro, transformando-a. Após essa elaboração mental, baseada no conhecimento prévio do indivíduo, a unidade de informação se transforma em uma unidade conceitual, que é representada por um termo, o qual possui um único significado, geralmente expresso por símbolos e palavras, com o objetivo de comunicação.

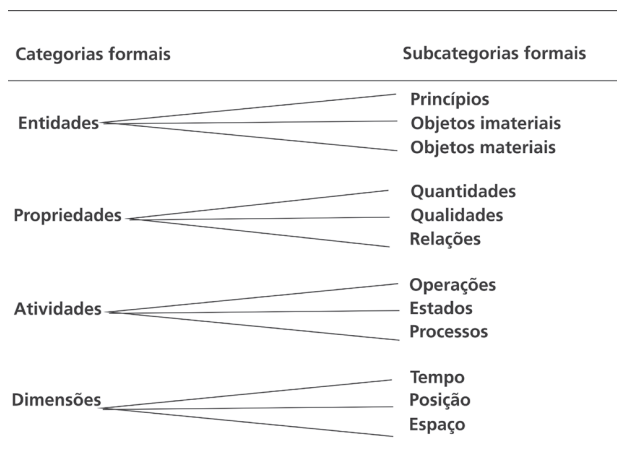
Figura 2: Modelo de construção do conceito.



Fonte: Dahlberg, 1978a, p. 143.

Além disso, Dahlberg propõe quatro categorias formais que se desdobram em 12 subcategorias (Figura 3)<sup>3</sup>, tendo por base as categorias aristotélicas, que permitem combinação entre elas.

Figura 3: Referentes de conceitos.



Fonte: Dahlberg, 1978a, p. 145.

Outra proposta de categorização foi criada por Shiyali Ramamrita Ranganathan (1892-1972), matemático indiano que se tornou bibliotecário e classificacionista. É um sistema quase matemático, que estrutura o conhecimento de maneira que os assuntos compostos sinteticamente derivam de conceitos “elementares”.

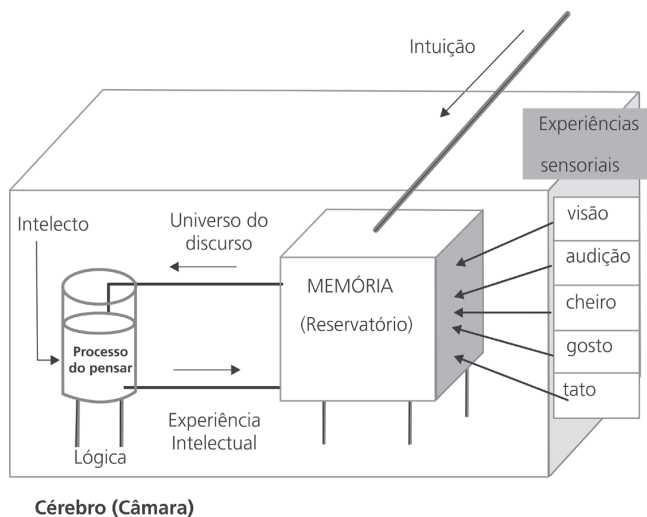
Para Ranganathan (1967), o conceito é produto do pensamento; o conhecimento é representado pela totalidade das ideias armazenadas pela memória humana, e a informação existe quando os conceitos são comunicados. Todos esses processos são sensoriais, permeados pelas experiências cognitivas de cada indivíduo.

<sup>3</sup> Não se pretende neste capítulo trazer definições das categorias. Para aprofundar nos estudos, deve-se consultar a literatura citada.

Ranganathan, no *Prolegomena to Library Classification* (1967), define Conceito como “unidade de pensamento de qualquer nível de complexidade”. Considerou o conhecimento como 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.

Nessa abordagem, Ranganathan apresenta uma alternativa à Árvore de Porfírio, que é um método dicotômico composto de dez divisões, organizadas em um eixo único. Propõe, então, um padrão para mapear o conhecimento com a policotomia ilimitada (número ilimitado de divisões das áreas do conhecimento), representado pela Árvore Baniana (*Banyan Tree*) que, em adição a um tronco original, vários troncos secundários podem ser formados de tempos em tempos, ideia que associou a uma árvore de assuntos (Ranganathan, 1967, p. 368). Para ilustrar a formação das ideias, Ranganathan nos apresenta um diagrama esquemático (Figura 4).

Figura 4: Compartimento do cérebro: formação das ideias.



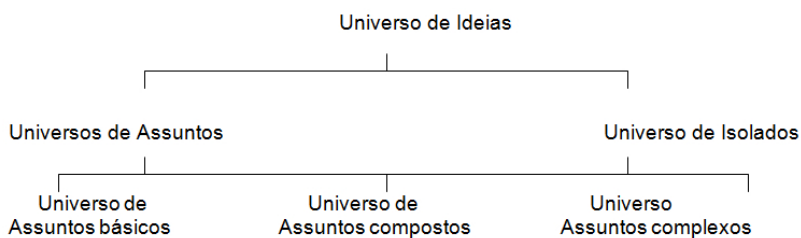
Fonte: Ranganathan, 1967, p. 8.

Assim, Ranganathan (1967, p. 82) descreve o assunto como: "Um 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 uma pessoa qualquer."

Todo assunto é oriundo de um assunto básico e seu componente vem de um ou mais conceitos isolados, 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 para formar um componente de assunto. Mas, sem seu contexto, ela não é considerada um assunto. Ao estabelecer um Conceito, o modelizador/classificacionista descreve o objeto a partir de um entre inúmeros cenários contextuais possíveis.

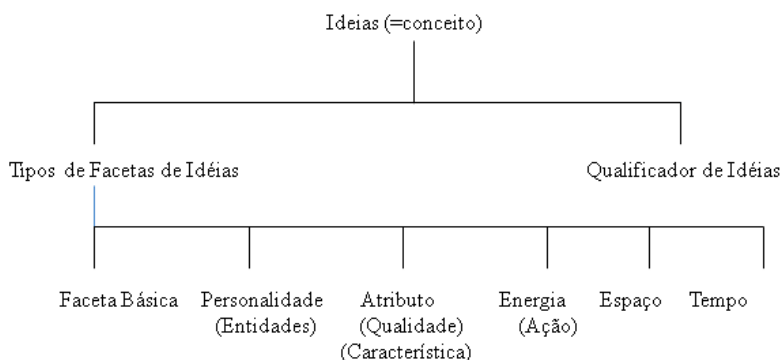
Os esquemas representativos a seguir (Figura 5 e Figura 6) mostram (1) o universo original das ideias em suas relações, seus subuniversos de assuntos e isolados propostos por Ranganathan e (2) os tipos de facetas de ideias propostos por Raghavan.

Figura 5: Universo original das ideias e suas relações segundo Ranganathan.



Fonte: Ranganathan, 1967, p. 85.

Figura 6: Exemplo da categorização de isolados dentro das facetas.



Fonte: Raghavan, 1985, p. 27.

Uma das maiores contribuições de Ranganathan foi a criação das categorias fundamentais, expressas por meio do acrônimo PMEST<sup>4</sup> (*personality, matter, energy, space, time* – em português: personalidade, matéria, energia, espaço e tempo), que é o princípio normativo adotado para organizar um Universo de Assuntos, ou seja, um “corpo” de conhecimento organizado e sistematizado.

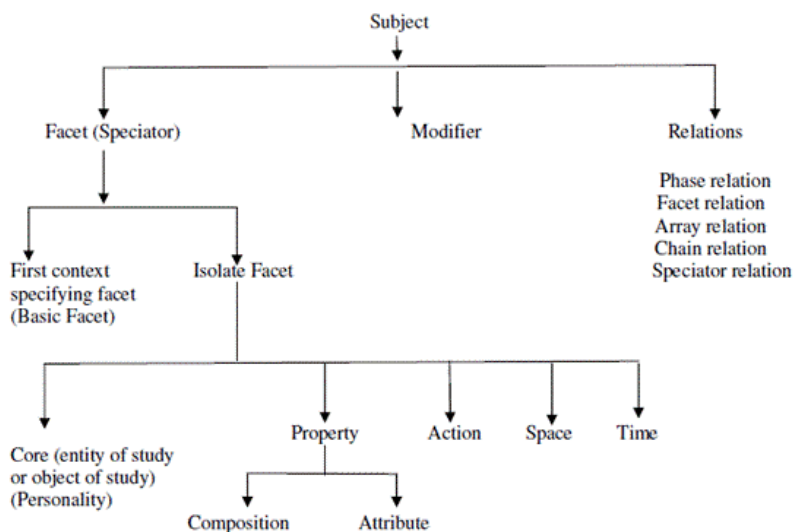
Os fundamentos da Linguística e da Terminologia também contribuem para a modelagem conceitual, porque é preciso considerar o mapeamento de um determinado domínio como um sistema de conceitos, cujos elementos são definidos a partir de diferentes tipos de relações: equivalência (sinonímia), lógicas (gênero-espécie, analítica e de oposição), ontológicas (partitiva, sucessão e material-produto), de efeito (causalidade, instrumental e descendência), entre outras, que formam a “estrutura” de uma linguagem comunicativa. Sendo assim, a Terminologia contribui, segundo Campos

<sup>4</sup> Não se pretende, neste estudo, detalhar as categorias de Ranganathan. Caso queira aprofundar nos estudos, deve-se consultar a literatura citada.

(2010), com os princípios teóricos para a criação de sistemas de conceitos bem estruturados. Isso acontece seja com a Teoria Geral da Terminologia, iniciada por Wüster, que se baseia nas dicotomias “termo e palavra”, seja com os princípios da Teoria Comunicativa da Terminologia (TCT), proposta por Cabré. Esta 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. A determinação do termo para representar o conceito tem papel essencial na Recuperação da Informação (RI), pois sempre podem haver diferentes termos para representar um conceito e, em certa medida, a escolha por um deles será prescritiva.

Neelameghan (1971) desenvolveu um modelo para representar o conhecimento a partir de uma estrutura de facetas generalizadas de assuntos, conforme Figura 7, no qual o processo de formação de conceitos (assuntos) se realiza a partir das formas linguísticas universais conforme expostas por linguistas como Chomsky, Fodor e Katz. Essa estrutura pode ter modelos específicos para diferentes campos de assunto, com a formação das palavras e a formação de assuntos que ocorrem de forma instantânea. Nesta proposta, o autor baseou-se nos princípios da Teoria Geral da Classificação para criar uma sequência útil, utilizando a proposta de Ranganathan, para a formação de assuntos a partir da criação de facetas sintáticas. O modelo proposto por Neelameghan (1971) pode auxiliar no mapeamento de um domínio do conhecimento, a partir da modelagem contextual, por meio de análise sintática.

Figura 7: Estrutura de facetas generalizadas de assuntos desenvolvida por Neelameghan.



Fonte: Dutta & Dutta, 2014, p. 60.

Dutta e Dutta (2014) descrevem os modos de formação de palavras polilexicais na linguística e os comparam com os diferentes modos de formação de assuntos compostos e complexos, como retratado por Ranganathan (1967), Neelameghan (1971, 1973) e Sen (2009). Os autores concluem que há semelhanças entre os modos de formação de palavras e de assuntos.

Esses elementos lógicos e sistematizadores de domínios apresentados por teóricos do âmbito da Biblioteconomia e da Ciência da Informação evidenciam toda uma linha de discussão travada neste campo com vista à representação de conhecimento para a ordenação de documentos e informação.

Nesse sentido, essas teorias dão conta de um espaço onde a interpretação é realizada por humanos: são modelos complexos, mas que não foram construídos para serem processados automática-



mente. Assim, como produtos dos contextos onde estamos inseridos, o ensino também, até o advento das propostas da Web Semântica, vinha se conformando a estratégias teóricas e metodológicas da Web Sintática para dar conta de tais contextos.

### **3. Modelos de representação de domínios para no novo contexto da Web Semântica**

Até o surgimento da Web Semântica, a separação entre documento e instrumento de tratamento estava bem definida. Tesouros e Taxonomias são exemplos de instrumentos que explicitam os conceitos de um domínio e permitem que se tenha um quadro de conceitos e relações a partir de um dado contexto de uso. Esses instrumentos são utilizados para a representação de conhecimentos de um dado domínio, porém, não possuem mecanismos para a derivação de conhecimentos, como é o caso das Ontologias, que se inserem no âmbito da Web Semântica. Nas ontologias, mais precisamente nas ontologias formais, o conhecimento pode ser obtido por meio de inferências sobre o conhecimento imediato disponível. Uma ontologia formal, ao mesmo tempo em que pode ser considerada uma ferramenta de representação, explicita e deriva também o conhecimento sobre um dado domínio (Campos, 2010).

Este novo cenário vai requerer, do profissional da informação, a aquisição e articulação de novos saberes, assim como novas estratégias para o ensino e a pesquisa na área.

A definição mais clássica de Web Semântica é a de Berners-Lee, Lassila e Hendler (2001, p.1), que a apresenta “como uma extensão da Web atual, na qual é dada à informação um significado bem definido, permitindo que computadores e pessoas trabalhem em cooperação”. Dessa forma, a proposta da Web Semântica é incorporar ligações semanticamente mais ricas, capazes de dar suporte às

aplicações em negociações. Para isso, a Web Semântica requer que as pessoas façam um esforço extra na codificação de informações em representações passíveis de processamento automático. Com esse esforço, os computadores terão condições de melhor processar, interpretar e concatenar conceitos.

Por outro lado, apesar de os mecanismos de busca terem avançado significativamente, o que ocorre é que a seleção dos dados recuperados, na maioria das vezes, não sofre nenhuma seleção. Recuperamos a informação, mas, por vezes, sem nenhuma precisão. A Semântica que se propõe a esta nova geração da Web, a Web Semântica, pode auxiliar a conseguirmos recuperação com maior consistência. No entanto, é necessário que se tenha claro que a “Web Semântica não é uma Web separada, e sim uma extensão da Web (sintática) atual. Nessa nova Web, a informação terá significado bem definido através de linguagens de marcação semântica” (Breitman, 2005, p. 9). E ainda:

A Web semântica não é Inteligência Artificial. O conceito de documentos compreensíveis por máquinas não implica uma inteligência artificial mágica que faz com que os computadores passem a entender o que os seres humanos falam. Esse conceito apenas indica que computadores passarão a ter a habilidade de resolver problemas bem definidos através do processamento de operações que se utilizam de dados. Em vez de querer que computadores “entendam” a linguagem das pessoas, pedimos que as pessoas façam um esforço extra (codificando a informação em representações passíveis de processamento automático, ontologias, por exemplo). (Breitman, 2005, p. 9).

Para que isso venha a acontecer, será necessário utilizar formalismos. Dito de outra forma, que possa existir uma semântica formal

para os conceitos definidos em linguagem natural. Nesse sentido, as ontologias de fundamentação exercem um papel preponderante.

Segundo Chandrasekaran, Josephson e Benjamins (1999, p. 20), na área de Inteligência Artificial, uma ontologia:

é um vocabulário de representação, freqüentemente especializado em um domínio ou assunto. Mais precisamente não é o vocabulário, como tal, que se qualifica como uma ontologia, mas as conceitualizações que os termos do vocabulário destinam-se a capturar.

Segundo Guarino (1997, p. 16), conceitualização pode ser definida como “uma estrutura semântica intensional que codifica as regras implícitas, restringindo a estrutura de um pedaço da realidade”.

As Ontologias, diferentemente de outras ferramentas semânticas, apoiam-se em mecanismos de representação baseados em lógica, que permitem definir, por meio de axiomas, os conceitos nelas contidos e suas relações. As relações também possuem atributos que facilitam a descoberta de conhecimentos. Esses atributos indicam, por exemplo, se a relação é inversa (ex. *filho de e pai de*), ou simétrica (ex. *irmão de*), ou ainda transitiva (ex. *pré-requisito de*), entre outros. Porém, o uso de tais recursos depende da forma de modelagem dos conceitos adotada na ontologia.

Para que o seu potencial seja plenamente realizado, as ontologias na Web Semântica devem lançar mão de recursos formais (lógicos), de modo que permitam explicitar conceitos de diferentes graus de expressividade e dar conta de requisitos computacionais complexos (Ding et al., 2006). Nesse sentido, espera-se que seja possível expressar não só a identidade dos conceitos e seus atributos, mas também as suas relações com outros, que dependem do contexto e de seu compromisso ontológico. Algumas dessas relações expressam aspectos dinâmicos, como, por exemplo, o papel que uma pessoa

exerce e que pode mudar com o tempo, enquanto outras podem ser influenciadas pelo contexto social. Estas últimas colocam em destaque a interação das pessoas com os recursos descritos pelas ontologias, podendo, de acordo com o perfil e contexto do usuário, sugerir associações com outros recursos potencialmente de interesse (Mendes et al., 2010).

Na prática, a Ontologia Formal de Fundamentação, um tipo de ontologia, pode ser entendida como a teoria das distinções *a priori* sobre as entidades do mundo (objetos físicos, eventos, regiões, quantidades de matéria) e as categorias de metanível para modelar o mundo (conceitos, propriedades, qualidades, estados, papéis e partes) (Guarino, 1998).

Mais precisamente, as ontologias de fundamentação: têm um âmbito de aplicação extenso, podem ser reutilizadas para modelar diferentes situações, são bem fundamentadas do ponto de vista filosófico-conceitual, são axiomatizadas e, por isso, semanticamente transparentes. Elas se concentram em conceitos muito gerais e basilares, como os conceitos de objeto, evento, qualidade, papel, e em relações como constituição, participação, dependência, parte-todo. Tais conceitos e relações não são especificados para um domínio restrito, mas podem ser aperfeiçoados para atender a várias exigências aplicativas. A transparência semântica é oferecida por uma seleção cuidadosa dos termos primitivos e por uma caracterização precisa de seu significado.

Podemos observar que, no espaço de uma Web Semântica, as ferramentas utilizadas para o controle terminológico e recuperação de informação requerem a expansão de seus mecanismos de atuação, ou seja, se os vocabulários controlados e tesauros visavam ao controle terminológico que permitia a tradução da linguagem apresentada em documentos em uma linguagem artificial, a ontologia objetiva algo mais: ser um artefato tecnológico que contém um conjunto de

regras que delimitam o significado intensional<sup>5</sup> de um vocabulário formal, permitindo que, a partir de um compromisso ontológico<sup>6</sup>, o conhecimento existente em um domínio possa ser representado e, assim, compartilhado e inferido.

Pode-se observar que, neste novo espaço, novas estratégias intelectuais serão requeridas para representar o conhecimento de um domínio e codificá-lo em uma linguagem que permita a leitura por máquinas. As representações devem ser feitas por meio de proposições lógicas, ou seja, os conteúdos das informações devem ser declarativos. Essas proposições são apresentadas na forma de axiomas, que permitam a realização de inferências com base nas proposições lógicas estabelecidas. Além da estrutura taxonômica, a qual é responsável pelo mapa estrutural do domínio coberto pela ontologia, os axiomas são uma chave essencial à construção de ontologias. O detalhamento dos axiomas é necessário para que as respostas oferecidas pelo sistema sejam confiáveis e baseadas no compromisso ontológico estabelecido pela própria ontologia (Medeiros, 2009).

Consideramos importante apontar que, para a elaboração de modelos de representação de domínios para o novo espaço da Web Semântica, o profissional da informação deverá se apropriar de conhecimentos que vão além das teorias terminológicas e das lógicas de classificação. São importantes as noções filosóficas e epistemológicas contidas em ontologias formais. O entendimento dessas noções e as notações adotadas para a sua modelagem são complexas, mas oferecem um campo vasto de aplicação, do qual o profissional da informação não pode se omitir.

---

<sup>5</sup> O conceito de “intensional”, neste contexto, pode ser considerado como as características e atributos de um conceito que permitem determinar relações entre os conceitos.

<sup>6</sup> Compromisso ontológico – acordo estabelecido sobre o que se compreende sobre um conceito, no interior de um domínio ou contexto de atividade.

#### 4. Considerações Finais

A questão que se coloca é a necessidade de nos empenharmos na busca de estratégias de ensino e aprendizagem que considerem novas formas de representação de informação e de conhecimento, para que a máquina “entenda” o conteúdo a representado.

No evento CIFORM/UFBA de 2015 (Campos, 2015), foram discutidas algumas questões relacionadas à formação do profissional de informação para este novo ambiente. Em evento da comunidade ontologista, no domínio da Ciência da Computação, o “Ontology Summit” já está traçando um núcleo básico para a formação do ontologista do futuro ou mesmo, pode-se dizer, do “classificacionista do futuro”, ou seja, aquele que constrói modelos conceituais de domínios. Tal discussão é necessária no espaço das pesquisas e ensino em Ciência da Informação.

Esta formação envolve conteúdos como: 1. **Fundamentos Teóricos:** Lógica (primeira ordem, de segunda ordem, lógica descritiva, lógica de definições); Teoria dos conjuntos, Ontologia filosófica (universais e particulares, tempo); Filosofia da linguagem (a confusão entre sentido e referência, a teoria dos atos de fala,...); Representação do conhecimento, modelagem conceitual, modelagem de dados, metadados; 2. **Linguagens:** RDF, OWL,...; 3. **Construção e edição de ontologias:** aspectos humanos (aplicação de princípios de classificação), ferramentas de software (Protégé,...), problemas de interoperabilidade semântica; 4. **Estratégias e teorias de avaliação de ontologias;** 5. **Teorias e Metodologias de elaboração de ontologias:** ontologias de fundamentação, ontologias de domínio. Além disso, registram-se também as disciplinas afins, a saber: **Lógica avançada** (lógica modal, lógica temporal, lógica padrão,...); **Ontologia Filosófica Avançada** (mereotopology, tropos...); **Ciência da Computação:** linguagens formais, máquinas formais, computabilidade, raciocínio automatizado, teoria de banco de dados; **Linguística / Ciências**

**Cognitivas:** distinção entre sintaxe, semântica, pragmática, processamento de linguagem natural, geração de linguagem natural, teorias cognitivas de categorização.

Assim, no âmbito deste capítulo, nossa intenção foi lançar algumas discussões sobre os conteúdos até hoje ministrados e aqueles que deverão ser revistos e acrescentados, visando à formação de profissionais de informação aptos a atuar efetivamente no desenvolvimento de modelos conceituais de domínios para uso na Web Semântica.

## Referências Bibliográficas

- Aristóteles. (2010). *Órganon*. São Paulo: Edipro.
- Barros, L. A. (2004). *Curso básico de terminologia*. São Paulo: EdUSP.
- Berners-Lee, T., Hendler, J., & Lassila, O. (2001). The Semantic Web. *Scientific American*, 284(5), 34-43. Recuperado de <http://www.sciam.com/article.cfm?articleID=00048144-10D2-1C70-84A9809EC588EF21>. [Acessado em 05/11/18]
- Bliss, H. E. (1929). *The organization of knowledge and the system of the sciences*. New York, NY: Holt and Company.
- Breitman, K. (2005). *Web Semântica: a internet do futuro*. Rio de Janeiro: LTC.
- Campos, M. L. A. (2010). O documento e as ferramentas de tratamento e recuperação de informações na Web Semântica: um novo espaço de identidade. In L. S. Freitas, C. H. Marcondes, & A. C. Rodrigues (Orgs.), *Documento: Gênese e contextos de uso – Estudos da informação* (Vol. 1, pp. 223-234). Niterói: EdUFF.
- Campos, M. L. A. (2012). Ontologia de Fundamentação como uma teoria de representação para a Web semântica: estratégias interdisciplinares na pesquisa e no ensino. In E. I. Murguia & M. E. F. Rodrigues (Orgs.), *Arquivologia, Biblioteconomia e Ciência da Informação: identidades, contrastes e perspectivas de interlocução* (Vol. 2, pp. 101-116). Niterói: EdUFF.
- Campos, M. L. A. (2015). Representação de Domínios na Web Semântica: Desafios para a formação de profissionais de informação. In *Anais do XII CINFORM – Encontro Nacional de Ensino e Pesquisa em Informação* (pp. 73-93). Salvador, BA, Brasil: ICI/ EdUFBA. Recuperado de <https://repositorio.ufba.br/ri/bitstream/ri/18051/1/ANAIS%20DO%20XII%20CINFORM%20ANO%202015.pdf>
- Chandrasekaran, B., Josephson, J. R., & Benjamins, V. R. (1999). What are ontologies, and why do we need them? *IEEE Intelligent Systems*, 14(1), 20-26. doi:10.1109/5254.747902
- Dahlberg, I. (1978a). A referent-oriented analytical concept theory of interconcept. *International Classification*, 5(3), 142-150. doi:10.5771/0943-7444-1978-3-142

- Dahlberg, I. (1978b). Teoria do conceito. *Ciência da Informação*, 7(2), 101-107. Retirado de <http://revista.ibict.br/ciinf/article/view/115/115>. [Acessado em 05/11/18].
- Ding, L., Kolari, P., Ding, Z., & Avancha, S. (2007). Using ontologies on the semantic web: a survey. In R. Sharman, R. Kishore, & R. Ramesh (Eds.), *Ontologies: A Handbook of Principles, Concepts, and Applications in Information Systems* (pp. 79-113). Boston, MA: Springer.
- Dutta, B., & Dutta, C. (2014). A linguistic view of subject formation process as described by Ranganathan and Others. *Annals of Library and Information Studies*, 61(1), 56-64. Recuperado de <http://hdl.handle.net/10760/23304> [Acessado em 05/11/18]
- Feibleman, J. K. (1954). Theory of Integrative Levels. *British Journal for the Philosophy of Science*, 5(17), 59-66. doi:10.1093/bjps/V.17.59
- Green, R. (2008). Relationships in Knowledge Organization. *Knowledge Organization*, 35(2/3), 150-159. doi:10.5771/0943-7444-2008-2-3-150
- Guarino, N. (1997). Semantic Matching: Formal Ontological Distinctions for Information Organization, Extraction, and Integration. In Pazienza, M. T. (Ed.), *Information Extraction: A Multidisciplinary Approach to an Emerging Information Technology. Proceedings of SCIE '97 International Summer School on Information Extraction, Frascati, Italy* (pp. 139-170). Heidelberg: Springer. Recuperado de <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.20.7515>. [Acessado em 05/11/18]
- Guarino, N. (1998). Formal Ontology and Information Systems. In N. Guarino (Ed.), *Formal Ontology in Information Systems: Proceedings of FOIS'98, Trento, Italy* (pp. 3-15). Amsterdam: IOS Press. Recuperado de <https://klevas.mif.vu.lt/~donatas/Vadovavimas/Temos/OntologiskaiTeisingasKonceptinisModeliavimas/papildoma/Guarino98-Formal%20Ontology%20and%20Information%20Systems.pdf>
- Lima, G. A. (2018). Aspectos cognitivos na categorização: buscando o contexto. In G. A. Lima (Org.), *Bibliotecas digitais: novas tendências na navegação em contexto* (Vol. 1, pp. 1-12). Rio de Janeiro: Interciência.
- Medeiros, J. S. (2011). *Tesouros conceituais e ontologias de fundamentação: análise comparativa entre as bases teórico-metodológicas utilizadas em seus modelos de representação de domínios* (Dissertação de mestrado em Ciência da Informação, Instituto de Arte e Comunicação Social, Universidade Federal Fluminense, Niterói, RJ, Brasil). Recuperado de [http://www.ci.uff.br/ppgci/arquivos/Dissert/Diss\\_JacksonMedeiros.pdf](http://www.ci.uff.br/ppgci/arquivos/Dissert/Diss_JacksonMedeiros.pdf)
- Mendes, P. N., Kapanipathi P., Cameron, D. H., & Sheth, A. P. (2010). *Dynamic associative relationships on the linked open data web*. Paper presented at the Web Sci10: Extending the Frontiers of Society On-Line, Raleigh, NC, USA. Retirado de <https://corescholar.libraries.wright.edu/knoesis/760>
- Morujão, C. (2006). A lógica modernorum: lógica e filosofia da linguagem na escolástica dos séculos XIII e XIV. *Revista Filosofia de Coimbra*, 14(28), 301-322. Recuperado de [http://www.uc.pt/fluc/dfci/publicacoes/a\\_logica\\_modernorum](http://www.uc.pt/fluc/dfci/publicacoes/a_logica_modernorum). [Acessado em 05/11/18].
- Neelameghan, A. (1971). Sequence of Component Ideas in a Subject. *Library Science with a Slant to Documentation*, 8, 323-334.
- Neelameghan, A. (1973). Basic Subject. *Library Science with a Slant to Documentation*, 10. Papers F to N.



- Piedade, M. A. (1983). *Introdução à teoria da classificação* (2ª ed.). Rio de Janeiro: Interciência.
- Raghavan, K. S. (1985). The general theory of classification as the basis for structuring of subject headings. In *Proceedings of the Regional Conference of International Federation for Documentation Committee On Classification Research (FID-CR), New Delhi, India* (pp. 24-48). New Delhi: Delhi Library Associations.
- Ranganathan, S. R. (1967). *Prolegomena to library classification* (3rd ed.). London: Asia Publishing House.
- Sen, B. K. (2009). Universe of knowledge from a new angle. *Annals of Library and Information Studies*, 56(1), 7-12.
- Vickery, B. C. (1997). Ontologies. *Journal of Information Science*, 23(4), 227-286. doi:10.1177/016555159702300402
- Wilson, T. D. (1972). The Work of the British Classification Research Group. In H. Wellisch, & T. D. Wilson (Eds.), *Subject retrieval in the seventies* (pp. 62-71). Westport, CT: Greenwood Publishing Co.

