

WALTER MOREIRA

A CONSTRUÇÃO DE INFORMAÇÕES DOCUMENTÁRIAS: APORTES DA  
LINGUÍSTICA DOCUMENTÁRIA, DA TERMINOLOGIA E DAS ONTOLOGIAS

Tese apresentada à Escola de Comunicações e Artes da  
Universidade de São Paulo como requisito parcial para  
obtenção do título de Doutor em Ciência da Informação.

Área de Concentração: Cultura e informação

Linha de pesquisa: Organização da informação e do  
conhecimento

Orientadora: Profa. Dra. Marilda Lopes Ginez de Lara

SÃO PAULO, 2010

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

Catálogo da Publicação  
Escola de Comunicações e Artes  
Universidade de São Paulo

Moreira, Walter.

A construção de informações documentárias: aportes da linguística documentária, da terminologia e das ontologias / Walter Moreira; Orientadora: Marilda Lopes Ginez de Lara. - São Paulo, 2010.

155 f. : il.

Tese (Doutorado em Ciência da Informação)--Universidade de São Paulo, 2010.

1. Língua documentária. 2. Terminologia. 3. Ontologias. 4. Informação documentária. I. Lara, Marilda Lopes Ginez de. II. Título.

CDD 025.3

MOREIRA, Walter. **A construção de informações documentárias**: aportes da linguística documentária, da terminologia e das ontologias. Tese apresentada à Escola de Comunicações e Artes da Universidade de São Paulo para obtenção do título de Doutor em Ciência da Informação. Área de Concentração: Cultura e informação. Linha de pesquisa: Organização da informação e do conhecimento.

Aprovado em: 23 de abril de 2010.

Banca Examinadora

Profa. Dra. Marilda Lopes Ginez de Lara  
(Universidade de São Paulo)

Prof. Dr. Carlos Henrique Marcondes de Almeida  
(Universidade Federal Fluminense)

Prof. Dr. Renato Rocha Souza  
(Universidade Federal de Minas Gerais)

Profa. Dra. Anna Maria Marques Cintra  
(Universidade de São Paulo)

Profa. Dra. Nair Yumiko Kobashi  
(Universidade de São Paulo)

Para Marilda, meu AMOR, com quem tenho uma união absoluta, que foi, certamente, planejada pela Mãe.

Para Raíssa e Larissa, nossas adoráveis filhas, por todo amor que me inspiram.

Para minha mãe, Maria, meu pai, João (*in memoriam*) e meus irmãos, que sempre me ensinaram sobre o verdadeiro sentido do amor.

# AGRADECIMENTOS

À Profa. Dra. Marilda Lopes Ginez de Lara, minha querida e competente orientadora, que me deu inúmeras demonstrações de que a vida intelectual não se separa da vida afetiva e que orientou este trabalho com a paixão pelo conhecimento e a segurança que é típica apenas dos grandes Mestres.

À Profa. Dra. Nair Yumiko Kobashi, por acreditar em mim, pelo exemplo de dignidade e coragem e por me apontar o caminho da pesquisa com seus comentários sempre repletos de valor.

Ao Prof. Marcos Luiz Mucheroni, pela contribuições oferecidas no exame de qualificação.

À Profa. Dra. Anna Maria Marques Cintra, que me abriu as portas da Escola de Comunicações e Artes e me ensinou com práticas o que é ser Professor.

Ao meus amigos da Biblioteca Conde de Moreira Lima, das Faculdades Integradas Teresa D'Ávila (Fatea): Cristina (minha bibliotecária *virtual*), Luis, Graça, Cléia, Nilce, Vânia, pelo sorriso constante e pela disposição permanente em ajudar.

À Silvia (Multiofício), que me surpreendeu com sua delicadeza, artigo cada vez mais raro.

Aos colegas da pós, pelos debates, pelos cafés e pela partilha do conhecimento.

*O senhor sabe o que o silêncio é? É a gente mesmo, demais.*

*Vivendo, se aprende; mas o que se aprende,  
mais, é só a fazer outras maiores perguntas.*

*Guimarães Rosa, Grande sertão: veredas*

# RESUMO

MOREIRA, Walter. **A construção de informações documentárias**: aportes da linguística documentária, da terminologia e das ontologias. 2010. 156 f. Tese (Doutorado em Ciência da Informação) – Universidade de São Paulo, Escola de Comunicações e Artes, São Paulo, 2010. [Orientadora: Profa. Dra. Marilda Lopes Ginez de Lara]

Investiga a interface teórico-prática entre terminologia, ontologia filosófica, ontologia computacional e linguística documentária e os subsídios que oferecem para a construção de informações documentárias. Foram estabelecidos como objetivos específicos, a análise das condições de produção, desenvolvimento, implementação, uso e integração de ontologias com base no referencial teórico da ciência da informação; a investigação sobre a contribuição das ontologias para o desenvolvimento de tesouros e vice-versa; a discussão acerca do fundamento filosófico da aplicação de ontologias com base no estudo das categorias ontológicas presentes na filosofia clássica e nas propostas contemporâneas. Defende que a compreensão das ontologias por meio da teoria comunicativa da terminologia colabora para a organização de um acesso menos quantitativo (sintático) e mais qualitativo (semântico) à informação. Observa que, conquanto compartilhem alguns objetivos comuns, ainda há pouco diálogo entre a ciência da informação (e, no seu interior, a linguística documentária) e a ciência da computação. Defende que as ontologias concretas e as ontologias filosóficas não são eventos completamente independentes que guardam entre si apenas a similaridade de denominação e constata que a discussão sobre categorias e categorização na ciência da computação, nem sempre possui a ênfase que recebe na ciência da informação no âmbito dos estudos sobre representação do conhecimento. A abordagem do rizoma, de Deleuze e Guattari, foi tratada como provocadora de reflexões sobre a validade do modelo hierárquico arborescente e sobre as possibilidades de sua ampliação. Conclui que a construção de ontologias não pode prescindir do tratamento terminológico-conceitual, como compreendido pela terminologia e pela ciência da informação, acumulado nos referenciais teóricos e nas metodologias para construção de linguagens documentárias e que, por outro lado, a construção de linguagens documentárias mais flexíveis não pode ignorar o modelo de representação das ontologias, com mais predisposição para a formalização e para a interoperabilidade.

Palavras-chave: Terminologia. Ontologia. Linguística documentária. Informação documentária.

# ABSTRACT

MOREIRA, Walter. **Documentary information**: contributions of ontology, terminology and documentary linguistic. 2009. 131 f Thesis (Ph.D. in Information Science) – Universidade de São Paulo, Escola de Comunicações e Artes, São Paulo, 2009. [Advisor: Profa.. Dra. Marilda Lopes Ginez de Lara]

Investigates theoretical and practical interfaces between terminology, philosophical ontology, computational ontology and documentary linguistics and the subsidies that they offer for the construction of documentary information. It was established as specific objectives, the analysis of the production, development, implementation and use of ontologies based on the information science theories, the research on the contribution of ontologies for the development of thesauri and vice versa and the discussion of the philosophical foundation of the application of ontologies based on the study of ontological categories present in classical philosophy and in the contemporary proposals. It argues that the understanding of ontologies through the communicative theory of terminology contributes to the organization of a less quantitative access (syntactic) and more qualitative (semantic) of information. Notes that, in spite of sharing some common goals, there is little dialogue between the information science (and, inside it, the documentary linguistics) and computer science. It argues that the computational and philosophical ontologies are not completely independent events, which have among themselves only the similarity of name, and notes that the discussion of categories and categorization in computer science, does not always have the emphasis it receives in information science in studies on knowledge representation. The approach of Deleuze and Guattari's rhizome, was treated as instigator of reflections on the validity of the hierarchical tree model structure and the possibilities of its expansion. It concludes that the construction of ontologies can not ignore the terminological and conceptual analysis, as it's understood by the terminology and by the information science accumulated in the theoretical and methodological basis for the construction of indexing languages and, on the other hand, the construction of flexible indexing languages can not ignore the representational model of ontologies which are more capable for formalization and interoperability.

Keywords: Terminology. Ontology. Documentary linguistics. Documentary information.

# LISTA DE ILUSTRAÇÕES

Figura 1 - A escola de Atenas.....	30
Figura 2 - Árvore de Porfírio .....	37
Figura 3 - Árvore de Brentano .....	39
Figura 4 – Exemplo de aplicação da tríade peirceana .....	44
Figura 5 – Exemplo de componentes de ontologias .....	54
Figura 6 – As variações linguísticas geográficas como empecilho para o alinhamento de ontologias .....	66
Figura 7 - Triângulo semântico .....	100

# LISTA DE QUADROS

Quadro 1 - Enunciados lógicos .....	32
Quadro 2 - Categorias aristotélicas .....	35
Quadro 3 - Tábua dos juízos e categorias de Kant .....	42
Quadro 4 - Categorias universais de Peirce.....	45
Quadro 5 - Conjunto de elementos Dublin Core.....	58
Quadro 6 - Grupos de pesquisa sobre ontologias em ciência da informação .....	60
Quadro 7 – Relatividade cultural no significado das palavras .....	65
Quadro 8 - Níveis de representação do conhecimento .....	109
Quadro 9 - Comparação entre rizoma e internet.....	122

# LISTA DE TABELAS

Tabela 1 - Artigos sobre ontologia em periódicos nacionais de ciência da informação .....21

Tabela 2 - Artigos sobre ontologia em periódicos internacionais de ciência da informação ..21

Tabela 3 - Co-ocorrência das palavras-chave mais utilizadas .....92

# SUMÁRIO

<b>1 Introdução .....</b>	<b>14</b>
1.1 Tema e problema.....	18
1.2 Hipóteses .....	23
1.3 Objetivos.....	24
1.4 Justificativa .....	25
1.5 Estrutura da tese .....	27
 <b>2 A base filosófica das ontologias .....</b>	<b>28</b>
2.1 Categorias aristotélicas .....	34
2.2 A árvore de Porfírio.....	35
2.3 A leitura de Brentano.....	38
2.4 As categorias de Kant .....	39
2.5 A tríade de Peirce.....	43
2.6 Categorias, ontologias e informação documentária.....	46
 <b>3 Ontologias: a compreensão das ciências da informação e da computação.....</b>	<b>50</b>
3.1 Definições e características das ontologias.....	50
3.2 Metadados e ontologias: construção, comunicação e uso da informação.....	55
3.3 Interoperabilidade semântica.....	60
3.4 As relações entre interoperabilidade e categorias.....	67
 <b>4 Teorias da terminologia .....</b>	<b>72</b>
4.1 Teoria Geral da Terminologia.....	72
4.2 Teoria Comunicativa da Terminologia.....	73
 <b>5 Linguística Documentária .....</b>	<b>82</b>
5.1 Linguagem Documentária .....	85
5.2 Linguística documentária e terminologia.....	87

<b>6 As ontologias e a construção de informações documentárias .....</b>	<b>94</b>
6.1 <i>Sistemas de informação documentária: alguns esclarecimentos terminológicos .....</i>	94
6.2 <i>Tesauros e ontologias .....</i>	99
6.3 <i>Informação documentária e terminologia .....</i>	104
6.4 <i>Informação documentária e representação do conhecimento .....</i>	108
 <b>7 A árvore e o rizoma: a proposta de Deleuze .....</b>	 <b>112</b>
7.1 <i>Antes, uma pequena entrada na sociedade da informação .....</i>	112
7.2 <i>A lógica da árvore como modelo de representação .....</i>	115
7.3 <i>A multiplicidade e a árvore .....</i>	123
7.4 <i>Reverendo as estruturas .....</i>	125
7.5 <i>As provocações de Deleuze .....</i>	129
 <b>8 Considerações finais .....</b>	 <b>131</b>
 <b>Referências .....</b>	 <b>139</b>
 <b>Apêndices .....</b>	 <b>150</b>
Apêndice A – <i>Artigos sobre ontologia em periódicos internacionais de Ciência da Informação .....</i>	150
Apêndice B – <i>Grupos de pesquisa sobre ontologias no Brasil .....</i>	155

# 1 INTRODUÇÃO

Talvez estes tempos sejam marcados e reconhecidos no futuro como a era das redes. O conceito de rede tornou-se pré-requisito para pensar a sociedade contemporânea; não apenas na vertente mais filosófica do conceito, mas também em sua acepção tecnológica. Não se pode negar o avanço que representam as condições físicas (computadores e tecnologias necessárias para conexão) no estabelecimento dos meios de comunicação, de modo geral. Não seria possível a internet, por exemplo, sem sua complexa estrutura de *backbones* e *infovias*. As condições físicas, entretanto, possibilitam não apenas conexões entre máquinas, mas também propiciam, esta sim é a grande mudança, incontáveis formas de conexões entre pessoas que permitem construir ou ampliar redes de relacionamentos nos mais diversos níveis. Não se fala apenas do *facebook*, do *orkut*, do *twitter* e seus assemelhados, mas das facilidades de intercomunicação de modo amplo, mesmo, e principalmente, na esfera acadêmica, como acontece, por exemplo, nas comunicações informais dos pesquisadores (MOREIRA, 2005). Dadas a pluralidade de fontes e de usuários de informação, reais ou potenciais, há que se pensar em macrossoluções que tangenciem, por semelhança, aquelas que Otlet (1996) propusera.

As instituições preocupadas hoje com a organização e a circulação da informação, em todas as esferas e níveis, abraçam (ou precisam abraçar) o diálogo e o intercâmbio como condições para participar do mundo contemporâneo. Talvez a proposição pareça exagerada quando são consideradas as bibliotecas menores e suas preocupações regionalizadas, mas a condição de isolamento dessas pode converter-se exatamente na causa de seu fracasso. Em outras palavras: considerando-se a perspectiva instaurada pelas redes de comunicação não

se admite mais um sistema de informação que não dialogue com outros sistemas, mesmo porque sua própria existência tornar-se-ia neste caso um contrassenso.

Podem-se negar diversas coisas ao computador, pode-se considerar ingenuidade, como de fato é, atribuir-lhe o estatuto de maior invenção humana desde o alfabeto, como tem sido apregoado em diversos discursos, mas é inegável que o computador é, em todos os tempos, a ferramenta mais adequada para a organização e disseminação da informação, principalmente da informação documentária. De pontos isolados, quando do seu surgimento, os computadores transformaram-se em pontos virtuais de acesso a uma vasta rede hipermediática. Podem, potencialmente, facilitar o acesso a um *googol*<sup>1</sup> de informações produzidas pelo homem, registradas nos mais diversos suportes. Na realidade, contudo, ainda permanecem, sob certos aspectos, isolados, praticamente incomunicáveis entre si, quando não há qualidade suficiente na interação homem-máquina, em nível ainda sintático.

A proposta da *web* semântica volta-se exatamente para este problema. Na *web* atual (sintática) os computadores apenas apresentam a informação e deixam o processo de interpretação por conta dos seres humanos. Evidentemente que a tarefa de interpretação é mais complexa, mas ainda assim é possível diminuir a quantidade de interferência humana no processo, atribuindo à informação um significado formalmente definido por meio do uso de ontologias (BERNERS-LEE; HENDLER; LASSILA, 2001). Algumas definições de *web* semântica, aliás, soam extremamente familiares ao campo da ciência da informação. Veja-se, por exemplo, esta, tomada de Perojo e León (2005<sup>2</sup>, tradução nossa):

A *web* semântica é uma área fértil, localizada no cruzamento da inteligência artificial com as tecnologias *web*, que propõe novas técnicas e paradigmas para a representação da informação e do conhecimento, com o fim de facilitar tanto a localização quanto o compartilhamento, a integração e a recuperação de recursos.

---

<sup>1</sup> Um *googol* equivale ao número um seguido de 100 zeros, ou seja,  $10^{100}$ . Foi criado em 1938, pelo matemático Edward Kasner. A expressão inspirou o nome da empresa e do famoso mecanismo de busca.

<sup>2</sup> Algumas citações diretas aparecem sem indicação de paginação porque não foram encontradas nos documentos eletrônicos correspondentes.

Em se tratando de atribuir semântica aos fluxos de informação na *web*, as ontologias têm sido amplamente estudadas na ciência da computação e, na condição de instrumentos de representação do conhecimento, também interessam à ciência da informação.

O que se pretende com esta reflexão que se inicia, sobre as ontologias e suas aplicações na representação do conhecimento e na recuperação de informações, é identificar e discutir, com o referencial da ciência da informação, mecanismos que possam facilitar os processos de interação e de interatividade, por meio de agentes inteligentes. Estes mecanismos devem levar em conta, necessariamente, a redução das barreiras conceituais ou terminológicas.

Neste trabalho manifesta-se, essencialmente, preocupação com as condições pelas quais se dá essa intercomunicação no domínio da ciência e tecnologia. Acredita-se que a linguística documentária acumulou conhecimento suficiente para participar mais ativamente da discussão com aporte de campos como a terminologia e as ontologias.

Os estudos já realizados sobre as linguagens documentárias atestam sua importância como fator de qualidade para os processos de produção e recepção (mais especificamente a recuperação) da informação. Acredita-se que o avanço teórico alcançado não é proporcional à sua aplicação prática, a qual ainda se ressentir de subsídios ou metodologias que possam oferecer garantias para a equivalência de sentido entre a informação documentária (representação) e o texto fonte, como já advertiu Kobashi (1994), que dedicou uma tese à elaboração de metodologia para a construção de informações documentárias.

A razão deste descompasso é o dinamismo inerente aos processos de representação e recuperação da informação documentária. Muitas metodologias e técnicas desenvolvidas para sistemas de informação implantados sob o paradigma do acervo ou de redes domésticas rapidamente tornaram-se obsoletas e precisam ser continuamente repensadas numa sociedade que pensa e age globalmente e em rede. Os índices de concordância *kwic* (*keyword in context*) e *kwoc* (*keyword out of context*) e o sistema unitermo, por exemplo, mostraram-se interessantes num contexto inicial de automação dos sistemas de informação

documentária, mas foram rapidamente superados, porquanto ignoravam aspectos semânticos da informação.

Do entusiasmo inicial com a pesquisa em texto completo passou-se à frustração com a qualidade da informação recuperada por esse meio. Evidentemente diversas barreiras estão sendo rompidas e inúmeros projetos de sucesso facilitam a vida do pesquisador contemporâneo. Vejam-se, por exemplo, os repositórios de informação, os arquivos abertos e suas facilidades. Uma quantidade desafiadora equivalente, senão superior, de problemas tem sido igualmente proposta, notadamente os que se relacionam com a qualidade da recuperação da informação, já que os aspectos de quantidade estariam supostamente garantidos.

Do mesmo modo como, a partir de Otlet, diversas escolas brasileiras de biblioteconomia inseriram (tardamente) o termo *documentação* nas denominações de seus cursos, vem se tornando comum tratar qualquer contexto relacionado com a informação como concernente ao campo da ciência da informação. Neste caso, como naquele, a alteração da designação não possui correspondente qualitativo na reflexão, como já apontaram Barreto (1993) e Kobashi, Smit e Tálamo (2001). Nas palavras destas estudiosas, “a racionalidade técnica que, enquanto tal, é circunstancial e histórica, não foi substituída pela racionalidade acadêmico-reflexiva, fundamental para a institucionalização da identidade do campo” (KOBASHI; SMIT; TÁLAMO, 2001, p. 2). Em outras palavras, ainda faltam reflexões.

A incorporação e a associação definitiva da tecnologia à ciência da informação ocorrem a partir da década de oitenta. Anteriormente, contudo, em 1973, Foskett<sup>3</sup> (*apud* PINHEIRO; LOUREIRO, 1995) já demonstrava preocupação com os rumos da ciência da informação e sua exagerada ênfase na tecnologia, o que poderia reduzir a informação a uma *commodity*. Uma nova disciplina, defende Foskett (*op. cit.*, p. 49), “não surge porque aqueles

---

<sup>3</sup> FOSKETT, D. J. Ciência da informação como disciplina emergente: implicações educacionais. In: \_\_\_\_\_. *Ciência da informação ou informática?* Rio de Janeiro: Calunga, 1980. p. 53-69.

que a praticam aperfeiçoaram suas tarefas, mas porque dinamicamente novas relações aparecem com outros campos”.

Busca-se escapar, portanto, da tentação fácil de encarar o computador e suas tecnologias associadas como panaceia ou, por outra associação com a cultura grega, como o *deus ex machina* (o artefato que o teatro grego produziu para resolver inesperadamente, todos os problemas da trama). Este trabalho de pesquisa sustenta-se, então não apenas no caráter da racionalidade técnica, mas também da racionalidade reflexiva. Refletindo sobre a aplicação de ontologias na representação do conhecimento e na recuperação da informação, conforme as compreende a ciência da informação, procura compreender suas interpenetrações, seus empréstimos, suas mútuas influências.

## 1.1 TEMA E PROBLEMA

Um dos princípios fundamentais da comunicação relaciona-se ao compartilhamento razoável de conceitos que garantam algum nível de inteligência na emissão e na recepção da mensagem, ou seja, não será possível estabelecer comunicação sem algum grau de compreensão mútua. A ciência da informação compartilha com a ciência da comunicação esta preocupação quando toma a informação como objeto de estudo em seus aspectos de construção, comunicação e uso. Esta abordagem exige, naturalmente, que se estude igualmente a natureza da informação e sua gênese.

As condições para a comunicação efetiva da informação e do conhecimento formam, na definição clássica de Saracevic (1996, p. 47), a preocupação central da ciência da informação; entendida como

um campo dedicado às questões científicas e à prática profissional voltadas para os problemas de efetiva comunicação do conhecimento e de seus registros entre os seres humanos, no contexto social, institucional ou individual do uso e das

necessidades de informação. No tratamento destas questões são consideradas de particular interesse as vantagens das modernas tecnologias informacionais.

Ressaltam-se aqui as preocupações com “os problemas de efetiva comunicação do conhecimento” e as “vantagens das modernas tecnologias informacionais”. A primeira dá conta de um aspecto pragmático do objeto da ciência da informação, isto é, ainda que não se possa garantir totalmente a produção de conhecimento com a simples comunicação da informação, é essa a finalidade dos sistemas documentários. Podem-se discutir as garantias relativas à transferência da informação, mas sua finalidade parece mais estável.

A segunda preocupação destacada, referente às tecnologias, confere à definição um caráter indispensável às definições clássicas: a peremptoriedade. Pode-se, perfeitamente, compreender no escopo desta definição, por exemplo, o trabalho pioneiro de Paul Otlet. No que foi considerado por muitos como uma utopia, Otlet (1996) propõe utilizar, em 1934, a melhor das “modernas tecnologias informacionais” disponíveis no momento como suporte para sua ousada proposta: a ficha de cartolina no formato 12 x 7,5 cm. Outro aspecto que precisa ser destacado da fala de Saracevic (1996) diz respeito aos “seres humanos”. Compreende-se, para os efeitos desta pesquisa, que as ontologias, conforme o tratamento que a ciência da computação lhe confere, avançaram consideravelmente no que tange à representação do conhecimento em ambientes digitais. O que parece ainda problemático é que o elemento humano fica, desconcertantemente, relegado a *segundo plano*, em certo sentido. A interatividade máquina-máquina só tem sentido completo se e quando o homem aparece como ator nas duas pontas: homem-máquina-máquina-homem. Ainda que se possa *informar* a um agente inteligente ou a um sistema de informação sobre uma estrutura ontológica, ainda que se possa dotar uma máquina de uma base de conhecimento, qualquer uso que se possa fazer disso com um computador tem sido restrito, em última instância, ao domínio das relações sintáticas.

Se, por um lado, há relativa ausência de literatura sobre as ontologias no domínio da ciência da informação, por outro, este se tornou justamente um dos principais motores desta pesquisa. O fato é que a ciência da computação avança nestas questões, por meio da

inteligência artificial, e a ciência da informação por meio do refinamento de propostas das linguagens documentárias, notadamente com o aporte da terminologia. Há pouco diálogo entre os dois campos em relação aos produtos gerados, mais especificamente entre ontologias e tesouros; apesar das evidentes semelhanças e graus de complementaridade, ambos pouco se beneficiam dos avanços. Assumindo-se que a representação do conhecimento ou a representação documentária jamais são fins em si mesmas, o compromisso maior que tanto um campo quanto outro adota relaciona-se, no limite, com o desenvolvimento da ciência e da tecnologia, já que o fluxo do conhecimento científico e tecnológico também se dá por meio de suas representações, sejam ontologias ou tesouros.

Um levantamento quantitativo, sem maiores pretensões, sobre o número de publicações relacionadas a esses campos em alguns dos periódicos nacionais e internacionais mais importantes da ciência da informação revela a ausência de pesquisas sobre o tema no seu interior. Considerou-se o período 2004-2008 e buscou-se exclusivamente pelo termo *ontologia*, nos campos de palavras-chave e título.

Foram verificados os seguintes periódicos nacionais: *Ciência da Informação*, *DataGramaZero*, *Em Questão*, *Encontros Bibli*, *Informação & Informação*, *Informação & Sociedade: estudos*, *Perspectivas em Ciência da Informação*, *Revista Digital de Biblioteconomia e Ciência da Informação* e *Transinformação*. Foram encontrados apenas 12 artigos no período coberto (Tabela 1), concentrados em apenas cinco dos periódicos analisados. O interesse pela temática aponta tímido crescimento ao longo dos anos que compreendem o período coberto<sup>4</sup>.

Para o levantamento da literatura internacional, foi utilizada a base de dados da *Library and Information Science Abstracts (LISA)*, que cobre cerca de 440 periódicos de mais de 68 países em mais de 20 idiomas diferentes. No período coberto, que também

---

<sup>4</sup> Nos anos subsequentes a este levantamento verificou-se crescimento nas discussões sobre ontologias por pesquisadores do campo da ciência da informação tanto em eventos mais específicos deste campo, como nos *Encontros Nacionais de Pesquisa em Ciência da Informação*, quanto em eventos com abordagem mais interdisciplinar, como nos *Seminários de Pesquisa em Ontologia no Brasil*. A título de ilustração, elencam-se: Campos et al. (2009), Costa; Marcondes (2008); Freitas et al. (2008); Sales; Café (2008); Silva; Souza; Almeida (2008).

compreendeu os anos de 2004 a 2008, foram localizados 261 artigos, em setenta periódicos diferentes. Na Tabela 2 foram elencados apenas os periódicos que apresentaram maior número de trabalhos; são seis periódicos que respondem por 41,38% do total de artigos encontrados (N=261). A Tabela completa é apresentada no Apêndice A.

Tabela 1 - Artigos sobre ontologia em periódicos nacionais de ciência da informação

PERIÓDICO	2004	2005	2006	2007	2008	Total	
						N	%
Encontros Bibli	-	2	-	2	1	5	41,7
Ciência da Informação	-	-	-	-	2	2	16,7
Perspectivas em Ciência da Informação	-	-	-	1	2	3	25
DataGramaZero	1	-	-	-	-	1	8,3
Revista Digital de Biblioteconomia e Ciência da Informação	-	-	-	1	-	1	8,3
Em Questão	-	-	-	-	-	-	-
Informação & Informação	-	-	-	-	-	-	-
Informação & Sociedade: estudos	-	-	-	-	-	-	-
Transinformação	-	-	-	-	-	-	-
Total	N	1	2	-	4	5	12
	%	8,3	16,7	-	33,3	41,7	

Tabela 2 - Artigos sobre ontologia em periódicos internacionais de ciência da informação

PERIÓDICO	2004	2005	2006	2007	2008	Total	
						N	%
Journal of the China Society for Scientific and Technical Information	6	4	10	6	7	33	12,6
IEEE Intelligent Systems	11	3	6	-	1	21	8,0
Knowledge-based systems		2	8	2	4	16	6,1
Nauchno-tekhnicheskaya informatsiya	1	9	4	1	-	15	5,7
Journal of the American Society for Information Science and Technology	2	3	-	1	6	12	4,6
Information: Wissenschaft & Praxis	1	1	6	1	2	11	4,2
Total	N	21	22	34	11	20	
	%	8,0	8,4	13,0	4,2	7,7	

Os artigos não foram analisados; isto inviabiliza a verificação de sua pertinência no que se refere à discussão sobre ontologias sob a ótica da ciência da informação, mas é

importante observar que a LISA destina-se, como o informa o *site* do fornecedor<sup>5</sup>, aos profissionais de biblioteca e outros especialistas em informação. Os anos de maior produção, em números absolutos, 2005 e 2006 (N=66 em cada um deles. Em 2004 N=55; em 2007 N=32 e em 2008 N=42) coincidem com o silêncio da produção brasileira no mesmo período. Vale destacar ainda que apenas um periódico, o *Journal of the China Society for Scientific and Technical Information*, responde por mais de um décimo de toda a produção levantada no período coberto.

Curiosa e estranhamente, os cinco artigos encontrados nos periódicos nacionais, *Ciência da Informação* e *Perspectivas em Ciência da Informação*, os quais também são indexados pela LISA, não apareceram no levantamento realizado na base. Procedeu-se então à pesquisa específica, por meio da LISA, nestes periódicos, e apareceram como resultados apenas dois artigos da *Ciência da Informação* publicados em 2003, fora, portanto, do escopo do levantamento.

O levantamento, como se disse, foi pouco profundo e não considerou a análise dos artigos em si, mas se ateve às escolhas de palavras-chave conforme manifestadas no artigo. Dessa forma não cobre aqueles trabalhos que tratam da temática utilizando outras expressões como palavras-chave. Acredita-se que não houve prejuízos para a compreensão do fenômeno porque não há variação terminológica significativa entre os termos utilizados. Ainda relativamente pouco comum no campo da ciência da informação, o termo ontologia é mais frequente e mais antigo na ciência da computação, principalmente no subcampo denominado inteligência artificial.

---

<sup>5</sup> Disponível em: <<http://www.csa.com/factsheets/lisa-set-c.php>> . Acesso em: 12 ago. 2009.

## 1.2 HIPÓTESES

A maioria dos métodos automáticos de registro e localização da informação considera basicamente os seus aspectos sintáticos e estatísticos. Acredita-se, então, que a compreensão mais terminológica das ontologias pode oferecer colaborações para a organização de um acesso menos quantitativo e mais qualitativo à informação. Mais de dois mil anos depois do surgimento das primeiras formas de índice, ainda se dá primazia, *grosso modo*, para a frequência e/ou ocorrência como indicadores de relevância.

Sustentam esta reflexão os seguintes pressupostos:

- a) sem o estabelecimento de ontologias comuns não se poderão formar bases de conhecimento que garantam a qualidade das trocas informacionais necessárias entre os diferentes sistemas de informação ou sistemas documentários interconectados por redes de informação;
- b) a construção de ontologias não pode prescindir do tratamento terminológico-conceitual, como compreendido pela terminologia e pela ciência da informação, acumulado nos referenciais teóricos e nas metodologias para construção de linguagens documentárias;
- c) a construção de linguagens documentárias mais precisas e mais flexíveis expressas em relacionamentos conceituais mais dinâmicos pode beneficiar-se do conhecimento acumulado na construção de ontologias de domínio e suas regras de inferências e relacionamentos.

As hipóteses de trabalho são as seguintes:

- a) o conceito clássico de ontologias, conforme compreende a ciência da computação, é adequado à representação do conhecimento em campos estáveis, que possuem um objeto concreto delimitado;

- b) o uso de ontologias clássicas apresenta dificuldades quando aplicado a objetos teóricos, notadamente nas humanidades e nas ciências sociais;
- c) o uso de ontologias clássicas oferece dificuldades de aplicação em áreas nas quais não há estabilidade de conceitos, em áreas em formação ou transformação;
- d) ao menos teoricamente, a proposta deleuzeana do rizoma pode sustentar a representação de modo alternativo ao modelo arborescente de representação do conhecimento.

## 1.3 OBJETIVOS

### **Geral**

Investigar a interface teórico-prática entre terminologia, ontologia filosófica, ontologia computacional e linguística documentária e os subsídios que oferecem para a construção de informações documentárias em sistemas de informação documentária.

### **Específicos**

- a) analisar as condições de produção, desenvolvimento, implementação, uso e integração de ontologias com base no referencial teórico da ciência da informação, considerando domínios estáveis, bem como domínios em formação e transformação;
- b) investigar a contribuição das ontologias para o desenvolvimento de tesouros e vice-versa;
- c) discutir o fundamento filosófico da aplicação de ontologias com base no estudo das categorias ontológicas presentes na filosofia clássica (Aristóteles e Kant) e nas propostas contemporâneas (Deleuze);

- d) compreender as possíveis inter-relações entre ontologia filosófica, ontologia computacional e representação documentária baseada em categorias de assunto em domínios estáveis e em domínios em formação e transformação.

## 1.4 JUSTIFICATIVA

Paul Otlet, nome que certamente figura no *panteão* da ciência da informação (*latu sensu*), inicia seu hoje famoso *Tratado de documentação*, publicado em 1934, com a seguinte frase:

Para tornar acessível a quantidade de informações e artigos publicados a cada dia na imprensa cotidiana, nas revistas, para conservar os folhetos, informes, prospectos, documentos oficiais, para encontrar as matérias dispersas nos livros, para fazer um todo homogêneo destas massas incoerentes, são necessários novos procedimentos, muito distintos daqueles da antiga biblioteconomia, tal com têm sido aplicados (OTLET, 1996, p. 6).

A quantidade de informações em circulação atualmente, com todas as facilidades trazidas pelas tecnologias da informação e da comunicação, continua provocando a busca de novas metodologias para a organização destas massas que agora, em vista da maior oferta e diversidade de documentos, são ainda mais incoerentes. Não fossem outras razões, esta seria suficiente para motivar uma reflexão sobre este que tem sido um dos problemas mais caros à ciência da informação: as condições de produção, acesso e circulação da informação, bem como o comprometimento com a produção, acesso e circulação do conhecimento.

Fora do âmbito estrito dos sistemas de informação documentária, a eficácia e o sucesso do *Google™* e seu recurso de métodos automáticos para a recuperação da informação na internet – em detrimento dos primeiros diretórios, como o *Yahoo!* e sua indexação realizada por especialistas – ensinam sobre a necessidade de utilização de modelos automatizados também no escopo da representação e recuperação da informação

documentária. Ocorre, contudo, que a informação é um constructo e depende, nesta condição, dos elementos semânticos para sua efetivação. Neste aspecto falham os métodos automáticos dos mecanismos de busca, ainda baseados nos atributos sintáticos, o que resulta em perda da qualidade.

O tratamento semântico, pela relação contextual que visa a estabelecer entre signo e significado, é inerentemente complexo e tem sido proposto como tarefa para as ontologias computacionais, principalmente em aplicações para a internet. O uso de ontologias neste ambiente visa a sanar um problema histórico recente: a rede cresceu como meio para a troca de informações entre *pessoas* e deixou a desejar no que se refere às condições da troca de dados e informações que possam ser processados *automaticamente* (BREITMAN, 2005).

As ontologias vêm sendo estudadas com maior interesse pela ciência da computação desde a década de oitenta, com vistas à padronização da representação do conhecimento por diferentes sistemas de informação. Na ciência da informação o primeiro estudo sobre o assunto data de 1977, mas o interesse maior se deu apenas a partir de 2002 (NASCIMENTO et al., 2007), o que caracteriza de algum modo uma preocupação tardia, posto que a representação do conhecimento também é objeto da ciência da informação, como sempre foi, a seu modo, da biblioteconomia.

Discutir as interconexões entre as ontologias e as linguagens documentárias, mais especificamente os tesauros, justifica-se ainda pelo comprometimento direto deste tema com a qualidade da comunicação científica, que tem a precisão como condição para sua eficácia. O uso de instrumentos de representação do conhecimento adequados visa a reduzir as barreiras e ruídos na comunicação entre os produtores de ciência entre si e entre estes com o público especializado ou leigo.

Ainda que trate de um tema cuja literatura mais vasta encontra-se no interior da ciência da computação, esta pesquisa tem como fundamentação teórica a literatura da ciência da informação e da terminologia, com especial atenção para a teoria comunicativa da

terminologia, cuja concepção de unidade terminológica é mais flexível e admite a possibilidade de variação conceitual e denominativa, além de admitir também relações entre conceitos que ultrapassam as relações lógicas e ontológicas estabelecidas pela teoria clássica.

## **1.5 ESTRUTURA DA TESE**

O texto foi estruturado da forma que segue. Após este capítulo introdutório, é apresentado o capítulo dois, que trata das bases filosóficas das ontologias, compreendendo a ontologia em seu sentido filosófico como base para a compreensão das relações que guardam a ontologia (filosofia) e as ontologias (ciência da computação e ciência da informação). Neste capítulo busca-se compreender também algumas das bases filosóficas das categorias. O capítulo três apresenta algumas questões gerais a respeito das ontologias, conforme a compreensão do termo pela ciência da computação, principalmente, e pela ciência da informação, e retoma as relações que estabelecem entre si ontologias e categorias. No capítulo quatro são apresentadas as teorias clássica e comunicativa da terminologia, considerando-se suas semelhanças e diferenças e sua aplicabilidade na construção de ontologias. O capítulo cinco trata da compreensão da linguística documentária e de suas relações com a terminologia. A aplicação das metodologias propostas para a construção de ontologias na construção de informações documentárias é discutida no capítulo seis. O capítulo sete apresenta a proposta deleuzena do rizoma como alternativa teórica para as representações arborescentes do conhecimento. Ao final, no capítulo oito, são apresentadas as considerações finais.

## 2 A BASE FILOSÓFICA DAS ONTOLOGIAS

Neste capítulo serão apresentados os conceitos gerais de ontologia (considerando-se o entendimento da filosofia) em busca da compreensão de alguns dos aspectos que a aproximam e de outros que a distinguem das ontologias (conforme compreendidas pela ciência da computação e pela ciência da informação). Exploram-se neste capítulo também as relações que a ciência da informação estabelece com os estudos sobre categorias, sobre ontologias e representação documentária.

Considerada em quaisquer de seus aspectos, pode-se atribuir à ontologia a função de fornecer sistemas de categorização que permitam ao homem organizar a realidade. Insere-se aí sua proximidade com o domínio da ciência da informação e objetiva-se, desse modo, com o presente capítulo: a) discutir o fundamento filosófico da aplicação de ontologias com base no estudo das categorias ontológicas presentes em alguns filósofos, notadamente Aristóteles, Kant e Peirce e b) fornecer subsídios iniciais para a compreensão da inter-relação entre ontologia filosófica, ontologia formal e representação documentária baseada em categorias de assunto.

A adoção de princípios ontológicos no campo de pesquisa que envolve a representação e a recuperação da informação e do conhecimento, ainda que a base tecnológica atual possa fazer parecer incipiente, não é recente, embora ainda seja, por sua complexidade, insipiente. Aristóteles, *um sábio de Estagira cuja cabeça sustenta ainda hoje o Ocidente*<sup>6</sup>, inaugura, com o que chamou de metafísica<sup>7</sup>, estes estudos.

---

<sup>6</sup> Verso da música *Alexandre*, de Caetano Veloso, gravada em seu disco *Livro*, de 1997.

A biblioteconomia, a documentação e a ciência da informação também contribuíram, certamente, para que o nome de Aristóteles se repetisse *desde esses tempos até os nossos tempos e além*<sup>8</sup>. Uma das diferenças mais significativas entre os tesouros e as taxonomias, por exemplo, é a presença de relações ontológicas no primeiro, o que o torna, aliás, um instrumento mais complexo.

A lógica, que também tem sido largamente utilizada, é um importante instrumento para atestar que alguma coisa existe, mas como instrumento de argumentação, não pode – nem pretende – descrever as coisas que existem. Pode-se afirmar com o exemplo clássico, que Sócrates é mortal, e isso é perfeitamente representável num tesouro, mas não se pode discutir, usando-se exclusivamente um tesouro, o que é e como se relaciona, em termos ontológicos, o conceito de morte com os outros conceitos que fundam a noção de realidade; em outras palavras, não se pode, seguindo a linha da lógica aristotélica, compreender a essência do conceito.

A ontologia desponta como preocupação filosófica a partir da metafísica aristotélica. A célebre *Escola de Atenas* (Figura 1), de Rafael (1483-1520), pintor renascentista italiano, retratando a continuidade do pensamento da Academia de Platão, representa o momento desta mudança. Os dois grandes filósofos do pensamento clássico aparecem no centro da figura. Enquanto Platão aponta para o alto, para o mundo das ideias, Aristóteles diverge e aponta para o que o rodeia, indicando sua preocupação com o mundo concreto. Posteriormente, Porfírio (*explicador* de Aristóteles), Brentano, Kant e Peirce entre outros, retomam e atualizam o pensamento aristotélico, direta ou indiretamente, concordando com ou divergindo dele, conforme apontado na discussão que se segue.

A ontologia, numa definição sumária, preocupa-se essencialmente com o ser, com a identificação das características comuns a todos os seres. Esta identificação se dá por meio

---

<sup>7</sup> Aristóteles qualificou os pré-socráticos, por seu esforço em compreender o universo (*physis*) de físicos (*physikoi*) e propôs, para diferenciá-los deles, o termo metafísica.

<sup>8</sup> Verso da música *Alexandre*, de Caetano Veloso, gravada em seu disco *Livro*, de 1997.

da observação, possibilitando o conhecimento do mundo físico, e do raciocínio, produzindo uma estrutura de abstrações.



Figura 1 - A escola de Atenas

Como esta reflexão se realiza no interior de uma ciência social aplicada, é importante enfatizar que a descrição e a representação possibilitadas pela ontologia interessam também aos estudos sobre criação e manutenção de bases de dados, já que a escolha das categorias ontológicas é (ou deveria ser) o primeiro passo na sua definição, conforme Sowa (2000). O campo que compreende a geração e uso de base de dados, contudo, utiliza as ontologias de forma distinta da que a filosofia faz. Enquanto esta última elabora grandes concepções independentes de aplicação (*top-down*), a primeira trata de micromundos, com um número de conceitos limitados por suas aplicações (*bottom-up*). Em vez de discutir o antagonismo que a distinção dicotômica pode fazer transparecer é necessário considerar suas aplicações diferenciadas. Quando se utiliza a estrutura *bottom-up* com o fim de resolver (ou tentar resolver) questões localizadas de representação e uso do conhecimento, a ciência da informação o faz de forma dialética com os empréstimos tomados da filosofia. É preciso observar ainda que quando se trata de compartilhamento de bases de conhecimento o entendimento filosófico da ontologia torna-se se não mais relevante, pelo menos merecedor de um olhar atento, porque mais generalizante.

Neste ponto tem-se o impasse que caracteriza a complexidade destes estudos. Os recortes da realidade trazem como consequente agravante os problemas relativos às garantias de interoperabilidade das ontologias entre si. Por outro lado, admitindo-se que a concepção de ontologias de nível superior, que descrevem conceitos gerais independentes de domínio e neste sentido mais próximas da concepção *top-down* da ontologia filosófica (SOWA, 2000), seja mais adequada para o compartilhamento de conhecimentos (porque buscam modelos *universais*), é preciso considerar que não há na intenção *inicial* da ontologia filosófica uma proposta de formalização, do modo como a compreende a ciência da computação, o que torna esta concepção praticamente não-operacionalizável no universo dos sistemas de informação.

Dentre as preocupações que remontam às mais antigas das aspirações filosóficas da humanidade, estão a compreensão do melhor método para o correto pensamento e a compreensão da realidade em si mesma, de como o mundo, por sua ordem aparente, se separa do não-mundo, do *i-mundo*. À primeira busca denominou-se lógica, à segunda ontologia. Enquanto a lógica preocupa-se com a relação mútua entre proposições, a ontologia preocupa-se com a mútua relação entre coisas. Insiste-se novamente na distinção entre lógica e ontologia: a lógica pode afirmar que alguma coisa existe, mas em si não possui vocabulário para descrever as coisas que existem (SOWA, 2000). A lógica permite verificar a veracidade ou falsidade das relações mútuas entre proposições por meio da análise de argumentos ou conjunto de enunciados. Considerando-se com Porfírio de Tiro (2002, p. 31) que “a linguagem é fundamentalmente um instrumento de comunicação” concernente ao mundo da experiência, a lógica “estuda o modo pelo qual funciona esta comunicação”.

Um dos modos pelos quais se pode compreender o seu funcionamento ocorre por meio da lógica de enunciados (segmentos linguísticos que possuem sentido completo e podem ser verdadeiros ou falsos). Há duas condições que um segmento linguístico deve cumprir para que seja um enunciado: a) deve ter um sentido completo e b) deve poder ser verdadeiro ou falso. Assim, a frase ‘A água está fria’ é um enunciado, pois preenche ambas

as condições, já ‘Traga-me um copo d’água’, não, pois possui sentido completo, mas não se pode afirmar sobre sua veracidade ou falsidade.

Se o enunciado é verdadeiro ou falso, contudo, não é a lógica que determina, são as diversas ciências particulares ou a observação empírica sobre o mundo. Valer-se-á aqui do exemplo citado por Nahra e Weber (2001, p. 85): seja a conjunção “A terra é redonda e a URSS é um país”. Primeiro enunciado simples conjuntivo: “A terra é redonda” (representação: p). Segundo enunciado conjuntivo: “A URSS é um país” (representação: q).

	p	q	p.q
1	V	V	V
2	V	F	F
3	F	V	F
4	F	F	V

Quadro 1 - Enunciados lógicos

Em 1960, a realidade correspondia à situação 1, hoje corresponde à situação 2. No ano de 1960, quem afirmasse ‘p.q’ estaria afirmando uma verdade; hoje quem afirmar ‘p.q’ afirmará uma falsidade, como se pode ver visualizar no Quadro 1.

Retornando da pequena digressão acima, é preciso salientar que neste trabalho não se ocupa da lógica, mas da ontologia e da sua interface com a ciência da informação no que tange à categorização do conhecimento como fundamento de sua representação e da recuperação da informação em redes ou sistemas de informação documentária. As relações entre lógica e ontologia interessam, contudo, aos estudos sobre ontologias quando a lógica é utilizada como instrumento de formalização das ontologias.

As ontologias vêm ocupando cada vez mais a atenção de estudiosos da ciência da computação e, mais recentemente, da ciência da informação, com a possibilidade de melhorar significativamente a representação computacional de um domínio do conhecimento. Esta melhoria dá-se inicialmente num nível máquina-máquina e depois máquina-homem, para posteriormente interferir na relação homem-homem. Sendo as ontologias conjuntos estruturados de conceitos especializados representados por

denominações, a terminologia pode igualmente contribuir para o aprofundamento da compreensão.

Defende-se, juntamente com Smith (2004), que se corre o risco de realizar um uso *pobre* da ontologia se isso ocorrer desvinculado da compreensão filosófica. A maioria dos trabalhos sobre metodologias para desenvolvimento de ontologias para sistemas de informação, como foi possível comprovar tomando-se como ponto de partida a sustentação de Smith (2004), ignoram o trabalho feito na filosofia.

Assim como as classificações bibliográficas são tributárias das classificações filosóficas, acredita-se que não se pode desenvolver ontologias (conforme compreendidas pela ciência da computação/inteligência artificial) sem o necessário aporte da ontologia (conforme compreensão da filosofia). Por esta razão, apresentam-se neste trabalho as bases filosóficas da ontologias, por meio dos estudos de categorização. Como se poderá perceber, não foi objetivo deste estudo aprofundar a compreensão dos pensadores, mas compreender brevemente algumas das contribuições que prestaram para os estudos sobre as categorias.

A ontologia (teoria do ser) forma, juntamente com a estética (teoria do belo), a ética (teoria da moral), a epistemologia (teoria do conhecimento) e a lógica, o conjunto das divisões didáticas da filosofia. A palavra ontologia é resultado da junção de duas outras: *onto* e *logia*. Chauí (2000, p. 266) ensina sobre a etimologia da palavra *onto*:

*Onto* deriva-se de dois substantivos gregos, *ta onta* (os bens e as coisas realmente possuídas por alguém) e *ta eonta* (as coisas realmente existentes). Essas duas palavras, por sua vez, derivam-se do verbo ser, que, em grego, se diz *einai*. O particípio presente desse verbo se diz *on* (sendo, ente) e *ontos* (sendo, entes). Dessa maneira, as palavras *onta* e *eonta* (as coisas) e *on* (ente) levaram a um substantivo: *to on*, que significa o Ser.

Trata-se nesta acepção, naturalmente, do ser que é realmente e que se opõe ao que parece ser, à aparência. Dessa forma a ontologia é definida como o estudo ou conhecimento do ser, dos entes ou das coisas tais como são em si mesmas, real e verdadeiramente. A ontologia é, deste modo, a própria filosofia e o conhecimento do ser, pois representa a

passagem das opiniões sobre as coisas sensíveis mutáveis para o pensamento sobre as essências imutáveis.

## 2.1 CATEGORIAS ARISTOTÉLICAS

O conceito de categoria é escorregadio. Ferrater Mora (1978) sustenta que foi Aristóteles quem primeiro usou o termo categoria em sentido técnico. Às vezes, tergiversa o dicionarista, pode ser traduzida por ‘denominação’ e, com maior frequência, por ‘predicação’ e ‘atribuição’. Nesta pesquisa optou-se pelo uso da expressão categoria, considerando-se, inclusive, sua popularização.

A querela ontológica inicia-se quando Platão (428-348 a.C.) propõe sua hoje célebre teoria das ideias, fundamentada no fato de que os sentidos revelam as coisas como múltiplas e mutáveis e que somente a inteligência pode revelar sua unidade e permanência. Se Platão separa o mundo em dois – um mundo onde tudo é ilusão e um mundo das ideias, onde tudo é verdade – Aristóteles une-o novamente: verdade e aparência estão no mesmo mundo, a questão é diferenciá-las. A lógica e a ontologia, por meio da linguagem, formam, por excelência, os instrumentos necessários para a compreensão e formalização das diferenças.

Aristóteles (384-322 a.C.) compreende o mundo físico como a realidade última. Isso está sustentado no estabelecimento das categorias com as quais quer organizar a realidade: substância, qualidade, quantidade, relação, lugar, tempo, posição, estado, ação e paixão. O Quadro 2 apresenta um exemplo que pode facilitar a compreensão de sua aplicabilidade. A substância pode ser considerada como a categoria primeira, pois em torno dela giram as afirmações. As demais categorias possibilitam os vários tipos de afirmação que se pode fazer a respeito da substância.

Categoria	Exemplo
Substância	mulher
Qualidade	administradora
Quantidade	1,65 m
Relação	mãe de Larissa e Raíssa
Lugar	Cruzeiro
Tempo	2010
Posição	sentada
Estado	bem vestida
Ação	atenta
Paixão	recebendo informações

Quadro 2 - Categorias aristotélicas

Sowa (2000) estabelece um interessante paralelo entre as categorias de Aristóteles e seus desdobramentos diretos ou indiretos. Neste trabalho, que acompanha seu raciocínio em alguns aspectos, ater-se-á somente à já clássica representação de Porfírio e às representações posteriores de Brentano, Kant e Peirce por considerá-las mais próximas do que se pretende como organização e representação do conhecimento no universo da ciência da informação em sua necessidade de operar com esses conceitos. Estudos posteriores deverão comparar a evolução da representação ontológica em outros pensadores.

## 2.2 A ÁRVORE DE PORFÍRIO

A dicotomia proposta por Porfírio (232-305) deve sua larga e longa aceitabilidade, certamente, ao que se pode chamar de vontade humana de que o mundo pudesse assim ser compreendido e representado, notadamente no medievo. Feliz ou infelizmente a realidade é mais complexa do que isto e não pode ser representada em face da presença/ausência de determinados traços. Dessa forma o tipo de categorização proposto nesta representação perde facilmente sustentação.

Porfírio foi discípulo do filósofo Cássio Longino em Atenas, o qual verteu-lhe o nome sírio *Malco* para o correspondente grego *Porfírio* (SANTOS, 2002). Em 263 esteve em Roma

onde a convivência com Plotino amadureceu seu pensamento. Dentre suas contribuições mais importantes estão os comentários sobre as categorias de Aristóteles: a *Isagoge* (uma variação linguística arcaica equivalente a preliminar ou preâmbulo). Segundo Santos (2002), seu tradutor e comentarista, a *Isagoge* pode interessar ao leitor de vários modos:

- a) como introdução à lógica e, como tal, à filosofia toda, ou
- b) como síntese das soluções à questão da relação entre física, lógica e gramática, ou
- c) como transmissor do legado da Grécia antiga às escolas do Ocidente medieval.

Nesta obra, Porfírio formulou a questão que ficaria conhecida como a *querela dos universais*<sup>9</sup>, um dos principais problemas lógicos e metafísicos da Idade Média no Ocidente. Ainda na *Isagoge*, Porfírio construiu a estrutura lógica hierárquica que ficou conhecida como sua árvore, embora ele mesmo não tenha proposto explicitamente esta ou qualquer outra forma de representação gráfica. Citando-o literalmente:

Esclareçamos o que queremos dizer tomando o exemplo de uma única categoria. A essência é em si mesma um gênero: abaixo dela acha-se o corpo; sob o corpo, o corpo animado; sob o corpo animado, o animal; sob o animal, o animal racional; sob o animal racional o homem; sob o homem acham-se Sócrates, Platão e os homens particulares (PORFÍRIO DE TIRO, 2002, p. 40).

Baseando-se na teoria dos predicados de Aristóteles, a *Árvore de Porfírio* constitui-se num conjunto hierárquico finito de gêneros e espécies, identificados por dicotomias sucessivas, como se pode ver na Figura 2.

A influência de Porfírio estende-se pela posterior teoria da definição e classificação dos seres e igualmente pelas classificações das ciências, seja originando classificações

---

<sup>9</sup> A querela dos universais origina-se de três questões colocadas por Porfírio no início da *Isagoge*: “Antes de mais, no que tange aos gêneros e às espécies, acerca da questão de saber (1) se são realidades subsistentes em si mesmas ou se consistem apenas em simples conceitos mentais (2) ou, admitindo que sejam realidades subsistentes, se são corpóreas ou incorpóreas e, (3) neste último caso, se são separadas ou se existem nas coisas sensíveis e delas dependem [...]” (PORFÍRIO DE TIRO, 2002, P. 35).

dicotômicas, seja exigindo-lhes simetria (POMBO, 1998). Essa influência estende-se ainda, naturalmente, com maior ou menor intensidade pela classificação bibliográfica e outras linguagens documentárias de um modo geral e também está presente nas taxonomias e nas ontologias.

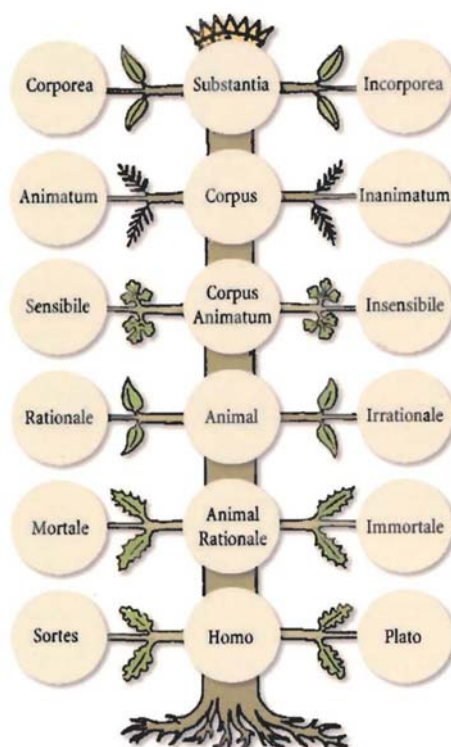


Figura 2 - Árvore de Porfírio<sup>10</sup>

A linha moderna de compreensão das classificações bibliográficas, contudo, é devedora de Ranganathan, que ratifica a dinamicidade da classificação a partir da adoção de facetas que permitem a adoção de policotomias. Anteriormente a Ranganathan, informa Pombo (1998), o problema já fora detectado por outros estudiosos, como, por exemplo, Buffon, em 1749, quando discutia a impossibilidade da adoção de um sistema geral em face

<sup>10</sup> Fonte: PORFÍRIO DE TIRO. *Isagoge*: introdução às categorias de Aristóteles. São Paulo: Attar, 2002. [quarta capa]

da dinamicidade (ou ainda da multiplicidade como denominaria Deleuze, séculos mais tarde) inerente à natureza:

[...] a natureza caminha por gradações desconhecidas e, por consequência, ela não se pode prestar totalmente a essas divisões uma vez que ela passa de uma espécie a outra, e muitas vezes de um gênero a outro, por nuances imperceptíveis, de tal forma que se encontra um grande número de espécies intermédias e de meios-objetos que não sabemos onde colocar e que perturbam necessariamente o projeto de um sistema geral (BUFFON<sup>11</sup> *apud* POMBO, 1998).

A *Árvore de Porfírio* deve sua limitação à relatividade de sua construção, conforme apontado por Tálamo *et al.* (1992), pois depende da seleção do atributo de qualidade que compreenderá as subdivisões sucessivas. O homem, no exemplo de Eco (1984), também citado pelas autoras, só é mortal numa hierarquia que focalize o problema da duração da vida. Focalizando-se outros problemas, pode-se, por exemplo, inseri-lo numa árvore que considere sua natureza orgânica ou inorgânica. A seleção se dá, então, pelo que Tálamo *et al.* (1992) chamam de '*pressão*' *contextual*, podendo-se considerá-la verdadeira "apenas em relação a um determinado código e não em relação às propriedades dos objetos em si mesmos". Neste tipo de divisão, o contexto não é incorporado, ou então incorpora-se apenas um contexto determinado, o qual irá suportar a interpretação.

## 2.3 A LEITURA DE BRENTANO

Franz Brentano (1838-1917), criador do intencionalismo, defendeu a existência real do objeto de conhecimento independentemente do sujeito pensante, e foi um precursor da moderna fenomenologia. Reorganizou as dez categorias aristotélicas como folhas de uma mesma árvore, dando-lhes um arranjo mais complexo (SOWA, 2000), cujos espaços foram preenchidos com outros termos tomados de Aristóteles (Figura 3). A tradução dos termos

---

<sup>11</sup> BUFFON, G. L. L. Histoire naturelle, générale et particulière. In: \_\_\_\_\_. *Oeuvres*: vol. 1. Paris: Hachette, 1845.

apresentados nesta figura é de Moreira, Alvarenga e Oliveira (2004), também leitoras de Sowa (2000), o qual chama atenção para a já conhecida dificuldade de tradução dos termos gregos para línguas baseadas no latim.



Figura 3 - Árvore de Brentano

Brentano acrescentou à concepção porfiriana das categorias aristotélicas algumas modificações significativas. A substância deixa de ser a categoria definidora máxima e passa a ser uma instância do ser, que comporta também o acidente; este se subdivide em propriedade e relação. A propriedade cuja divisão primeira é inerência, direcionamento e contenção, forma a maior ramificação, em cuja base estão as outras categorias aristotélicas (qualidade, quantidade, ação, paixão, estado, posição, lugar e tempo). Com este arranjo mais sistematizado Brentano presta um serviço fundamental para as futuras formas de representação do conhecimento, principalmente para a construção de ontologias de topo, propostas por Sowa (2000).

## 2.4 AS CATEGORIAS DE KANT

A razão, para Immanuel Kant (1724-1804), é uma estrutura vazia, inata, independente da experiência. A razão seria assim uma forma pura, sem conteúdos,

universal. Os conteúdos são dados pela experiência. “Sem sensibilidade”, pode-se ler em Kant (1996, p. 92), no que é considerado como uma síntese entre racionalistas e empiristas, “nenhum objeto nos seria dado, e sem entendimento nenhum seria pensado. Pensamentos sem conteúdo são vazios, intuições sem certos conceitos são cegas [...]”.

Na *Crítica da razão pura*, Kant propôs uma primeira mudança radical nas categorias aristotélicas: a combinação dos relacionamentos num julgamento, considerando que a realidade se dá na possibilidade de emissão de juízos, afirmações ou negações. Em outros termos, não basta apenas que algo esteja num contexto de tempo e de espaço para que seja considerado real, antes é preciso que se possam emitir juízos a seu respeito. Apresentando-se a realidade nos juízos, pode-se afirmar a correspondência condicionada entre juízo e realidade, o que permite ao homem formar um conjunto de juízos que sustenta sua expressão das coisas reais.

Um juízo pode ser definido como “a maneira como o conhecimento afirma ou nega o que uma coisa é ou não é [...] a afirmação ou a negação da realidade de um objeto pela afirmação ou negação de suas propriedades” (CHAUÍ, 2000, p. 295).

Os juízos não possuem, contudo, sempre a mesma natureza. A afirmação ou negação das propriedades de algum objeto, variam e condicionam a percepção do objeto, como ocorre, por exemplo, com as seguintes afirmativas: ‘A USP é uma universidade brasileira’, que não possui a mesma natureza de: ‘A USP é a melhor universidade do Brasil’. Ambas referem-se a propriedades da USP; são portanto juízos e podem dar a impressão de simplesmente descrevê-la. Basta um exame mais atento, entretanto, para que sejam percebidas suas diferenças: a primeira afirmativa refere-se ao que se denomina por fatos, enquanto a segunda envolve um ‘juízo de valor’, representando, neste caso, uma opinião, para a qual não se pode esperar o mesmo grau de concordância.

Enquadrando-se o juízo na teoria dos atos linguísticos proposta por Austin, tem-se que ele é uma forma de declaração. Uma declaração, diferentemente de uma afirmação, cria

novas realidades linguísticas. Segue um exemplo, tomado de Echeverría (2003, p. 66-67, tradução nossa):

[...] Quando se anuncia que a Miss Venezuela é a mais bela das concorrentes e se lhe confere o título de Miss Universo; em que momento se tornou a mais bela? Naquele em que foi pronunciada como tal? Ou já o era, antes que este julgamento fosse emitido? Devemos reconhecer que, do ponto de vista das propriedades físicas das concorrentes, não há grandes mudanças nelas antes e depois do pronunciamento. No entanto, antes da declaração, a Miss Venezuela, sendo o que era, não era a mais bela das participantes. Só chega a sê-lo quando alguém emite um juízo de que é. Com base neste juízo o mundo muda; muda para aqueles que aceitam o juízo emitido e muda, obviamente, para Miss Venezuela. A partir de então a sua identidade, tanto para si mesma como para os demais, é outra. [...]

A partir de uma *tábua dos juízos*, Kant propõe e lhe faz derivar e corresponder uma *tábua das categorias*, ambas organizadas em quatro grupos de três (Quadro 3). Esta é considerada a primeira grande mudança imposta às categorias aristotélicas, mas ainda parte delas (SOWA, 2000). É o próprio Kant (1996, p. 108), aliás, quem informa sobre seu desejo de rever Aristóteles: “segundo Aristóteles, denominaremos tais conceitos categorias na medida em que nossa intenção, em princípio, identifica-se com a de Aristóteles, se bem que se afaste bastante dele na execução”. Considera que por meio desses conceitos puros é possível compreender “algo do múltiplo da intuição” e segue sua crítica ao modelo aristotélico:

[...] a procura desses conceitos fundamentais constituiu um plano digno de homem perspicaz como Aristóteles. Entretanto, por não possuir nenhum princípio catou-os como se lhe deparavam, reunindo primeiramente dez, que denominou categorias (predicamentos). A seguir, creu ter encontrado ainda mais cinco conceitos que acrescentou sob a denominação de pós-predicamentos [...] (KANT, 1996, p. 109).

Aristóteles, na visão de Kant, enumera, mas não justifica as categorias, não apresenta provas de sua derivação. A diferença em Kant é o foco no sujeito como princípio orientador da descoberta. O mundo é ordenado pelos seres humanos segundo a distinção que possuem. Se não se conhece a distinção mesa, por exemplo, não se pode observá-la. Pode-se ver diferenças de cor, forma, textura etc., mas não uma mesa (SALATIEL, 2006). Os esquimós podem observar mais distinções de branco do que os brasileiros, a diferença entre os povos, neste caso, não é biológica.

	JUÍZOS	CATEGORIAS
1. QUANTIDADE	Universais	Unidade
	Particulares	Pluralidade
	Singulares	Totalidade
2. QUALIDADE	Afirmativos	Realidade
	Negativos	Negação
	Infinitos	Limitação
3. RELAÇÃO	Categóricos	Substância
	Hipotéticos	Causalidade
	Disjuntivos	Comunidade
4. MODALIDADE	Problemáticos	Possibilidade
	Assertórios	Existência
	Apodícticos	Necessidade

Quadro 3 - Tábua dos juízos e categorias de Kant<sup>12</sup>

Em Kant, são as categorias que permitem organizar o conhecimento, por esta razão se diz que o entendimento kantiano é apriorístico (anterior à experiência e independente dela). Munido das categorias, o sujeito forma os conceitos. As categorias permitem ao sujeito o "enquadramento" do mundo, por isso condicionam, num certo sentido, sua visão de mundo e sua racionalidade.

As categorias, em Kant, formam a condição para o conhecimento da realidade e sustentam o modo como a razão organiza os dados da experiência. Com Kant a metafísica clássica, cujo representante máximo é Aristóteles, é abalada, pois surge a compreensão de que o conhecimento é relação e não contemplação. Na concepção aristotélica (e também na platônica) conhecer equivale a classificar as coisas do mundo segundo regras formais do pensamento obedecendo a hierarquias.

<sup>12</sup> Adaptada de Salatier (2006, p. 82).

Ainda há ecos do pensamento kantiano em pensadores da ciência da informação, como, por exemplo, em Barreto (2002), de quem se transcreve o argumento-base para a crítica que realiza sobre o conceito de ‘gestão do conhecimento’

[...] Entende-se o conhecimento como um fluxo de acontecimentos, isto é, uma sucessão de eventos, que se realizam fora do estoque, na mente de algum ser pensante e em determinado espaço social. É um caminho subjetivo e diferenciado para cada indivíduo. [...]

Chamou-se ‘revolução copernicana’ à síntese que Kant realizou do embate entre racionalistas e empiristas. Sua concepção filosófica aceita argumentos de uns e de outros na compreensão de que o conhecimento começa com a experiência, mas que esta não age sobre um ser sujeito; o sujeito constrói esquemas interpretativos de organização dos dados da experiência. Estes esquemas permitem transformar em conhecimento os dados da experiência.

Em vez de explicar os conceitos em função da experiência, Kant dedicou-se a explicar a experiência em função dos conceitos, considerando que “antes que as impressões que irrompem nos sentidos se transformem em conhecimento, devem ser organizadas ou unificadas de algum modo pela atividade do entendimento” (RUSSELL, 2001, p. 345).

## 2.5 A TRÍADE DE PEIRCE

O princípio das categorias é retomado por Charles Sanders Peirce (1839-1914), que se declara, aliás, “fã ardoroso de Kant” durante sua juventude (SOWA, 2000), mas não aceita seu apriorismo.

A concepção teórica triádica que já havia encantado pensadores como Kant, conforme apontado acima, e também Freud (inconsciente, subconsciente e consciente ou id, ego e superego), Lacan (real, simbólico, imaginário) e que aparece em diversas outras concepções tais como a tríade figura divina nas religiões cristãs: Pai, Filho e Espírito Santo, além de estar presente em diversas outras concepções, também seduziu Peirce. Tomando como base o pensamento triádico de Kant, Peirce propõe outra tríade: primeiridade, secundidade e terceiridade (Figura 4), anteriormente denominadas qualidade, reação e mediação.

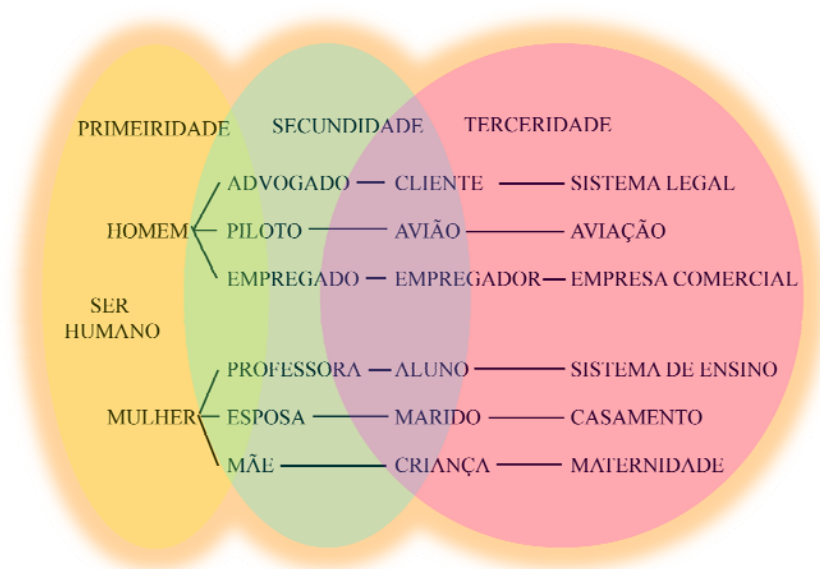


Figura 4 – Exemplo de aplicação da tríade peirceana

A *primeiridade* refere-se à concepção de ser ou de existir independente de qualquer outra coisa. Um indivíduo pode ser reconhecido como um ser humano ou um subtipo, tal como homem ou uma mulher, por impressão dos sentidos, independentemente de relacionamentos externos. A *secundidade* refere-se à concepção de ‘ser relativo a’, de reação a alguma outra coisa. O mesmo indivíduo pode ser classificado em relação a muitos outros tipos: mãe, advogado, esposa, piloto, empregado, pedestre etc.. Esta classificação depende de relacionamento externo com algum ser, tais como criança, cliente, marido, avião, empregador, tráfego. A *terceiridade* relaciona-se com a concepção de mediação por

meio da qual a primeiridade e a secundidade se relacionam, tais como a maternidade, o sistema legal, o casamento, a aviação etc. (SOWA, 2000).

Peirce, informa Braga (1999), levou trinta anos (1867-1897) para completar sua teoria das categorias. O cuidado excessivo reflete-se na complexidade e no nível de completeza de seus estudos. Sowa (2000) preocupa-se apenas com o ponto de vista ontológico da qualidade nas páginas que dedica ao estudo de Peirce. Isto é, evidentemente, natural, pois sua obra busca os fundamentos lógicos, filosóficos e computacionais da representação do conhecimento; sua preocupação é essencialmente ontológica. Braga (1999), entretanto, completa o quadro e permite compreender melhor a abordagem peirceana da ontologia (Quadro 4).

	Qualidade	Objeto	Mente
Do ponto de vista ontológico da qualidade	Qualidade ou primeiridade	Reação ou secundidade	Mediação ou terceiridade
Do ponto de vista do objeto ou secundidade	Qualia (fatos de primeiridade)	Relações (fatos de secundidade)	Representação (signos ou fatos de terceiridade)
Do ponto de vista da mente ou terceiridade	Sentimento ou consciência imediata (signos de primeiridade)	Sensação ou fato (sentido de ação ou reação ou signos de secundidade)	Concepção ou mente nela mesma (sentido de aprendizagem, mediação ou signos de terceiridade)

Quadro 4 - Categorias universais de Peirce

Em sua *nova lista de categorias* Peirce (1868) mantém alguns princípios das categorias lógicas aristotélicas, como o da cópula, por exemplo. A unidade de uma proposição, sustenta, é a unidade à qual o entendimento reduz as impressões. Esta unidade está na ligação do sujeito ao predicado. A cópula, portanto, é o que completa o trabalho dos conceitos de reduzir a multiplicidade à unidade. Assim, há expressões sem ligação, como “homem” ou “é vencedor” e expressões com ligação como “o homem é vencedor”. Expressões sem ligação, por si mesmas não afirmam nem negam nada, por isso não são analisáveis e podem ser agrupadas em categorias.

Peirce, por considerar sua obra sobre signos inseparável de seu trabalho sobre lógica (LECHTE, 2002) interessa de perto às questões referentes à construção e ao fluxo da informação documentária. Os signos, assim como a articulação das formas lógicas, são veículos do pensamento e elementos para a construção de conhecimentos.

## **2.6 CATEGORIAS, ONTOLOGIAS E INFORMAÇÃO DOCUMENTÁRIA**

Com o fim de organizar o texto que conclui este capítulo lança-se a pergunta: como as categorias se relacionam com as ontologias e estas com a informação documentária? Os sistemas de informação, como se sabe, têm sua eficácia condicionada a alguns princípios ordenadores basilares. A adoção de critérios explícitos e os recortes institucionais no estabelecimento de categorias, por exemplo, ajudam a evitar a tentação de tomar casos concretos por princípios fortemente estabelecidos. Nas linguagens documentárias as categorias herdam alguns princípios filosóficos, pois se referem a agrupamentos que manifestam particularidades em relação aos campos do conhecimento ou de atividade, mas como contraexemplo do que ali se dá, não são universais, como sugere Aristóteles, pois servem a propósitos institucionais ou pontos de vista.

A institucionalização, a necessidade de um olhar apriorístico que atenda às necessidades institucionais de categorização do conhecimento para fins de representação e posterior recuperação da informação em sistemas de informação pode fazer parecer que a abordagem kantiana é a solução para tais problemas. Com base em que princípios, contudo, poderão ser produzidas as categorias do conhecimento (categorias do entendimento, diria Kant)? O único princípio norteador da escolha é a própria experiência concreta dos sistemas de informação e sua complexidade, que envolve os usuários da informação (considerados de forma coletiva), as linguagens documentárias, os profissionais da informação e a base

tecnológica disponível. Neste sentido o entendimento de Peirce parece mais justo aos problemas ali verificados.

A ciência da informação, uma ciência social aplicada – ou socialmente aplicável, conforme sugere a versão preliminar da nova tabela das áreas do conhecimento do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), proposta para discussão<sup>13</sup> – vale lembrar, pretende organizar o conhecimento não apenas para representá-lo, mas também para garantir seu acesso e apropriação. Sem a adoção de critérios explícitos arrisca-se a reduzir a discussão sobre a apropriação ao nível do indivíduo e dessa forma perder o parâmetro social. Exatamente neste ponto interessa a discussão sobre a contribuição que as ontologias, garantidas suas condições de interoperabilidade, podem prestar.

Num nível de utilização mais elementar dos sistemas de informação documentária, pode-se verificar que o acesso aos documentos necessários para o desenvolvimento de pesquisas mudou significativamente, com o advento das tecnologias. A busca em diferentes catálogos e bases de dados é realizada na própria mesa de trabalho e é possível, opcionalmente, imprimir os resultados (as informações documentárias ou o texto completo) ou lê-los diretamente na tela, entre outras possibilidades. Qualquer pessoa, basicamente, possui essa expectativa hoje. A concepção fundamental de acesso facilitado pelo uso de ontologias, contudo, é outra. Interessa explorar que mudanças as tecnologias de rede provocaram no modo de construção e uso de sistemas de informação, incluindo-se a organização do conhecimento e a recuperação da informação. As ontologias permitem novas metodologias nestes níveis.

Aristóteles, Porfírio, Brentano, Kant e Peirce, assim como todos os homens, cada qual conforme o aparato intelectual de que dispõe, querem compreender o mundo buscando sua ‘ordem’, sua ‘estabilidade’, ainda que reconheçam a dinamicidade que lhe é inerente. O desconforto causado pela impossibilidade de categorizar o mundo, que se pode encontrar,

---

<sup>13</sup> Confira <<http://www.memoria.cnpq.br/areas/cee/proposta.htm>> Acesso em: 01 nov. 2009.

por exemplo, na análise que Lara (2001) faz da obra de Umberto Eco, *Kant e o ornitorrinco*, transforma qualquer novo modelo em uma hipótese para a revisão da categorização adotada.

Tendo em vista que as linguagens documentárias operam no terreno do simbólico, e que este é marcado pela convenção, é necessário revisitar algumas de suas concepções fundamentais para que se possam compreender as relações ontológicas que linguagens documentárias expressam em seu aparato conceitual. A contribuição que os estudos sobre informação e linguagem podem prestar à consolidação de projetos mais ambiciosos de representação e recuperação de informações, como a *web* semântica, por exemplo, condiciona-se à compreensão da ontologia em nível conceitual, portanto filosófico, e simbólico.

Enquanto a filosofia centra suas preocupações ontológicas no que existe no mundo, a aplicação de ontologias no domínio da representação computacional do conhecimento é bem mais modesta; entende pragmaticamente que *o que existe* é apenas o que pode ser representado (GRUBER, 1995). Neste último caso, as preocupações ontológicas restringem-se a domínios específicos e visam antes soluções de problemas absolutamente demarcados, por meio de aplicações práticas, que reflexões.

Nestes tipos de aplicações, envolvem-se basicamente um conjunto de termos e conceitos que são utilizados para a descrição do domínio (uma ontologia); uma base de conhecimento que permita aplicar isso a uma determinada realidade e os chamados agentes inteligentes, programas de computador com capacidade *autônoma* de ação. Esses agentes são denominados inteligentes porque visam a simular o comportamento humano na resolução de tarefas. Os agentes, mediante as instruções do usuário, coletam dados na internet e processam as informações por meio de regras lógicas de inferências, como as ontologias, por exemplo. Outra característica dos agentes é a cooperação, ou seja, a troca de informações que podem realizar entre si.

Por vezes tem-se a apressada impressão de que as ontologias, tal como aplicadas no domínio da engenharia do conhecimento ou dos sistemas de informação, guardam apenas relação de empréstimo terminológico com a ontologia filosófica. Na verdade os campos guardam entre si relações mais profundas, pois enquanto a ontologia filosófica se preocupa com as entidades da realidade ontológica, a engenharia do conhecimento preocupa-se com as entidades da *realidade virtual*, isto é, com a representação do conhecimento e sua implementação em sistemas computacionais.

Ainda que Gruber (1995) tente simplificar a questão afirmando que para a inteligência artificial o que existe é o que pode ser representado, não escapa do problema ontológico e filosófico fundamental: o que de fato existe para ser representado? Quais são as entidades dessa *condição virtual*? Do modo como a questão está colocada faz parecer que a oposição se dá entre o real (a realidade, o que existe) e o potencial (o que pode ser representado). Optar pela potencialidade é algo simplesmente irrealizável, pois não existem opções neste caso. Não há como separar efetivamente o que pode ser representado do que não pode sê-lo. O problema da representação não se resolve assim. Talvez fosse melhor pensar a relação *atual* e *virtual*, conforme a concepção deleuzeana para definir o escopo das ontologias. O virtual, esclarece Deleuze (1988, p. 203), “não se opõe ao real, mas apenas ao atual. O virtual possui plena realidade enquanto virtual”.

## **3 ONTOLOGIAS: A COMPREENSÃO DAS CIÊNCIAS DA INFORMAÇÃO E DA COMPUTAÇÃO**

Este capítulo aprofunda a compreensão sobre ontologias computacionais, esboçada no capítulo anterior e busca compreendê-la na intersecção entre as concepções da ciência da informação e da ciência da computação. Para diferenciar as propostas das duas ciências, retoma a discussão sobre as categorias e suas relações com as ontologias.

Apresentam-se também neste capítulo alguns aspectos tecnológicos envolvidos nas aplicações de ontologias em sistemas de informação, discutindo-se especificamente o aporte ao tratamento da informação em nível sintático possibilitado pelos metadados e as condições de interoperabilidade semântica.

### **3.1 DEFINIÇÕES E CARACTERÍSTICAS DAS ONTOLOGIAS**

A definição de ontologias tem ocupado espaço considerável na literatura da ciência da computação e, mais recentemente, da ciência da informação. Guarino e Giaretta (1995), num trabalho que discute a formação e o significado do termo, partem da definição clássica de Gruber (1995) para construir um (já então necessário) ‘esclarecimento terminológico’, como denuncia o subtítulo de seu artigo. Passados quinze anos, o campo ainda é incipiente e

o significado do termo é amplo, ambíguo e ainda está em negociação, variando conforme sua aplicação.

Em ciência da computação (mais especificamente no domínio da inteligência artificial) compreendem-se as ontologias, numa abordagem sintética, como um conjunto de conceitos e termos que podem ser usados para descrever alguma área do conhecimento ou construir sua representação. Trata-se de uma especificação formal explícita de uma conceitualização compartilhada (GRUBER, 1995). Por *formal* entende-se que esta representação seja expressa num formato legível por computadores; *explícita* significa que os conceitos, as propriedades, as relações, as funções, as restrições e os axiomas devem estar formalmente definidos e passíveis de serem manipulados por computadores. Entende-se por *conceitualização* que tal representação seja referente a algum modelo abstrato de algum fenômeno do mundo real e, finalmente, por *compartilhada*, compreende-se que esse conhecimento seja consensual.

Guarino e Giaretta (1995) coletaram diversas interpretações possíveis para o termo *ontologia*, as quais denunciam sua polissemia: a) ontologia como um sistema conceitual subjacente a uma base de conhecimento; b) ontologia como um tipo especial de base de conhecimento ou aplicação comercial; c) ontologia como um vocabulário usado por uma teoria lógica e d) ontologia como uma especificação de uma conceitualização. Guarino e Giaretta (1995) veem como principal problema nesta última interpretação, a que Gruber defende, o fato de ser baseada numa noção extensional de conceitualização.

A definição de Jacob (2003), reproduzida abaixo, tangencia a de Gruber e acrescenta-lhe aspectos importantes: as ontologias focam, sem prejuízo da informação sintática, o compartilhamento da informação semântica entre sistemas de informação.

[...] uma ontologia é uma conceitualização parcial de um determinado domínio do conhecimento, compartilhado por uma comunidade de usuários, que foi definida numa linguagem formal processável por máquina para o propósito explícito de compartilhar informação semântica entre sistemas automatizados (JACOB, 2003, p. 20, tradução nossa).

Vale citar ainda, pela abrangência, a definição de Jasper e Uschold<sup>14</sup> (*apud* BREITMAN, 2005, p. 31): independentemente do formato que assuma, uma ontologia “deve incluir um vocabulário de termos e alguma especificação de seu significado. Esta deve abranger definições e uma indicação de como os conceitos estão inter-relacionados, o que resulta na estruturação do domínio e nas restrições de possíveis interpretações de seus termos”.

A expressão *restrições possíveis* merece ser observada com mais cuidado. O compartilhamento de conhecimentos, na definição de Gruber (1995) deve ser consensual, isto é, espera-se que os atores envolvidos possam, de alguma forma, expressar interpretações que possuam pontos de contato entre si, que possuam potencial dialógico. Esta parece ser uma situação idealizada, que desconsidera a problemática da interpretação e da polissemia, a não ser que Gruber refira-se simplesmente ao compartilhamento de bases de conhecimento apenas por agentes inteligentes, realizado já no interior de contextos completamente formalizados.

Diversos projetos utilizando ontologias têm surgido recentemente. Almeida e Bax (2003) realizaram um extenso levantamento sobre estes projetos, classificando-os em: a) projetos relacionados à gestão do conhecimento, b) projetos relacionados a comércio eletrônico, c) projetos relacionados ao processamento de linguagens naturais, d) projetos relacionados à recuperação da informação na web e e) projetos relacionados à educação.

O projeto CYC, desenvolvido pela *Cycorp*, uma empresa americana de tecnologia, não relacionado no estudo acima, pode ser apontado como exemplo de projetos relacionados ao processamento de linguagens naturais. O objetivo deste projeto, que já pode ser considerado antigo para os padrões da rede, pois foi iniciado em 1984, foi codificar em linguagem de máquina os milhões de pedaços de conhecimento compreendidos pelo senso

---

<sup>14</sup> JASPER, R.; USHOLD, M. A framework for understanding and classifying ontology applications. In: INTERNATIONAL WORKSHOP ON KNOWLEDGE ACQUISITION MODELING AND MANAGEMENT, 12., 1999, Banff (Canada). **Proceedings...** Banff: [s.n.]: 1999.

comum. O exemplo apresentado no site do projeto<sup>15</sup> a respeito da aplicabilidade das ontologias é contundente. Considere as frases: a) *Fred viu o avião sobrevoando Zurique* e b) *Fred viu as montanhas sobrevoando Zurique*. Um ser humano tem pouca, ou mesmo nenhuma, dificuldade em reconhecer que na primeira frase o verbo *sobrevoar* refere-se ao avião, enquanto na segunda, refere-se a Fred. Este (re)conhecimento é intuído pelo conhecimento ontológico que se tem de que aviões voam e montanhas não. Um ser humano com inteligência mediana pode facilmente adaptar sua ‘ontologia pessoal’ para novos contextos e compreender que é possível ainda que o verbo *sobrevoar* da primeira frase refira-se a Fred, ou até mesmo que as montanhas possam sobrevoar Zurique, em contextos de ficção. Esta flexibilidade é o ideal dos sistemas de informação de modo geral, uma quimera, já que a informação no contexto desses sistemas é sempre construída, sempre devedora de um ponto de vista e, quase sempre, institucionalizada, realizada a partir de convicções ou representações individuais ou coletivas.

No que se refere aos elementos que compõem uma ontologia, segue-se a descrição de Gómez-Pérez (1999, p. 35, tradução nossa): a) conceitos organizados numa taxonomia (classes e subclasses) – os conceitos podem ser abstratos ou concretos, simples ou compostos (*composite*), reais ou fictícios; b) relações entre os conceitos, formalmente definidas; c) funções – casos especiais de relacionamento em que um conjunto de elementos tem uma relação única com um outro elemento; d) axiomas – regras, sempre verdadeiras; e) instâncias – usadas para representar elementos.

Estes componentes podem ser visualizados na Figura 5, que representa uma ontologia (um modelo) construída com base nas informações apresentadas na norma da Associação Brasileira de Normas Técnicas (ABNT) para elaboração de referências (NBR 6023).

---

<sup>15</sup> Disponível em: <[http://www.cyc.com/cyc/technology/whatis\\_cyc\\_dir/whatsincyc](http://www.cyc.com/cyc/technology/whatis_cyc_dir/whatsincyc)>. Acesso em: 26 out. 2009.

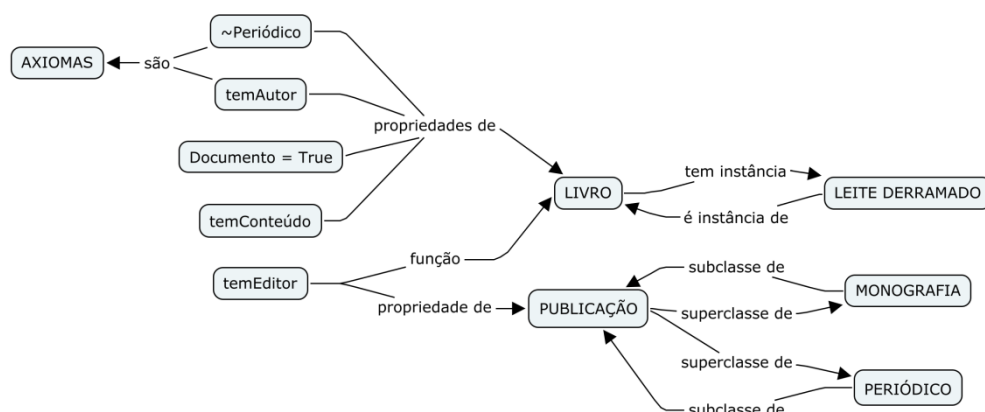


Figura 5 – Exemplo de componentes de ontologias

A indicação da instância como componente da ontologia causa, na literatura de ciência da computação, discussão semelhante à que provoca na ciência da informação a utilização de conceitos individuais na organização de tesouros. Diferentemente dos tesouros, entretanto, os quais demandam a delimitação de campos conceituais específicos como garantia da integridade das relações entre os conceitos e que raramente são aplicados em casos (domínios) extremamente específicos ou particulares, as ontologias, porque cobrem um espectro mais amplo de aplicações, podem exigir, num nível determinado, este tipo de compreensão. Apresenta-se a seguir a classificação das ontologias, com base em Maedche (2002):

- ontologias de alto nível – descrevem conceitos genéricos, como espaço, tempo, evento e outros, os quais são independentes de domínios particulares; ajustam-se melhor às ações de interoperabilidade, pois podem ser compartilhadas por grande número de usuários;
- ontologias de domínio – descrevem o vocabulário relacionado com um domínio genérico por meio da especialização dos conceitos introduzidos na ontologia de alto nível;
- ontologias de tarefa – descrevem o vocabulário relacionado com uma tarefa ou atividade genérica por meio da especialização dos conceitos introduzidos na ontologia de alto nível;

- d) ontologias de aplicação – são as ontologias mais específicas; os conceitos neste tipo de ontologia frequentemente correspondem às especializações dos conceitos usados nas ontologias de domínio ou de tarefa.

Há diversos requisitos tecnológicos para que se possa utilizar com sucesso as ontologias na representação do conhecimento. Muitos deles, como os que se referem aos protocolos de comunicação, por exemplo, não serão tratados nesta reflexão. Interessam neste trabalho, pela necessidade de recorte temático, os aspectos que envolvem mais proximamente o referencial teórico da ciência da informação na discussão.

Deste ponto de vista, importa compreender qual o papel reservado aos metadados no desenvolvimento de sistemas de informação documentária para que estes sistemas possam cooperar. Ainda na perspectiva da cooperação, discute-se o conceito de interoperabilidade como característica intrínseca das ontologias. Retoma-se, no próximo subcapítulo, a discussão sobre as categorias, com o fim de compreender em que sentido o conhecimento que a ciência da informação desenvolveu neste tema pode auxiliar a repensar as ontologias, ainda em seus aspectos de interoperabilidade.

## **3.2 METADADOS E ONTOLOGIAS: CONSTRUÇÃO, COMUNICAÇÃO E USO DA INFORMAÇÃO**

A biblioteconomia ocupa com estudos sobre catalogação parte considerável do investimento teórico-prático que qualquer campo do conhecimento necessariamente realiza. O contexto das tecnologias digitais certamente contribui para que haja continuidade neste investimento. Grande parte dos problemas que se colocam para o tratamento dos objetos digitais de informação referem-se, conforme discutido em outro trabalho (MOREIRA, 1998), a um dos questionamentos mais antigos e mais elementares nos campos que têm

como objeto o tratamento da informação: por que e para quem são descritas as fontes de informação e qual o nível apropriado para descrevê-las?

A catalogação, ou seja, a operação de criação de entradas para um catálogo, é anterior à profissão de bibliotecário e à própria biblioteconomia. Aliás, é exatamente desse esforço histórico de sistematização que emerge este campo. Sem padrões mínimos de descrição não se pode falar em bibliotecas e nem mesmo em sistemas de informação.

Nada mais natural, portanto, que o estudo sobre este tema fosse retomado no contexto das tecnologias digitais, pois é exatamente quando o conjunto de recursos de informação cresce e se torna mais complexo que o nível de exigência aumenta com relação à catalogação. Em outros termos, quanto maior e mais diversificada forem a coleção de objetos de informação e sua comunidade usuária, proporcionalmente necessário se faz a tarefa de gerar índices que possam representar e recuperar (localizar) a informação.

A catalogação dos recursos digitais disponíveis eletronicamente na internet ganhou o nome de metadados e ambos têm sido confundidos. Na prática, entretanto, catalogação e metadados não significam exatamente a mesma coisa. O que diferencia a informação documentária gerada no interior de catálogos nos tradicionais sistemas de recuperação de informação dos metadados é não apenas o ambiente digital, mas também a finalidade. O que há de diferente é que o nível de descrição de metadados é, de modo geral, menos profundo do que o que pretende a catalogação. Não poderia ser diferente, claro, em se tratando de rede mundial. Dessa forma mesmo os não-especialistas e até mesmo os mecanismos de busca podem descrever os metadados, o que não ocorre com o MARC, por exemplo, que pode ser considerado, aliás, o primeiro padrão de metadados, criado na década de 60, nos Estados Unidos.

Pode-se imaginar que em face das modernas técnicas de localização de informação, disponíveis atualmente, das quais o *Google™* é, certamente, o grande expoente; e das possibilidades de pesquisa em texto completo, a descrição de recursos de informação tenha se tornado anacrônica, mas trata-se de engano. Localizar, na internet, um texto que faça

referência a determinado autor, ou mesmo textos produzidos por este autor, por exemplo, é relativamente simples, mas não é igualmente simples localizar *apenas* os textos produzidos por este autor determinado. Os metadados possibilitam essa localização por meio do uso da sintaxe `<autor="nome do autor pesquisado">`, com a condição de que tenham sido respeitados os níveis de descrição na geração do recurso.

O termo metadado (*metadata*) nasceu na ciência da computação e no seu interior é definido, apressadamente, como dado para descrever outros dados. Caplan (2003) informa que o termo foi criado por Jack E. Myers, na década de 60, e registrado em 1986, como marca da *Metadata Company*, empresa norteamericana que desenvolve *softwares* e serviços relacionados com medicina e saúde.

No início dos anos 90, o termo era usado com o sentido de “informação necessária para tornar arquivos de computadores úteis para seres humanos” (CAPLAN, 2003, p. 1). Com o surgimento da internet, o metadado ganhou a acepção de descrição de objetos da rede. O termo passa a integrar a terminologia da ciência da informação por volta de 1995, com a criação do *Dublin Core Metadata Element Set (Dublin Core)*, formado, principalmente, por participantes do então nascente W3C (*World Wide Web Consortium*).

Metadado, conforme a definição presente no documento da *National Information Standards Organization* – NISO (2004, p. 1, tradução nossa), é “informação estruturada que descreve, explica ou facilita a recuperação, o uso e a gestão do recurso de informação”. Caplan (2003, p. 3) define metadados como “informação estruturada sobre um recurso de informação de qualquer tipo de mídia ou formato”, e apresenta uma tipologia que permite agrupá-los, conforme sua aplicação, em *descritivos* (utilizados com finalidade de descoberta, identificação, seleção, localização e aquisição), *administrativos* (informação destinada a facilitar a gestão de recursos; inclui, por exemplo, informação sobre como e quando o recurso foi criado) e *estruturais* (cuja função é manter os objetos digitais complexos agrupados, facilitar a navegação e a apresentação). Esta classificação é apenas funcional e não se refere às qualidades inerentes dos metadados, já que todo metadado é, por definição, descritivo de alguma coisa.

Há diversos padrões ou esquemas de metadados, os quais são conjuntos de regras de conteúdo, semântica e sintaxe. Dentre os mais conhecidos estão o *Dublin Core*<sup>16</sup>; o *Visual Resources Association Core Categories* (VRA Core)<sup>17</sup> e o *Encoded Archival Description* (EAD)<sup>18</sup>. Enumeram-se no Quadro 5, para efeitos de ilustração, os elementos *Dublin Core*.

Elemento	Identificador	Definição
Colaborador	<i>Contributor</i>	Entidade responsável pelas contribuições intelectuais ao conteúdo do recurso
Cobertura	<i>Coverage</i>	Extensão ou alcance (de tempo ou de espaço) do recurso
Criador	<i>Creator</i>	Entidade originalmente responsável pela existência do recurso
Data	<i>Date</i>	Data associada a um evento do ciclo de vida do recurso
Descrição	<i>Description</i>	Descrição do conteúdo do recurso
Formato	<i>Format</i>	Manifestação física ou digital do recurso
Identificador	<i>Identifier</i>	Referência não ambígua ao recurso, dentro de determinado contexto
Idioma	<i>Language</i>	Idioma do conteúdo intelectual do recurso
Editor	<i>Publisher</i>	Entidade responsável por tornar o recurso disponível
Relação	<i>Relation</i>	Um recurso relacionado
Direitos	<i>Rights</i>	Informações de direitos sobre o recurso ou associadas a ele
Fonte	<i>Source</i>	Recurso relacionado, do qual o recurso descrito é derivado
Assunto	<i>Subject</i>	Conteúdo (assunto) do recurso
Título	<i>Title</i>	Nome dado ao recurso
Tipo	<i>Type</i>	Natureza ou gênero do recurso

Quadro 5 - Conjunto de elementos Dublin Core

Por mais utópico que possa parecer, considerando-se a base tecnológica de que se dispõe nesta primeira década do século XXI, é preciso empreender a grande tarefa de indexar recursos da internet. Num primeiro nível, sintático, relativamente mais simples de

<sup>16</sup> Disponível em: <<http://dublincore.org>>. Acesso em: 28 fev. 2010.

<sup>17</sup> Disponível em: <<http://www.vraweb.org>>. Acesso em: 28 fev. 2010.

<sup>18</sup> Disponível em: <<http://www.loc.gov/ead/>>. Acesso em: 28 fev. 2010.

resolver porque envolve padrões ou protocolos já aceitos pela comunidade de usuários; já há avanços. Cita-se novamente o *Google™* e seu poderoso algoritmo como exemplo do sucesso da localização de informação sintática na *web*. Entretanto, quando se trata de informação semântica, nada é tão pacífico, pois neste caso é preciso dispor de tecnologias que insiram a informação em contextos determinados, nos quais ela (a informação) possa assumir significados relacionais.

É necessário, então, produzir informação sobre informação (metadados) de forma padronizada, com vistas não somente à descrição ou localização de objetos de informação, mas também como forma de viabilizar as trocas entre sistemas de informação (numa concepção ampla). Dito de outro modo: é preciso ratificar os formatos ou padrões de catalogação utilizados entre sistemas de informação documentárias, em conjunto com os metadados, para que haja condições de, superado este nível, investir-se mais fortemente em teorias e metodologias para a adoção de padrões que viabilizem a interoperabilidade semântica entre estes sistemas. Neste aspecto, acredita-se, as ontologias podem favorecer a tarefa da ciência da informação.

Alguns grupos de pesquisa veem reforçando esta tendência e desenvolvendo pesquisas sobre a aplicação de ontologias em contextos de informação documentária. Dentre os grupos de pesquisa cadastrados no *Diretório dos grupos de pesquisa no Brasil* do CNPq, ainda são poucos os que, pela linha de pesquisa apresentada, são mais proximamente interessantes às pesquisas sobre o tema em ciência da informação. Realizou-se consulta à base de dados com os seguintes parâmetros: 'grande área do grupo = ciências sociais aplicadas' e 'área do grupo = ciência da informação'. Os resultados apresentaram apenas cinco grupos (Quadro 6).

Sem os parâmetros restritivos de *grande área* e de *área*, foram encontrados outros grupos, todos relacionados no Apêndice B.

---

Nome: Redes de conhecimento e informação

Instituição: Universidade Estadual de Londrina – Centro de Educação Comunicação e Artes

Líder: Maria Inês Tomaél

Linhas de pesquisa: Compartilhamento da informação e do conhecimento. Uso de ontologias para a representação do conhecimento.

---

Nome: Ontologia e taxonomia: aspectos teóricos e metodológicos

Instituição: Universidade Federal Fluminense – Departamento de Ciência da Informação

Líder: Maria Luiza de Almeida Campos

Linhas de pesquisa: *Datawarehouse*. Fluxos e mediações sócio-técnicas da informação.

---

Nome: Estudos e pesquisa em mediação e comunicação da informação

Instituição: Universidade Federal da Bahia – Departamento de Documentação e Informação

Líderes: Henriette Ferreira Gomes e Maria Isabel de Jesus Sousa

Linhas de pesquisa: Informação e contextos sócio-econômicos.

---

Nome: Educação, conhecimento e tecnologia

Instituição: Universidade Federal do Rio Grande – Instituto de Ciências Humanas e da Informação

Líder: Gisele Vasconcelos Dziekaniak

Linhas de pesquisa: Informação, educação e sociedade. Organização do conhecimento.

---

Nome: Gestão da informação e do conhecimento

Instituição: Universidade Federal de Minas Gerais - Escola de Ciência da Informação

Líder: Ricardo Rodrigues Barbosa e Marta Macedo Kerr Pinheiro

Linhas de pesquisa: Ontologias organizacionais.

---

Quadro 6 - Grupos de pesquisa sobre ontologias em ciência da informação

### 3.3 INTEROPERABILIDADE SEMÂNTICA

Quando se organiza um sistema de informação, adota-se como condição necessária à noção de sistema a promoção da cooperação em quaisquer das fases do processamento. Naturalmente nem sempre se pode realizar isso sem que outros fatores concorram positiva ou negativamente. Ainda que se considere a oferta de fontes de informação por meio de

arquivos disponíveis na internet como um avanço realmente significativo, é sempre necessário considerar que toda informação tem custo e tem autoria, isto é alguém se responsabiliza pela criação, inserção, representação e manutenção do arquivo num *site* ou numa base de dados. A idéia de cooperação, então, não passa necessariamente pela benemerência, isto é, não se trata de pessoas físicas ou jurídicas dotadas de boa vontade e de condições de oferecer acervos digitalizados ou conjuntos de metadados isolados. Longe disso, requer o estabelecimento de políticas de informação que ofereçam condições para viabilizar a produção e a recepção da informação científica e tecnológica e que garantam condições tecnológicas para que o processo ocorra.

“O que existe num conjunto de livros? Que espetáculo teríamos”, pergunta o imaginador Otlet (1996, p. 107, tradução nossa), “se, por um milagre bibliográfico, nos fosse dado de repente o poder de ler todos eles ao mesmo tempo, todas as suas partes, todas as suas páginas?”. Ainda que alguns dos ideais de Otlet já tenham sido de algum modo colocados em prática, pelo menos aqueles dependentes apenas da evolução tecnológica, a árdua tarefa da interpretação ainda não foi (e nem se vislumbra um horizonte no qual será) resolvida. Há, contudo, não se pode negar, tecnologias, como o hipertexto, que modificam o estatuto da leitura, que promovem novos modelos de interação entre autor, texto e leitor, por conta da modificação que realizam na fisiologia da leitura. “A revolução do texto eletrônico será ela também uma revolução da leitura. Ler sobre uma tela”, ensina Chartier (1994, p. 100), “não é ler um *códex*”.

A quantidade e a qualidade da interação presente nos fluxos de comunicação não se separam, evidentemente, da base tecnológica disponível. Em outras palavras, a tecnologia influencia diretamente o modo como as pessoas produzem e trocam informações, seja no âmbito pessoal, profissional ou acadêmico. A mudança mais significativa que a internet provocou no fluxo da comunicação, de modo geral, relaciona-se à possibilidade de interação *todos x todos*. Outras mídias, como a televisão ou o telefone, por exemplo, condicionam este tipo de interação, a *um x todos* ou *um x um*.

As novas possibilidades de interação influem, de modo natural, nos sistemas de informação e o interesse em interoperabilidade cresce proporcionalmente ao interesse em intercâmbio de informações documentárias. Superar as barreiras tecnológicas e de organização (representação temática, catalogação e garantias de acesso) tornou-se objetivo comum hoje a todos os envolvidos nos processos de tratamento da informação. Subsidiando esta concepção está o objetivo de criar condições para que os usuários possam acessar múltiplas bases de dados, usando uma interface definida e uma linguagem de busca normalizada. Além das vantagens econômicas, visam-se também o aumento das trocas de idéias entre campos científicos e a possibilidade de buscas mais exaustivas (MAI, 2003).

Define-se interoperabilidade como “a capacidade de um sistema de *hardware* ou de *software* de se comunicar e trabalhar efetivamente no intercâmbio de dados com outro sistema, geralmente de tipo diferente, projetado e produzido por um fornecedor diferente” (REITZ, 2004, p. 373)<sup>19</sup>. No universo das bibliotecas digitais a interoperabilidade encontra-se na “possibilidade de um usuário realizar buscas a recursos informacionais heterogêneos, armazenados em diferentes servidores na rede, utilizando-se de uma interface única sem tomar conhecimento de onde nem como estes recursos estão armazenados” (MARCONDES; SAYÃO, 2001, p. 27).

A promoção da interoperabilidade requer acordos de cooperação em três níveis: técnico, de conteúdo e de organização. O primeiro nível (sintático) refere-se aos formatos, protocolos, sistemas de segurança e outros que garantam a troca de mensagens; o segundo (sintático-semântico) refere-se aos metadados, e inclui acordos sobre semântica e interpretação da informação; o terceiro nível organizacional envolve as regras básicas de acesso, a manutenção de coleções e serviços, pagamento, autenticação etc. (ARMS *et al.*, 2002).

---

<sup>19</sup> A tradução é de Sayão e Marcondes (2008), que usaram a versão eletrônica *on-line* do *Online Dictionary for Library and Information Science (ODLIS)*.

A biblioteconomia desenvolveu, ao longo de sua história, diversas técnicas para viabilizar e facilitar o intercâmbio de informações entre unidades de informação. A garantia de interoperabilidade, de diálogos entre sistemas de informação diferentes (em termos de conteúdo ou de estrutura) formam o esteio de grandes idéias ou projetos, como a utopia universalista de Otlet ou o pragmatismo de Dewey. Talvez não seja exagero dizer que nunca a base tecnológica foi mais adequada do que neste início do século XXI.

Talvez não seja igualmente exagero afirmar, numa curta digressão, que Otlet tenha sonhado com condições assim, pois, ainda que fosse um entusiasta da documentação e, sem dúvida, um amante dos livros, demonstra, em algumas passagens de sua obra clássica, preocupações com a falta de sustentabilidade do modelo impresso, como se pode perceber pelas citações abaixo, as quais mostram também uma *precoce* preocupação ecológica:

É preciso assinalar os prejuízos da indústria do papel do ponto de vista do desmatamento. [...] Esses 30.000 periódicos, com uma tiragem de 10.800 milhões de exemplares, consomem a cada dia cerca de 1000 toneladas de polpa de madeira, exatamente 350.500 vagões conduzidos por 1.800 locomotivas [...]. Isso sem levar em conta os papéis de embrulho, cartões, folhetos, folhas para escrita etc. [...].

Uma semana de publicação de um dos jornais atuais [ca. 1934] de grande circulação é uma floresta que desaparece em algum lugar (OTLET, 1996, p. 50, tradução nossa).

A popularização da Classificação Decimal de Dewey (CDD) e posteriormente da Classificação Decimal Universal (CDU), entre outros sistemas de classificação, visou a garantir, no limite, que os mesmos documentos receberiam notações equivalentes em diferentes unidades de informação, o que tornaria a representação mais uniforme e poderia garantir condições de interoperabilidade e de fluxo informacional mais profícuos.

Numa abordagem nitidamente onomasiológica, Otlet (1996, p. 382, tradução nossa), num exemplo sobre o uso da classificação decimal, afirma:

A classificação também se chama decimal porque organiza dez classes e dentro de cada uma delas dez grupos que, por sua vez, subdividem-se em dez nas quais todos os assuntos são distribuídos. Bastará traduzir as mesmas tabelas a todos os idiomas para o número 541.7 tenha em todas elas a mesma significação. Bastará subdividir

este número em novos decimais, correspondente a subdivisões de 'alotropia' (isomeria, tautomeria, polimeria) para seguir o desenvolvimento científico futuro em todas as suas ramificações. Bastará anotar este número em todos os documentos, livros, folhetos, artigos, notas bibliográficas etc., relacionados com a alotropia para assegurar-se de sua classificação dentro da universalidade dos temas científicos.

Como se pode perceber mais claramente hoje, isso seria mais fácil se as bibliotecas estivessem conectadas em rede, formando sistemas e compartilhando soluções tecnológicas.

A interoperabilidade, entretanto, não depende apenas de requisitos técnicos, essa é apenas a sua face que está mais evidente, há outros níveis envolvidos também, tais como: as interoperabilidades semântica, política-humana, legal e internacional (SAYÃO; MARCONDES, 2008). O ideal de compartilhamento e de interoperabilidade, aliás, é observável em diversos momentos anteriores ao advento da telemática. Norris<sup>20</sup> (*apud* MAI, 2003) relata o esforço que os monges franciscanos ingleses faziam, viajando de mosteiro para mosteiro, no final do século XIII, com seu catálogo compartilhado, o *Registrum Librorum Angliae*, de 183 bibliotecas monásticas. Esse recurso era usado para que pudessem manter-se informados sobre o acervo disponível em outros mosteiros.

Um dos problemas mais claramente colocados à interoperabilidade, que, aliás, também influenciou o projeto de adoção *uniforme* dos sistemas de classificação, como queriam seus idealizadores, refere-se à questão da complexidade das culturas envolvidas nos processos de construção da informação documentária. Se num contexto particular de uma unidade de informação já é muito difícil garantir a coerência intraindexadores e interindexadores, o problema agrava-se muito em contextos de intercâmbio intercultural ou internacional. É raro que existam relacionamentos um-para-um entre palavras de idiomas diferentes. Hjørland<sup>21</sup> (*apud* MAI, 2003), utiliza e adapta o exemplo de Hjelmslev para

---

<sup>20</sup> NORRIS, Dorothy M. *A history of cataloguing and cataloguing methods*. London: Grafton, 1939.

<sup>21</sup> HJØRLAND, Birger. *Principia informatica: foundational theory of information and principles of information services*. In: CONFERENCE ON CONCEPTIONS OF LIBRARY AND INFORMATION SCIENCE, 4., 2002, Greenwood Village. *Proceedings...* Greenwich Village: Libraries Unlimited, 2002. p. 113

ilustrar a relatividade cultural no significado das palavras (Quadro 7), expressas no modo como cada cultura recorta a realidade.

inglês	alemão	dinamarquês	francês	italiano	Espanhol
	baum	træ	arbre	albero	arbol
wood	holz		skov	bois	legno
woods	wald	basco			basque
forest				foret	foresta

Quadro 7 – Relatividade cultural no significado das palavras

Mai (2003) defende que a interoperabilidade entre sistemas de informação só será possível se houver adoção híbrida de sistemas de classificação (organização mais ampla do conhecimento), em conjunto com sistemas de indexação especiais (organização orientada ao domínio). Pode-se acrescentar a esta hibridez também a construção de ontologias comuns, cujo controle mais rigoroso dos relacionamentos entre conceitos viabiliza a criação de estruturas comuns integradas.

Dentre as principais justificativas para o desenvolvimento e a adoção de ontologias está, certamente, sua possibilidade de reuso. Em comparação com as linguagens documentárias como o tesauro, por exemplo, as ontologias são mais *sustentáveis*, porque as relações que permitem estabelecer entre os conceitos são mais formalmente definidas, ou definidas com mais profundidade.

O conceito de reuso implica uma dupla caracterização das ontologias, por um lado as ontologias devem ser específicas o suficiente para descrever o campo de aplicação, por outro, não podem prescindir de certo nível de generalização que possa garantir seu intercâmbio em contextos diferenciados. A tarefa seria simples se os termos fossem rótulos e se fosse possível operar, em todos os contextos, com macro-ontologias, como se fora uma

linguagem universal e se os termos fossem estáticos em face dos conceitos que pretendem representar. Como operam no terreno do simbólico, as ontologias, como qualquer outra construção simbólica, demandam alinhamento. A tarefa de alinhamento de ontologias, contudo, está longe de ser simples, como permite ilustrar a Figura 6.

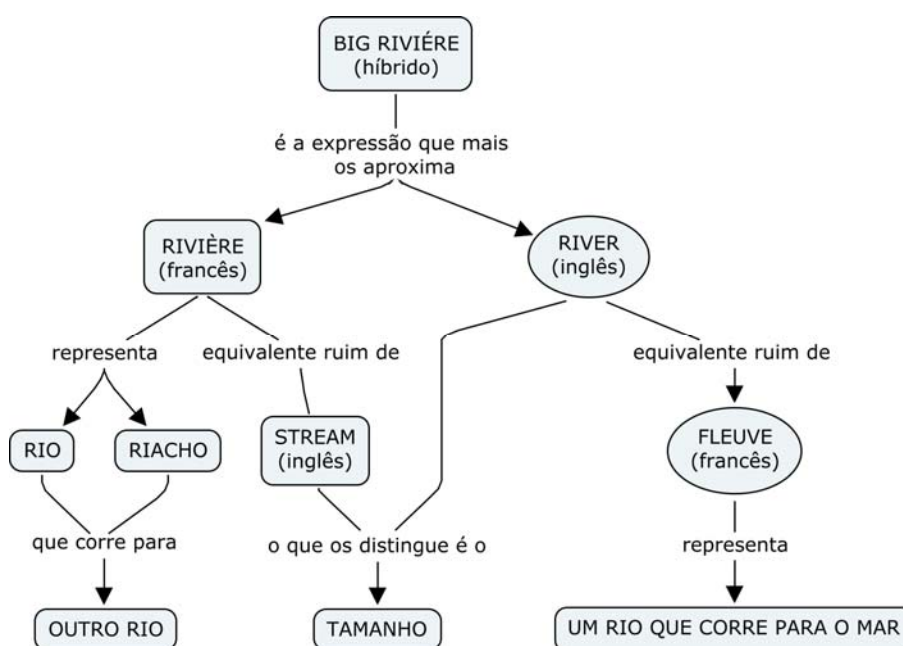


Figura 6 – As variações linguísticas geográficas como empecilho para o alinhamento de ontologias<sup>22</sup>

Este desalinhamento ocorre por uma variedade de diferenças culturais, geográficas, linguísticas, técnicas etc. No exemplo ilustrado na Figura 6, a questão é, principalmente, cultural, como decorrência de uma questão geográfica, pois os maiores rios da França correm para o Atlântico ou para o Mediterrâneo. Nos Estados Unidos, existem grandes rios, como o Ohio e o Missouri, que correm para o Mississippi. Este exemplo, proposto inicialmente por Saussure, em seu *Curso de linguística geral*, é utilizado, aliás, por diversos estudiosos da questão, em diferentes contextos, como, entre outros, Sowa (2000), Kalfoglou

<sup>22</sup> Baseada em Sowa (2000, p. 410).

e Schorlemmer (2003) e Ilari (2004). Esta arbitrariedade traz questionamentos óbvios como, por exemplo, os que se fazem à tarefa de tradução: “se cada língua recorta a seu modo a experiência, como explicar que as pessoas traduzam de uma língua para outra?” questiona Ilari (2004, p. 66).

### **3.4 AS RELAÇÕES ENTRE INTEROPERABILIDADE E CATEGORIAS**

A interoperabilidade aplicada aos sistemas de informação documentária refere-se à capacidade desses sistemas para possibilitar a cooperação mesmo em condições diferentes de representações. Tais sistemas, contudo, devem ser construídos visando a cooperação. Se as ontologias podem facilitar a tarefa, elas não prescindem das categorias. É preciso, então, recuperar o conceito de categoria e entender sua posição em relação às ontologias. De modo amplo, a categorização pode se compreendida como uma simplificação da realidade complexa em processos de generalização ou abstração, e discriminação. O homem depende da categorização para gerar pensamentos complexos e para compreender a realidade por meio da separação e junção de seus elementos. Sem generalização, defende Kostina (2003) não poderíamos chegar a uma estruturação conceitual e sem discriminação tudo seria uno, não havendo nem mesmo pensamento.

Sem categorização não é possível estruturar informações e produzir conhecimentos, seja no nível do indivíduo e seu processo cognitivo, seja no âmbito da coletividade dos sistemas.

A compreensão da realidade prima pelo conforto, pela ordem, como elementos quase naturais do homem. Se os elementos da realidade não se encaixam nas categorias de que o indivíduo já dispõe, se suas hipóteses são rejeitadas, é forçoso ajustá-las como

condição para o conhecimento; caso em que se propõe um *novo modelo*, uma *nova hipótese*, mas “embora no processo classificatório procuremos invariavelmente pela segurança da exclusão e da inclusão, tal como num leito de Procuste<sup>23</sup>, é necessário admitir a instabilidade e a possibilidade de coexistência de modelos”, ensina Lara (2001). A categorização é um dos aspectos mais importantes da linguagem, pois possibilita encontrar características semelhantes em objetos individuais, ainda que tais objetos sejam dessemelhantes.

Para Iyer (1995), que apresenta uma abordagem conciliatória entre os parâmetros clássicos e contemporâneos, os conceitos e as categorias formam as vigas mestras da organização e representação do conhecimento. A compreensão de conceito neste caso se dá como em Dahlberg (1978), condicionada à identificação de suas características, dos enunciados verdadeiros necessários e verificáveis sobre um referente e às relações que se dão pelo compartilhamento de características. A base de relacionamento entre conceitos é a posse de características comuns, características essas que formam o nível mais genérico das categorias. Assim, o conceito permite inferir mais informação do que a simples visão ou audição da palavra ou coisa permitem e é essa exatamente a sua função.

É oportuno revisitar os estudos de Iyer (1995) sobre os conceitos, e compreender suas denominações para o paradigma clássico e o *novo paradigma*. O primeiro tem como base Aristóteles e o segundo traz as contribuições de Rosch e Wittgenstein, principalmente.

Num paradigma clássico, as categorias emprestam estabilidade e ordem ao mundo percebido, por isso é impossível pensar sem formar categorias: qualquer objeto ou ideia encontrados são automaticamente comparados com outras coisas que se conhece; o que implica afirmar que todos os objetos em uma categoria são considerados similares ou equivalentes (IYER, 1995).

---

<sup>23</sup> Segundo a mitologia grega, Procusto atacava suas vítimas e as amarrava sobre um mesmo leito que usava como medida, ali cortava ou esticava as pernas das vítimas conforme fossem inadequadamente grandes ou pequenas (FRANCHINI; SEGANFREDO, 2007.).

A teoria clássica do conceito é baseada na suposição de que o mundo real é estruturado em grupos hierárquicos que compartilham propriedades inerentes. Para isso é preciso compreender as categorias em limites rígidos (IYER, 1995): um livro é ou não é um documento, por exemplo, neste modelo. Essa concepção está ancorada em três hipóteses: a) conceito como representação sintética da classe; b) os aspectos definidores de uma categoria são tanto individualmente necessários como suficientemente unidos para determinar a categoria e c) os conceitos são aninhados; a categoria subordinada possui todos os aspectos da categoria superordenada.

Na visão mais flexível de Rosch (*apud* IYER, 1995) as categorias variam conforme a linguagem e a cultura. Essa concepção baseia-se no reconhecimento de protótipos (exemplos típicos numa categoria), que ainda são formados a partir da hierarquização, como no paradigma clássico, mas sem a sua fixidez: rosa, exemplifica Iyer (1995), é um exemplo prototípico de flor. Isto é, rosa define melhor a categoria flor do que miosótis, por exemplo. As categorias, na proposta de Rosch (*apud* IYER, 1995) acomodam-se em estruturas graduadas com limites tênues entre seus membros. Estruturar gradualmente um conjunto de categorias significa admitir que os objetos podem ser organizados do melhor ao pior exemplo típico ou vice-versa.

No interior de uma categoria conceitual, há representantes mais nucleares e representantes mais periféricos, explica Finatto (2001) que também apresenta um exemplo clássico: um membro nuclear da categoria *ave* poderia ser *sabiá*, enquanto *pinguim* seria periférico.

Outra alternativa à visão hierárquica clássica é a de categorias formadas por *semelhanças de famílias*, de Wittgenstein (IYER, 1995; BLAIR, 2006). Diversão, acaso ou competição são, por exemplo, propriedades comuns a todos os jogos, sejam eles de tabuleiro, de cartas, com bolas ou outros. Alguns jogos, no entanto, como o xadrez, envolvem menos acaso do que outros, como as cartas. O agrupamento dos jogos, portanto, defende Wittgenstein (IYER, 1995; BLAIR, 2006), se dá por semelhanças de famílias, uma vez que não há um conjunto comum de propriedades compartilhado por todos os jogos. Bloor

(1983) objeta que em se considerando todos os tipos de similaridade, os agrupamentos (classes) podem incluir tudo. Uma faca, por exemplo, pode ser similar a um garfo, uma vez que ambos são usados para levar alimentos à boca, mas, em outra família, ambos são similares também a uma chave de fenda, porque são metálicos etc.

Identificar inicialmente a *família* e posteriormente as *semelhanças* entre seus membros são operações que se ajustam relativamente bem às metodologias de construção de ontologias. A identificação das semelhanças não prescinde do mapeamento ontológico, pois ainda que não se possa, nesta proposta, identificar um conjunto comum de propriedades, é preciso identificar e descrever essas mesmas propriedades com o fim de diferenciá-las, por um lado e identificar suas semelhanças, por outro. Posto que no caso das ontologias as relações entre os conceitos são, devem ser, formalmente definidas, o acordo fica mais facilitado.

Num *paradigma de estrutura graduável* (IYER, 1995) a definição de categorias, e a ordenação de seus membros, depende grandemente do sujeito que as realiza. O ambiente, a cultura e o conhecimento de uma pessoa interferem na seleção da categoria *fruta* ou *vegetal* que ela vai eleger para inserir o *tomate*, por exemplo, do mesmo modo que em alguns contextos culturais específicos *pinguim* pode ser melhor exemplo de *pássaro* do que *pardal*. Os dêiticos (aqui, agora, este, aquele, dela etc.) podem ser apontados como exemplos bastante enfáticos de palavras dependentes de contexto.

Iyer (1995, p. 43-44, tradução nossa) apresenta as seguintes características como definidoras das *categorias de estrutura graduável*:

- a) não possuem aspectos inerentes ou ‘naturais’, como as categorias clássicas;
- b) os membros da categoria não são iguais e sua estrutura é marcada por uma gradação de tipicidade, com alguns membros sendo melhores exemplos que outros;
- c) a estrutura da categoria é assimétrica, desse modo uma classe pode ter dez elementos e a outra apenas três;

- d) considera a abordagem comportamental e ecológica para a categorização de como as pessoas ordenam os objetos na categoria conforme sua tipicidade;
- e) tais categorias são essencialmente instáveis e constantemente redesenhadas se as olharmos em contextos diferentes e direcionadas por objetivos.

A categorização torna-se assim, não uma estrutura para definir ou enquadrar o universo, mas um processo para formar ferramentas para o pensamento, grupos momentaneamente úteis por associação. Esta abordagem é especialmente relevante no contexto de usuários buscando por informação e na identificação do modo como eles categorizaram os conceitos naquele momento, para expressar sua necessidade imediata de informação.

## **4 TEORIAS DA TERMINOLOGIA**

Ainda que nova como ciência, a prática terminológica já possui registros mais antigos. Como exemplo, Barros (2004) informa sobre a existência de dicionários temáticos monolíngues desde 2600 a.C., feitos pelos sumérios em forma de tijolos de argila. Não há neste modelo, evidentemente, preocupações com a linguagem de representação como irá aparecer posteriormente, trata-se apenas de compilação de termos.

Apenas no séc. XVII começa-se a compreender a terminologia como campo científico e pela forma ambígua como é conhecida até hoje: como conjunto de termos de uma área técnica ou científica e como disciplina de natureza linguística que estuda esse conjunto e propõe metodologias para sua construção. A maior contribuição dessa época vem do naturalista Karl Von Lineu (1707-1778) e sua taxonomia.

Já no século XX, outro representante oferece contribuição importantíssima: Eugen Wüster (1898-1977), engenheiro austríaco, ajuda a firmar a terminologia como disciplina científica, estabelecendo as bases da chamada teoria clássica da terminologia.

Com o auxílio de Cabré (1993) pode-se sintetizar a evolução histórica da terminologia em quatro períodos: as origens (de 1930 a 1960); a estruturação (de 1960 a 1975); a eclosão (de 1975 a 1985) e a expansão (a partir de 1985).

### **4.1 TEORIA GERAL DA TERMINOLOGIA**

As abordagens teóricas clássicas em terminologia são representadas pelos trabalhos pioneiros das primeiras escolas, a de Viena (ancorada em Wüster com sua concepção mais

filosófica e lógica da terminologia), a de Praga e a Russa (D.S. Lotte e Dredzen, cuja concepção da terminologia é mais linguística).

Wüster expôs os fundamentos da teoria geral da terminologia em sua tese de doutorado sobre normalização internacional da terminologia técnica, trabalho que motivou a criação do Comitê 37 (Terminologia, princípios e coordenação), da ISO. Em sua tese, manifestou preocupação com aspectos da normalização terminológica e criou uma metodologia baseada na sistematização dos conceitos para a elaboração de dicionários terminológicos (BARROS, 2004). Desenvolveu sua teoria geral da terminologia (TGT) nos anos 1970. A TGT concebe a terminologia como “uma ciência de caráter filosófico, que mantém relações privilegiadas com a Lógica, a Teoria da Classificação e a Linguística (posteriormente também com a Informática)” (BARROS, 2004, p. 55).

Essas bases teóricas iniciais possuem objetivos acentuadamente pragmáticos, em consonância com o estatuto científico daquele período; visam a facilitar o fluxo da comunicação das ciências num plano internacional, em contextos de comunicação expandidos. Referendando as críticas à Escola de Viena e aos seus seguidores, Krieger e Finatto (2004, p. 33) consideram que sua vocação inicial está associada à “missão de controlar e padronizar os usos terminológicos em escala mundial”, querendo ser universal sem metodologias adequadas tornou-se, por esta mesma razão, uma abordagem redutora.

## **4.2 TEORIA COMUNICATIVA DA TERMINOLOGIA**

A teoria comunicativa da terminologia (TCT), nascida das observações sobre as limitações da TGT, tem entre seus divulgadores Maria Teresa Cabré e os pesquisadores do Instituto de Linguística Aplicada, da Universidade Pompeu Fabra, em Barcelona. Para a TCT o modelo de Wüster é reducionista e idealista; neste modelo sustentam os pensadores da TCT,

o conhecimento especializado é julgado uniforme e independente das línguas e culturas. Considera-se igualmente ingênua a concepção de homogeneidade do conhecimento científico e a crença na possibilidade de reconstrução lógica da linguagem.

A TCT tem seus fundamentos assentados na teoria do conhecimento, na teoria da comunicação e numa teoria da linguagem (BARROS, 2004).

Epistemologicamente, a teoria clássica fundamenta-se no princípio da dissociação entre pensamento e linguagem e prioriza a onomasiologia como procedimento básico de identificação dos termos, isto é, “identificadas as noções ou aspectos conceituais, buscam-se os termos, entendidos como etiquetas denominativas desses conceitos” (KRIEGER; FINATTO, 2004, p. 55). A visão linguística não acolhe a dissociação denominação/conceito, pois considera a denominação (ainda que provisória) como condição para o acesso ao conteúdo, isto é, “o acesso ao plano do conteúdo se faz via componente linguístico [...] tal como se pode exemplificar com o surgimento de termos compostos<sup>24</sup>, de caráter analítico, que acabam se transformando em unidades simples, como a passagem da *estrada de ferro* para *ferrovia* [...] (id., ibid.)”. Em outros termos, não há como formar um determinado conceito sem denominá-lo.

Tálamo e Lara (2006, p. 210) colaboram na sustentação da crítica fundamental que separa a TGT da TCT:

A Teoria Geral da Terminologia (TGT) é tributária da linha cognitiva: concebe a formação do conceito como algo independente da língua e vê no signo linguístico apenas a parte significante, atribuindo-lhe uma função meramente instrumental a serviço da denominação. A Teoria Comunicativa da Terminologia, bem mais recente, faz críticas ao reducionismo da TGT (embora não negue sua coerência interna), propondo observar o fenômeno terminológico no interior da linguagem a partir de seu caráter comunicativo. Já a Socioterminologia ressalta o uso social dos termos, rejeitando a excessiva ênfase dada à normalização pela TGT.

Em busca de uma síntese sobre quais os princípios orientadores da perspectiva comunicativa da terminologia pode-se destacar:

---

<sup>24</sup> Melhor seria usar termos complexos, como se verá mais adiante (p. 79).

- a) a admissão da variação conceitual e denominativa das unidades terminológicas (KOSTINA, 2003; KRIEGER; FINATTO, 2004);
- b) a poliedricidade de termos: o conceito é percebido a partir de pontos de vista determinados por: grupo científico, conceitualização que uma língua faz da realidade, idéia prioritária que se quer destacar no tema etc. (KOSTINA, 2003). Dessa forma, os termos são compreendidos nas perspectivas linguística, cognitiva e comunicativa, ou seja, além de uma teoria da língua, incorpora uma teoria do conhecimento e uma teoria da comunicação (KRIEGER; FINATTO, 2004);
- c) ativação do valor especializado dos termos num contexto e situação profissional (ativado pelo uso): função pragmática (KOSTINA, 2003).
- d) singularização do conteúdo dos termos em função dos aspectos relevantes de uma situação comunicativa. Sendo um signo linguístico a unidade terminológica é uma unidade de conteúdo e forma e ambos ocorrem simultaneamente (KOSTINA, 2003);
- e) polivalência temática dos termos: um mesmo termo poderá ser utilizado em mais de um âmbito de especialidade com valor idêntico ou distinto (KOSTINA, 2003);
- f) integração entre discurso especializado e língua natural: rejeição da ideia de que o discurso especializado possa formar um sistema a parte, uma sublinguagem diferenciada (KRIEGER; FINATTO, 2004). Em princípio as unidades de significação especializada são unidades que fazem parte do léxico do falante e “adquirem valor especializado de acordo com seu uso em um determinado âmbito especializado” (*op. cit.*, p. 203).

Cabré (1999), cuja concepção é da teoria comunicativa, como se disse, entende a terminologia como uma matéria autônoma de caráter interdisciplinar, que configura sua própria especialidade selecionando elementos das matérias de que é devedora e construindo seu próprio âmbito científico. Nas suas palavras a terminologia é

uma *interdisciplina* constituída por elementos procedentes da base da linguística, da ontologia e das especialidades, ligada necessariamente à documentação, da qual se serve e a qual serve, e usuária, e ao mesmo tempo contribuinte, das novas tecnologias da informação (CABRÉ, 1999, p. 22, tradução nossa).

Cada uma dessas disciplinas apropria-se da unidade terminológica conforme seu domínio. Para a linguística os termos são maneiras de saber, formam conjuntos de signos linguísticos. A ontologia, subdomínio da filosofia, entende o termo como uma forma de conhecer. Neste caso o termo é tanto uma unidade de conhecimento, que dá acesso ao mundo, como unidade de representação, que indica como os falantes percebem o mundo organizado.

Esta noção de representação não se dá como em Descartes, separando *ser* e *mundo*, *sujeito* e *objeto*, mas como em Heidegger, que os considera inseparáveis: o *ser-no-mundo* (*dasein*); *ser* e *mundo* constituindo-se por simultaneidade e por referência ao outro. Na visão de Heidegger, explica Stein (2000, p. 7) o homem é “um ser que se projeta para fora de si mesmo, mas jamais pode sair das fronteiras do mundo em que se encontra submerso. Trata-se de uma projeção no mundo, do mundo e com o mundo, de tal forma que o eu e o mundo são totalmente inseparáveis”. Para as diferentes disciplinas técnico-científicas o termo é principalmente unidade de expressão e de comunicação, assume valor de intercâmbio.

A materialidade da língua também se torna, pelas razões apontadas acima, um importante aspecto a ser discutido no domínio da terminologia, já que esta disciplina não se ocupa apenas do plano conceitual ou das relações lógicas e ontológicas entre termos, mas também de sua expressão e das denominações, além de oferecer metodologias que possam subsidiar as comunicações necessárias entre as ciências e entre os cientistas.

A linguística parte da linguagem, por meio de uma mostra idealizada própria de um falante também idealizado para poder dar conta, explicar a competência; a terminologia, por sua vez, parte de uma linguagem real (para tanto, toma os dados da documentação) para poder dar conta da denominação especializada (CABRÉ, 1999).

A complexidade está, então, em compreender o tênue limite entre padronização, como parâmetro necessário para a organização, e flexibilidade de linguagem, como parâmetro necessário para a expressão. Cabré (1995), entre outros, propõe a dissociação entre terminologia e imposição de padrões, objetando que a própria concepção da área já é representativa de diversidade, expressa nas diferentes possibilidades de uso e nas diferentes disciplinas que a compõem. De qualquer modo, mesmo que impraticável por seu caráter coercitivo, é preciso discutir a padronização como elemento que permite ampliar as possibilidades de circulação da informação.

Neste ponto é muito útil, e merece transcrição pelo que tem de lúcida, a distinção entre normalização, recomendação e harmonização realizada por Barros (2004, p. 87-88):

A normalização se dá com base em medidas coercitivas, adotadas por uma autoridade política ou de outra natureza e, normalmente, é fruto de um contexto sociolinguístico particular. [...]

A perspectiva da recomendação é outra e significa que um termo deve ser empregado preferencialmente em relação a outros sinônimos [...] Um termo recomendado poderá eventualmente ser normalizado se ele conseguir eliminar seus concorrentes. [...]

Por sua vez, a harmonização constitui o resultado de um acordo estabelecido sobre o uso de conjuntos terminológicos empregados em um dado domínio. [...]

É importante lembrar que a linguagem documentária possui, na condição de instrumento de comutação de uma linguagem natural ou linguagem de especialidade para uma linguagem documentária normalizada (LARA, 1993), função normalizadora.

Apesar das ressalvas, que claramente rejeitam a imposição de padrões, ainda se discute padronização por conta da profusão de variantes terminológicas e de sua interferência no processo de comunicação nas línguas de especialidades. Além das variantes terminológicas linguísticas, as variantes terminológicas de registro interferem igualmente no processo. Em termos operacionais a documentação, dependente do controle terminológico como uma das instâncias da organização de informações documentárias, não pode prescindir da eleição de uma determinada unidade terminológica como a mais

representativa (os critérios variam, mas geralmente são concernentes à sua aceitabilidade no âmbito das línguas de especialidade). Por outro lado, a variação interessa igualmente ao campo, porquanto pode promover outros diferentes pontos de acesso à informação, enriquecendo o sistema. No caso das linguagens documentárias, a variação é integrada por meio de uma rede de equivalências.

As variantes terminológicas de registro, que preocupam mais a terminologia do que as variantes terminológicas linguísticas, são classificadas em: a) geográfica – “expressões para designar um mesmo conceito utilizado por falantes de uma mesma língua em regiões diferentes”; b) de discurso – “expressões utilizadas para um mesmo conceito nos diferentes níveis de discurso, nível científico, técnico ou de divulgação científica” e c) temporal – “designações para um mesmo conceito que concorrem durante um período de tempo simultaneamente, firmando-se a preferência por uma em substituição a anterior e deixando-a em desuso” (VAN DER LAAN; FERREIRA, 2000).

Aubert (2001, p. 14) insere na problemática da padronização a neologia e as compara para, então, distingui-las:

A padronização e a neologia são sempre possíveis, mas a neologia aparecerá frequentemente como um derivativo não planejado, ou seja, como algo nascido de uma ‘geração espontânea’, como algo não previsto no quadro de um esforço padronizador, configurando-se, antes, em desvio ou oposição a tal esforço. A padronização se faz, é bem verdade, necessária, pois a homonímia e a parassinonímia podem fugir do controle social e frustrar a interinteligibilidade dos usuários, tanto em situação monolíngue quanto, com maior razão, em situação bilíngue (tradução técnica e contextos similares). Mas tal esforço enfrentará sempre a necessidade de renovação periódica de adequação às cambiantes realidades sociais, econômicas, antropológicas e políticas e aos avanços científicos e tecnológicos.

A opção pela terminologia como metodologia para a normalização das línguas de especialidades relaciona-se também com a distinção que realiza entre a palavra e o termo. Trata-se de uma diferenciação problemática, pela ausência de definições esclarecedoras, pois, como já observara Cintra (1983, p. 9), “não há entre linguistas grande preocupação em conceituar palavra e muito menos termo, porquanto são denominações fortemente ligadas

à língua escrita, código secundário para a Lingüística que trabalha prioritariamente com a língua falada”.

Le Guern<sup>25</sup> (*apud* LARA, 1993) fala em palavra na língua (léxico) e palavra no discurso (em funcionamento) e aponta a referencialidade como fator que as distingue, fundamentalmente. Na língua as palavras exprimem apenas um conjunto de propriedades sem relação direta com o concreto, possuem significado, mas não referência. Trata-se da palavra em estado dicionário, como se denomina comumente. À *palavra* no discurso denomina-se, mais precisamente, termo. O termo vincula as palavras a determinados contextos o que lhe permite aproximar-se com mais precisão da função referencial.

Cabe, ainda, acrescentar ao quadro o *descriptor* conforme compreendido pela linguística documentária. O termo, embora empregado em textos de especialidade, configura-se, todavia, como elemento que guarda proximidade com a linguagem natural (admite, por exemplo, variação e sinonímia). O descriptor, ao contrário, integra o espectro das linguagens artificiais, construídas, neste caso, para fins específicos de gestão de sistemas de informação documentária.

Dependendo de sua estrutura, o termo pode constituir-se de um único lexema ou de uma sequência lexemática. O *termo* pode ser *simples*, definido pela norma ISO 1087 como “constituído de um só radical, com ou sem afixos”, como, por exemplo, martelo, serrote etc., ou *complexo*, isto é, “constituído de dois ou mais radicais, aos quais se podem acrescentar outros elementos”, como, por exemplo, *evasão de receitas*, *evidência contábil*, *fundo de reposição do ativo* (BARROS, 2004, p. 100). Na Documentação, a norma ISO 2788, que estabelece algumas diretrizes para a criação de tesauros, traz a expressão *termos compostos* em vez de *termos complexos*. Considerando-se a indivisibilidade do conceito, ou seja, admitindo-se que não existem conceitos compostos e que a relação que os conceitos guardam com os termos é de referencialidade, é inadequado falar em termo composto. O

---

<sup>25</sup> LE GUERN, M. Sur les relations entre terminologie et lexique. *Meta*, v. 34, n. 3, p. 340-343, 1989.

*termo composto*, conforme compreendido na norma, é um único termo que contém mais de uma palavra.

O termo e a palavra diferem também com relação à categoria gramatical de seu objeto, na lexicologia substantivos, verbos, adjetivos, advérbios, pronomes, preposições e conjunções, na terminologia a presença de nomes é quase exclusiva, a de verbos, adjetivos e locuções escassa (CABRÉ, 1999).

Antes da visão comunicativa da terminologia este campo de pesquisa era configurado como concernente às línguas de especialidade (tecnoletos), as quais formariam um subsistema dentro da língua geral. Hoje adota-se a compreensão de que a linguagem de especialidade (a língua em uso) forma o seu objeto e que esta não se configura como um subsistema da língua geral, pois as “particularidades sintáticas, pragmáticas, semióticas, além de terminológicas” do discurso especializado não deixam de ser “recursos linguísticos utilizados pela língua geral na qual são escritos esses textos” (BARROS, 2004, p. 43). Dito de outro modo, não existem termos *a priori*, mas unidades lexicais que podem adquirir estatuto terminológico no âmbito das comunicações especializadas.

A unidade terminológica pode ser estudada do ponto de vista de sua expressão ou de seu conteúdo. A do conteúdo semântico de um termo (análise conceitual), nas palavras de Boutin-Quesnel et al.<sup>26</sup> (*apud* BARROS, 2004, p. 106), “determina as características de um conceito, de sua compreensão, de sua extensão e das relações que o mesmo mantém com outros conceitos”.

Sistema de conceitos é o “conjunto estruturado de conceitos construído com base nas relações estabelecidas entre esses conceitos e no qual cada conceito é determinado por sua posição nesse conjunto”, conforme norma ISO 1087 (BARROS, 2004, p. 111-112). Esta mesma autora ainda destaca na mesma obra (p. 106) a confusão terminológica sobre o emprego dos termos *noção* e *conceito*, cuja origem se deve ao uso que a Norma ISO 1087 faz

---

<sup>26</sup> BOUTIN-QUESNEL, R. et al. **Vocabulaire systématique de la terminologie**. Québec: Publications du Québec, 1985.

do termo *notion*, em francês, como equivalente de *concept*, em inglês, o que acabou confundindo muitos terminólogos. Há uma corrente de terminólogos brasileiros para quem o termo *noção* transmite a ideia de algo vago, impreciso, por isso preferência pelo uso do termo *conceito*.

Antes que se estabeleça o sistema conceitual é preciso definir o *campo conceitual*, compreendido por Boutin-Quesnel<sup>27</sup> (*apud* BARROS, 2004, p. 109) como um “conjunto de conceitos ligados entre si e que podem ser agrupados em torno de um conceito-chave”. O campo conceitual é, como se vê, um subconjunto do sistema de conceitos.

A elaboração do sistema de conceitos irá influir de diversas formas e em diversas fases da elaboração dos produtos terminológicos e igualmente no desenvolvimento de ontologias: na organização da taxonomia, no tratamento dos dados terminológicos, no sistema de remissivas, nas relações que serão estabelecidas entre os conceitos na identificação dos axiomas.

---

<sup>27</sup> BOUTIN-QUESNEL, R. et al. **Vocabulaire systématique de la terminologie**. Québec: Publications du Québec, 1985.

## 5 LINGUÍSTICA DOCUMENTÁRIA

A filosofia da linguagem, na síntese de Chauí (2000, p. 185), tem como grandes preocupações as seguintes questões: “As palavras realmente dizem as coisas tais como são? Descrevem e explicam verdadeiramente a realidade?”. São questões que interessam também à ciência da informação justamente porque discutem a forma como é condicionada a compreensão. Exatamente por isso ela deve ter alguma utilidade para a compreensão dos processos de representação e recuperação da informação (BLAIR, 2003).

Admitir que os objetos constituem-se na linguagem implica admitir, ensina Echeverría (2003), que o objeto trará, sempre, a marca humana e revelará, sempre, algo sobre o *sujeito*, sempre um *ponto de vista*. Echeverría (2003) insiste: não existe outro caminho que não seja o da linguagem, fora da linguagem não existe um lugar em que se possa apoiar-se. Os seres humanos vivem em um mundo linguístico.

As informações documentárias, produzidas por linguagens documentárias, também não escapam a essa mediação. A recuperação da informação é fundamentalmente um processo linguístico. Na elaboração da estratégia de busca é preciso descrever aquilo que se quer e adequar essa descrição à descrição da informação que está disponível (BLAIR, 2003).

Um fato relativamente consensual em relação às negociações de sentido que realizam usuários e sistemas de informação documentária é que esta negociação é mediada por um profissional da informação, direta – no trato com o usuário – ou indiretamente – via construção e uso de linguagens documentárias. Os problemas originados desta negociação no que se refere à construção e uso das linguagens documentárias, de natureza linguística, exigem novos olhares para a questão. García Gutiérrez cunhou a expressão linguística

documentária para representar essa nova subárea da ciência da informação para a qual Tálamo e Lara (2006, p. 204) propõem como escopo o estudo “das estruturas simbólicas da documentação e das questões linguísticas advindas da mediação necessária entre os produtores e os consumidores de informação”.

A linguística documentária insere-se no atual contexto que tem como base tecnológica as chamadas tecnologias de informação e comunicação, mas não se sustenta apenas nelas ou mesmo apenas delas. As tecnologias de informação e comunicação permitem ampliar e aperfeiçoar alguns procedimentos metodológicos, mas a questão central, o problema maior, ainda é referente à filosofia da linguagem. Mais do que a afirmação de um paradigma tecnológico, sustentam Tálamo e Lara (2006, p. 204) encara-se “um modelo complexo que combina três vértices: a tecnologia, o meio social e a capacidade de processamento humano do conhecimento e da informação”.

Sustentar os estudos contemporâneos sobre o fluxo da informação nas tecnologias remete ao que Rodrigues (1994, p. 198) chama de naturalização da tecnicidade, a qual corresponde a uma das mudanças mais significativas da modernidade, “convertendo-a numa modernidade ‘tecno-lógica’, no sentido próprio deste termo, tal como outros períodos foram ‘mito-lógicos’ e ‘teo-lógicos’”. No primeiro caso, espera-se dos dispositivos técnicos a solução para os problemas, “tal como no passado se esperava a salvação da benevolência dos deuses ou da mediação da Igreja” (RODRIGUES, 1994, p. 18). As tecnologias de informação inserem-se no domínio dos *dispositivos técnicos* e diferenciam-se dos *utensílios* e dos *instrumentos*. Os primeiros incidem diretamente sobre a linguagem, os segundos são utilizados no domínio da produção, os últimos são usados para tornar a percepção sensorial mais precisa (RODRIGUES, 1994).

Considerando-se a proximidade conceitual, desponta como interessante verificar que relações guardam a análise documentária e a linguística documentária. A análise documentária, termo cuja inscrição no campo da informação deve muito a Jean Claude Gardin (LARA, 2009), preocupa-se com a identificação (análise, propriamente dita) e expressão (tradução) dos significados presentes num texto. Ocorre que a expressão dos

significados neste caso realiza-se com o apoio de uma linguagem documentária produzida em outra instância, em outro processo. A expressão do significado, então, se dá pela referência a algo que é *extraído* do texto e designado por signos (uma linguagem) que não é necessariamente a mesma usada no texto de origem. Ainda, portanto, que a análise documentária faça uso de linguagens documentárias, o desenvolvimento destas linguagens não se dá exatamente em seu interior. Em comum tem-se que ambas dependem dos conhecimentos produzidos na linguística. Em García-Gutiérrez (1990) tanto a análise documentária quanto a linguística documentária, entendidas como campos diferentes, possuem interface com a linguística. As relações entre a análise documentária e a linguística também já haviam sido apontadas por Cintra (1983).

Tálamo e Lara (2006) e Lara (2009) ampliam o escopo da linguística documentária para além da documentação *stricto sensu* (na qual são produzidas as linguagens documentárias), e integram também o espaço onde se dão suas interfaces evidenciadas pelas tecnologias da informação. Desse modo a aplicação do que se propõe para a linguística documentária alcança não só os ambientes documentários tradicionalmente concebidos, mas também aqueles que, por tratarem a informação em algum nível, podem utilizar referenciais das práticas documentárias convencionais, como é o caso das ontologias.

Os textos com os quais os sistemas documentários trabalham são submetidos a dois tipos básicos de tratamento: a condensação e a representação (classificação ou indexação). A condensação toma como base a superestrutura textual para desestruturação e reestruturação do texto original num texto de menor extensão que permita sua potencial reorganização. Este produto é denominado resumo e compartilha a mesma linguagem com o texto original. Na produção de informações documentárias utiliza-se um código comutador denominado linguagem documentária (CINTRA et al., 2001).

## 5.1 LINGUAGEM DOCUMENTÁRIA

Tratando-se de linguagem construída, a linguagem documentária é limitada em relação à linguagem natural e possui, conforme apontam Tálamo e Lara (2006), apenas duas de suas funções características: a referencial e a metalinguística. A linguagem documentária caracteriza-se também por possuir apenas uma articulação (não se fala uma linguagem documentária) e uma realização precária no eixo sintagmático. As línguas naturais, entre todos os sistemas semióticos, ocupam posição hierárquica que pode ser considerada superior porque constituem a única realidade imediata para o pensamento e o único código capaz de traduzir com a máxima eficiência e adequação qualquer outro sistema semiótico, sendo que o inverso não é verdadeiro (LOPES, 2005).

Como instrumento de representação documentária as linguagens documentárias caracterizam-se pela identificação de conceitos e por sua expressão terminológica. Tratar de conceitos e termos solicita de antemão o referencial da terminologia como campo de estudos e como objeto concreto. Dentre as várias linhas da terminologia, a teoria do conceito, formulada por Dahlberg, é bastante referenciada pela ciência da informação.

A teoria do conceito está expressa em um artigo (já clássico) de Ingetraut Dahlberg, publicado em 1978 no Brasil, no qual expõe alguns aspectos importantes sobre tipologia de conceitos (gerais, especiais e individuais) e sobre categorias. A intensão e a extensão dos conceitos também são tratadas a partir da identificação de suas características (essenciais, acidentais e individualizantes). A autora discute ainda a importância das definições (nominais e reais) para que se possam fixar os conceitos em termos de sua precisão de significado.

Na formalização dos conceitos reside um dos maiores problemas propostos à construção e uso de informações documentárias e que também tem reflexos na produção de ontologias: a definição limita (precisa limitar) o entendimento do conceito. Uma

definição, conforme a ISO 704, precisa definir o conceito como unidade com uma única intensão e extensão. Esta limitação ocorre principalmente em função de necessidades institucionais. Ou seja, um mesmo conceito pode assumir diferentes acepções, apropriações, em função da variedade de usos e contextos nos quais se insere. Em busca da precisão terminológica a informação documentária delimita o universo da interpretação. A noção de conteúdo, portanto, na atividade documentária, relaciona-se forte e pragmaticamente com os objetivos funcionais do sistema documentário. O termo deve ser aqui entendido como uma unidade terminológica determinada que assume um conteúdo específico dentro de um domínio específico, conforme Barros (2004).

Falar em domínios específicos não significa necessariamente falar em fixidez ou determinação. Domínios específicos são marcados por variantes culturais, as quais têm preocupado a socioterminologia, disciplina que visa a identificar e categorizar as variantes linguísticas e seus diferentes tipos de uso. Este tipo de conhecimento possui grande aplicação na criação de ontologias para fins de interoperabilidade (cf. cap. 3.3).

Ainda que se considerem as dificuldades inerentes à aplicação da metodologia proposta pela socioterminologia no âmbito da linguística documentária, não se pode desconsiderar o meio social como integrante do processo, pois tal postura descaracterizaria a linguagem documentária como interface. Entre outras qualidades, a linguagem documentária desempenha esse papel funcionando como código inteligível e como fonte para interpretação de sentido; caracterizando-se como metalinguagem e incorporando o usuário como integrante do processo (LARA, 2004) A informação documentária, pode-se depreender, é uma construção. Como reflexo desta concepção é que se tem buscado modelos de representação mais flexíveis, como é, por exemplo, o caso do tesouro em relação aos sistemas de classificação decimal e das ontologias em relação ao tesouro.

Conforme apontado em outro texto (MOREIRA, 2005), considera-se como uma alternativa muito promissora para a comunicação científica a perspectiva de bases de arquivos abertos (*open archives*) bem como outras formas de comunicação virtual. Do ponto de vista da representação e recuperação da informação, entretanto, essa nova via gera

novos e desafiantes problemas. Nesse cenário a linguagem documentária precisa aproximar-se mais da fluência da linguagem natural (o que não equivale a utilizá-la) e preservar a precisão da terminologia (o que equivale a ajustá-la). A adoção de meios semiautomatizados de representação, com apoio das ontologias e da terminologia, parece apontar uma direção. Mapeada a superestrutura (como, por exemplo, a clássica: introdução, material e método, resultados e discussão, denominada IMRAD, muito utilizada em artigos no campo das ciências da saúde) e estabelecidas as regras de inferência é possível, como demonstram Marcondes (2005) e Marcondes *et al.* (2007) publicar não apenas texto, mas também conhecimento, visando além dos seres humanos, os agentes de *softwares* inteligentes.

As condições necessárias para que a comunicação da informação se realize são facilitadas quando é possível compartilhar o entendimento dos termos e conceitos e isso só é possível a partir de um acordo *termino-ontológico*, ou seja, tomando-se como base a compreensão de que as ontologias prestam-se à organização efetiva do conhecimento, mas dependem de uma interface terminológica para os ajustes das representações (termos) e que, neste sentido, a terminologia pode prestar-lhe suporte teórico-metodológico. Adota-se neste ponto a expressão *termino-ontológico* para enfatizar que a abordagem que esta reflexão dedica-se às ontologias deve ser compreendida sob o prisma da base metodológica e epistemológica da terminologia, como o faz a linguística documentária.

## 5.2 LINGUÍSTICA DOCUMENTÁRIA E TERMINOLOGIA

A partir da constatação de que a terminologia parte da observação dos termos dentro de um domínio para ali qualificá-los, considerando principalmente o contexto sócio-cultural e institucional como determinante da carga semântica dos termos, propõe-se, então, um diálogo mais próximo entre a terminologia e as ontologias.

Enquanto a ontologia filosófica preocupa-se com o caráter lógico-epistemológico de formação dos conceitos antes de preocupar-se com sua expressão formal, as ontologias computacionais, por sua vez, tem na formalização uma das suas características essenciais. É precisamente a adoção da linguagem formal que permite, num nível de máquina, diminuir o *gap semântico*<sup>28</sup> existente em qualquer linguagem ou forma de representação. Além das regras de inferências, relacionamentos e definições legíveis por máquina, as ontologias também contemplam um conjunto de termos e suas inter-relações, ou seja, uma terminologia.

Observe-se que se fala em diminuição do *gap semântico*, mas não em sua eliminação, pois se esta impossibilidade existe na linguagem humana, que se dirá no nível da linguagem de máquina, que vai depender mais fortemente da intermediação das representações formais. Neste caso, a impossibilidade está na diferença entre o conhecimento que se pode formular com a linguagem natural e sua representação numa linguagem artificial como a linguagem de programação, por exemplo, já que a semântica de um objeto é sempre determinada por seu contexto.

O uso de ontologias pretende fornecer uma descrição exata (ou o mais próximo possível disso) do conhecimento, por isso fixa-o em um conjunto de axiomas e inferências explicitamente formuladas, pois não pode, de certa forma, prescindir disso, e deixa menos espaço para o *gap semântico* verificado com mais frequência na linguagem natural.

O entendimento da unidade terminológica como unidade de conhecimento leva à compreensão das ontologias como aporte necessário à operação de representação documentária. A terminologia depende das línguas de especialidades para se fazer concretamente, mas, pergunta-se, em que nível exatamente se pode localizá-las já que para cada especialidade existem subespecialidades assim como superespecialidades? Determinar o ponto de origem é questão puramente ontológica. A identificação de ontologias comuns às

---

<sup>28</sup> O *gap semântico* refere-se à distância que existe entre o problema no mundo real (domínio do problema) e o modelo abstrato construído (domínio da solução).

terminologias coloca-se como condição para a efetiva produção, representação e recuperação de informação documentária, bem como das garantias mínimas de interoperabilidade.

Nos contextos técnico-científicos a terminologia assume papel fundamental pelo fato de encontrarem-se exatamente aí as línguas de especialidades, responsáveis por garantir a comunicação rápida e precisa entre os interlocutores (pesquisadores, profissionais ou estudantes). Quanto maior o acordo relativo (explícito ou implícito) ao sentido que se atribui aos conceitos em negociação, maior a eficácia na comunicação. Em outras palavras, quanto mais distantes os interlocutores do ideal (inatingível) de univocidade de referência, maiores as perdas causadas à eficácia da comunicação. Em campos científicos nos quais a precisão terminológica é condição, como na comunicação científica, na qual se origina e se justifica, isso se agrava. Julga-se necessário, portanto, discutir a comunicação científica em termos de sua estruturação *termino-ontológica* para minimizar as ambiguidades neste tipo de comunicação, consideradas em seus aspectos de difusão ou de divulgação.

Compactua-se com a demarcação menos nítida a respeito das fronteiras que envolvem pesquisadores (cientistas) e público leigo, conforme compreendidas por Santos (2002, 2003) e Vogt (2008). Este último, aliás, prefere, por conta de sua maior abrangência, o termo *cultura científica* aos termos *difusão* ou *divulgação*. À divulgação científica, ensina Vogt (2008), não cabe apenas informar ao leigo os avanços da ciência, mas também criar condições para a “formação crítica do cidadão em relação à ciência”; isso, segue Vogt, “vai além da atitude inicial, na qual o cientista era o sábio, o cidadão era o ignorante e o jornalista científico ou divulgador da ciência era o construtor da ponte entre essas figuras”.

Há um tênue limite entre o que se convencionou chamar de ciência e o que, deste ponto de vista, chamou-se senso comum (SANTOS, 2002). Em tom provocativo, Santos julga necessário perguntar pelo contributo prático, efetivo que o conhecimento científico acumulado oferece para a felicidade do homem. A euforia causada pela ciência está, em parte, acredita, superada e isto torna o momento adequado para a visão pragmática que propõe. Nas palavras do autor:

[...] Depois da euforia cientista do século XIX e da consequente aversão á reflexão filosófica, bem simbolizada pelo positivismo, chegamos a finais do século XX possuídos pelo desejo quase desesperado de complementarmos o conhecimento das coisas com o conhecimento do conhecimento das coisas, isto é, com o conhecimento de nós próprios. [...] (SANTOS, 2002, p. 30).

Na medida em que a produção de ciência e de tecnologias cresce e que suas fronteiras deixam de ser nítidas, aumenta em maior proporção a demanda por informação com valor agregado, que possa ser recuperada de forma eficaz. Num mundo marcado pelas tecnologias de informação, mais acentuadamente pelos avanços da telemática, não se pode atender a essa demanda sem que se pense, conforme colocado anteriormente, em rede, em compartilhamento de recursos e de informação.

Na mesma proporção em que cresce a demanda, cresce a sua diversidade. Apenas para efeito de visualização, pode-se tomar como exemplo o caso da pós-graduação *stricto sensu* no Brasil. Conforme dados da Fundação de Amparo à Pesquisa do Estado de São Paulo – FAPESP (BICUDO; TAMIRI; SAMPAIO, 2002), entre 1989 e 1998 o número de mestres titulados mais que triplicou e o número de doutores quase quadruplicou em todo o país. Em 1998 foram formados 12.510 mestres (2.090 na Universidade de São Paulo – USP) e 3.945 doutores (1.465 na USP). Dadas a diversidade de público e a complexidade *endêmica* das especializações, como se pode garantir a referência contextual nas representações?

A questão das múltiplas representações tendo em vista as diferentes demandas de informação não é nova, ao contrário, trata-se de um dos fundamentos da análise documentária, conforme indicado por Cintra et al. (2001). O que ainda precisa ser discutido, o diálogo que ainda deve ser estabelecido é o da interoperabilidade entre as múltiplas representações, ou seja, interessa saber como essas múltiplas representações podem dialogar; para isso é preciso identificar seus pontos de contato.

O desafio que se coloca para a linguística documentária, portanto, diz respeito à construção de linguagens documentárias mais flexíveis, que possam acolher as diferentes formas de expressão dos usuários dos sistemas de informação documentária, bem como suas diferentes formas de representação.

A ciência da informação tem se valido, para representar e organizar informações, de instrumentos de controle terminológico, denominados de modo genérico por tesauros. A função do tesouro num sistema de informação documentária é subsidiar o processo de representação e consequente localização da informação documentária, seja por parte do sistema, ou por parte do pesquisador. Trata-se de uma modalidade de linguagem documentária que revela estruturação lógico-semântica entre os conceitos de um determinado domínio do conhecimento, sustentada pela norma de elaboração de tesauros, ISO 2788, e pela terminologia, oferecendo referenciais metodológicos.

A interseção (ou a imperiosa necessidade de seu efetivo reconhecimento) entre a ciência da informação, a terminologia e a ontologia já havia sido apontada na obra de Wüster, considerada, como se disse, o marco no aprofundamento das questões teóricas em terminologia. Considerando a terminologia em seu aspecto multidisciplinar, Wüster localiza-a na convergência da linguística, da lógica, da ontologia, das ciências da informação e das diferentes áreas do conhecimento científico. Pode-se acrescentar hoje a essa lista campos como a ciência da computação, a inteligência artificial e a engenharia do conhecimento, entre outros. Num mundo que se pretende governado pelas tecnologias de informação e comunicação, é fácil compreender o compartilhamento do domínio terminológico entre os *especialistas* e o público em geral, os usuários indiretos da terminologia, dentre os quais, aponta Krieger (2001b, p. 216), estão “os tradutores, intérpretes, documentalistas, os redatores técnicos, lexicógrafos e terminógrafos, estudantes universitários, entre outras categorias de profissionais que se envolvem com a linguagem”.

Moreira (2007) analisou as interconexões entre a lexicologia, a terminologia e a ciência da informação com os estudos sobre ontologias, expressas nas palavras-chave mais utilizadas para representação de assuntos em artigos brasileiros da grande área da ciência da informação. Na Tabela 3 apresenta-se uma matriz que revela a quantidade de co-ocorrências das palavras-chave mais utilizadas no *corpus* de análise. Nesse artigo, foi possível perceber que a terminologia concentra o maior número de combinações, revelando ser este o ponto de contato entre as disciplinas envolvidas neste estudo. Revela-se como

sintomático na matriz a ausência do termo *representação* já que este tema constitui-se numa das preocupações centrais da ciência da informação, notadamente nas interfaces com a terminologia e com as ontologias. A presença tímida da linguagem documentária (um instrumento da representação documentária) na matriz também não significa diálogo, pois aparece associada apenas às palavras-chave que lhe são mais evidentes, de forma quase redundante: terminologia e ciência da informação.

Tabela 3 - Co-ocorrência das palavras-chave mais utilizadas

	Ciência da informação	Linguagem documentária	Ontologia	Organização do conhecimento	Recuperação da informação	Terminologia
Ciência da informação		1	2	-	-	3
Linguagem documentária	1		-	-	-	3
Ontologia	2	-		-	-	2
Organização do conhecimento	-	-	-		-	-
Recuperação da informação	-	-	-	-		1
Terminologia	3	3	2	-	1	

O que se observou de mais significativo na matriz, contudo, foi a incomunicabilidade da palavra-chave *organização do conhecimento*. Localizada em seis artigos apareceu em apenas um deles situada implicitamente no escopo da terminologia, nos outros formava quase uma linha independente. Isso se configura como problema, vale lembrar, porque o *corpus* documental foi constituído por periódicos especializados em ciência da informação.

Novamente é preciso lembrar que a interdisciplinaridade inerente aos escopos dos campos estudados impede uma classificação estanque e que interessa, neste caso, apenas vislumbrar como tem sido construído o referencial da ciência da informação no trato dessas questões.

## **6 AS ONTOLOGIAS E A CONSTRUÇÃO DE INFORMAÇÕES DOCUMENTÁRIAS**

Neste capítulo serão retomados alguns conceitos que estão na intersecção entre as terminologias e a ciência da informação, mais especificamente aqueles desenvolvidos pela linguística documentária, com a finalidade de formar breve um estofo teórico pelo qual se possa identificar e compreender as relações de empréstimo e de interdisciplinaridade que estes campos guardam com as ontologias.

### **6.1 SISTEMAS DE INFORMAÇÃO DOCUMENTÁRIA: ALGUNS ESCLARECIMENTOS TERMINOLÓGICOS**

Houve um tempo em que se definiam as bibliotecas (tomadas como modelos de sistemas de informação documentária) como o que a palavra significa etimologicamente: um espaço físico destinado à guarda de livros. Mais recentemente, num ambiente mais complexo, o conceito tornou-se menos pacífico e passou a conviver com outros que lhe tangenciam, complementando-o, expandindo-o ou mesmo substituindo-o em contextos determinados, tais como: sistemas de informação, sistemas de recuperação da informação e unidades de informação, entre outros.

A variação terminológica enseja, em certos casos, não apenas ajustes (o que já significa mudança), mas também evolução na precisão da relação entre o termo e o conceito com o qual se relaciona. Disso advém, aliás, a dificuldade de identificação e de aceitação da sinonímia no uso de descritores. Para que sejam considerados sinônimos perfeitos é preciso que as unidades terminológicas sejam intercambiáveis em todos os contextos.

Nesta pesquisa optou-se pelo uso da expressão *sistemas de informação documentária* a *sistemas de informação* ou *sistemas de recuperação da informação*, com o fim de preservar o conceito de possíveis ambiguidades ou contaminações. *Grosso modo*, uma tenda com alguém que forneça informações turísticas pode ser apontada como exemplo de sistema de informação, mas possui, certamente, informação de natureza diversa da que compõe os sistemas de informação documentária.

*Sistemas de informação* é um termo mais comumente adotado pela literatura do campo da administração e da ciência da computação, cujo sentido se pode depreender, por exemplo, dos conteúdos programáticos dos cursos nesses campos ou pela co-ocorrência das palavras (co-texto) por meio de uma rápida busca no *Google™*. Alguns autores incluem os sistemas de informação numa categoria mais ampla, que comporta também os sistemas de comunicação de massa, as redes de comunicação de dados etc., independentemente da forma ou conteúdo das mensagens (ARAÚJO, 1995).

A expressão *sistemas de recuperação da informação* traz duas questões problemáticas. Em primeiro lugar, conforme já apontado em outro trabalho (MOREIRA, 1998), evidencia a tomada do todo pela parte, pois a recuperação é apenas uma das operações do sistema, neste sentido a denominação *subsistema* seria melhor. Outra questão envolve a utilização genérica que se faz da palavra *informação*. Não é qualquer informação que interessa à ciência da informação. Diversos outros campos adotam a informação como objeto, como o jornalismo, por exemplo. No caso da ciência da informação opera-se com o registro como forma de garantir a dimensão tempo-espaco (permanência e transporte) da informação, disso decorre o conceito de documento.

De que, exatamente, são compostos os modernos sistemas de informação documentária? Trata-se, naturalmente, de uma pergunta retórica, não se tem a pretensão, nem o objetivo, de respondê-la neste texto, serve apenas como mote para a discussão conceitual sobre este termo e sobre os seus usos. Respondê-la satisfatoriamente implicaria uma discussão sobre o conceito de informação que, embora pertinente e necessária, não está entre os objetivos desta reflexão. Algo que é aceito com relativo consenso é que o documento e suas representações formam a base de seu acervo. Muda-se então o foco: o que é o documento e qual a natureza de suas representações?

Pedauque (2004) num trabalho que visa a esclarecer e atualizar o conceito de documento, em face das modernas condições tecnológicas, esclarece que o documento neste contexto deixa de ser apenas um objeto material estável e propõe distinguir o conceito de documento como forma, como signo e como meio. A divisão triádica é baseada na já conhecida concepção linguística: sintaxe, semântica e pragmática.

O *documento como forma* é representado pela equação *documento = suporte + inscrição*, refere-se a um objeto de comunicação regido por regras de formalização colocadas de forma mais ou menos explícitas, que materializam um contrato de leitura entre um produtor e um leitor. O suporte tradicional dominante é o papel e a inscrição é a escrita.

A equação *documento = inscrição + sentido* representa a concepção de *documento como signo*. Neste caso o documento é percebido como portador de sentido, o que não permite dissociar sujeito e contexto, e confere importância especial à interpretação; o documento pode assumir, dependendo do contexto geral, sentidos diferentes.

A visão de *documento como meio* insere o documento nas relações sociais. O termo meio deve ser tomado com cuidado, adverte Pedauque (2004), neste caso é usado com a acepção de algo que reagrupa todas as aproximações que analisam o documento como um fenômeno social, um elemento tangível de comunicação entre pessoas. *Documento = inscrição + legitimidade* é a equação representativa aqui. O signo só se torna documento se for registrado para uso futuro.

Considerado como meio, o documento condiciona a mensagem. Questiona-se, contudo, a concepção de documento como “elemento tangível de comunicação” em sistemas de informação documentária. Tal tangibilidade é condicionada ao fluxo e, como decorrência, às condições de acesso aos documentos. Os fluxos e os acessos hoje se dão, prioritariamente, por meio das representações; é preciso, então, compreender o documento em conjunto com sua representação. Em outras palavras, não se separam a concepção de documento como signo e como meio, já que tanto o documento ou a representação não são portadores de sentidos, mas de significados e não se separam mais conforme compreendia a documentação clássica.

As diferenças entre sentido e significado ainda não foram estabelecidas de forma consensual. Nesta reflexão, sentido é utilizado conforme a compreensão que lhe dá Vygotsky (1991): a soma de todos os eventos que a palavra desperta na consciência. O sentido da palavra é adquirido no contexto em que surge, obtido no uso, como resultado de “um todo complexo, fluido e dinâmico que tem várias zonas de estabilidade desigual” (VIGOTSKY, 1991, p. 125). Em contextos diferentes, o sentido é alterado. O significado refere-se à apenas uma das zonas do sentido e tende a permanecer estável ao longo das alterações do sentido.

Se é permitido, busca-se o auxílio da ficção esclarecedora de Saramago (1998, p. 78):

[...] Ao contrário do que em geral se crê, sentido e significado nunca foram a mesma coisa, o significado fica-se logo por aí, é directo, literal, explícito, fechado em si mesmo, unívoco, por assim dizer, ao passo que o sentido não é capaz de permanecer quieto, ferve de sentidos segundos, terceiros e quartos, de direcções irradiantes que se vão dividindo em ramos e ramilhos, até se perderem de vista, o sentido de cada palavra parece-se com uma estrela quando se põe a projectar marés vivas pelo espaço fora, ventos cósmicos, perturbações magnéticas, aflições.

Retomando-se uma das questões iniciais deste subcapítulo, recorre-se novamente a Pedauque (2004) para que se compreenda a noção de documento digital, colocada de forma muito escorregadia pelo autor: um documento digital é a marca das relações sociais reconstruídas por meio de dispositivos de computação. Na proposta de uma equação que represente o documento digital, o autor reduz a concepção para o que denomina

*documento web*, dessa forma, *documento web* = *publicação* + *acesso localizado*. De qualquer modo, representa um avanço em relação à concepção mais ingênua expressa pelo senso comum detectada por Levy e Marshall<sup>29</sup> (*apud* MOREIRA, 1998) e expressa por meio das seguintes equações: *papel* = *fixo* + *permanente* e *digital* = *fluido* + *transiente*<sup>30</sup>.

A arquitetura mais rizomática das redes informação contemporâneas não pode mais comportar, como acontecera outrora, a publicação ou o acesso como eventos estáveis. A diminuição de algumas barreiras tecnológicas impostas à publicação, que também diminuiu o poder do editor, além do fenômeno da autopublicação, são certamente responsáveis pelo rápido crescimento no número de documentos postos em circulação. A problemática do acesso localizado é, neste caso, inversamente proporcional: quanto maior o número de documentos disponíveis, maiores a necessidade e a dificuldade de exatidão na representação (identificação) e maior a necessidade de recursos nas estratégias de localização.

Dessa forma, coloca-se em discussão a eficácia do tesauro como instrumento no processo de localização dos documentos digitais, uma vez que não tenha sido desenhado para isso e que não contemple os *web services*<sup>31</sup> entre os *usuários* aos quais se destina. Da mesma forma, questiona-se a eficácia das ontologias como instrumento no processo de representação do conhecimento sem o conluio das teorias e metodologias desenvolvidas para o desenvolvimento de tesauros, voltados prioritariamente para o universo da comunicação científica e tecnológica.

---

<sup>29</sup> LEVY, David M.; MARSHALL, Catherine C. What color was George Washington's white horse?: a look at assumptions underlying digital libraries. DIGITAL LIBRARIES, 1., 1994, Texas (EUA). **Proceedings...** Texas: College Station, 1994. Disponível em: <<http://www.csd.tamu.edu/DL94/paper/levy.html>> Acesso em: 20 fev. 2010.

<sup>30</sup> Curiosamente, o próprio texto de Levy e Marshall, referenciado anteriormente, converteu-se numa prova de que nem tudo é fluido e transiente na *web*, pois ainda está disponível no mesmo endereço após dezesseis anos.

<sup>31</sup> “Um *web service* é um sistema de *software* identificado através de uma URI cujas interfaces públicas e interconexões são descritas em XML. Sua definição é publicada de modo a poder ser ‘descoberta’ por outros sistemas de *software*. *Web services* podem interagir com outros sistemas ou *web services* do modo prescrito em sua definição, utilizando mensagens baseadas no padrão XML produzidas através de protocolos de internet” (W3C *apud* BREITMAN, 2005, p. 142).

No âmbito dos sistemas de informação documentária, cujo fim precípua é a recuperação da informação por meio de uma representação do conhecimento, o tesauro representa um grande avanço em relação às rígidas classificações bibliográficas baseadas em classificações filosóficas, principalmente as anteriores à Ranganathan.

O tesauro é utilizado tanto na análise documentária quanto na elaboração de estratégias de busca. Por sua característica de linguagem dinâmica, admite atualizações constantes e incorporação de novos conceitos e relações, mas é, por força de definição, circunscrito a domínios específicos do conhecimento, isto é, não se pode construir um tesauro geral. Pode-se, contudo, construir ontologias de alto nível, que descrevem os conceitos mais gerais, cujo sentido é o mesmo ou quase o mesmo em todos os domínios. A função mais importante de uma ontologia de alto-nível é apoiar a interoperabilidade semântica entre outras ontologias. No próximo subcapítulo busca-se compreender um pouco melhor as relações de complementaridade e de oposição entre os tesouros e as ontologias.

## 6.2 TESAUROS E ONTOLOGIAS

Um documento que define formalmente as relações entre os termos. Essa foi a definição que Berners-Lee, Hendler e Lassila (2001, tradução nossa<sup>32</sup>), coletaram para o termo *ontologia* em seu artigo basal para quem quer compreender a *web* semântica. A definição, por sua vagueza, ajusta-se também tanto à taxonomia quanto ao tesauro. O termo *ontologia* foi simplificado ao ponto do comprometimento terminológico e, se não dá

---

<sup>32</sup> No original: “[...] Artificial-intelligence and Web researchers have co-opted the term for their own jargon, and for them an ontology is a document or file that formally defines the relations among terms. The most typical kind of ontology for the Web has a taxonomy and a set of inference rules”.

origem, colabora para as inúmeras confusões conceituais que geraram uma infinidade de textos cujo fim principal é dirimi-las<sup>33</sup>.

Os tesouros, como as ontologias, são desenvolvidos para responderem a campos conceituais mais específicos, mais demarcados e, por esta mesma razão, mais precisos. Na abordagem da terminologia clássica, tomando-se como referência o triângulo semântico clássico (Figura 7), pode-se afirmar que os tesouros ajustam-se mais ao lado esquerdo do triângulo, preocupando-se com a relação entre termos e conceitos. A abordagem comunicativa da terminologia, assim com as ontologias, por sua vez, são mais adequadas ao seu lado direito, às aplicações ou apropriações dos conceitos simulando domínios particulares do conhecimento por meio de inferências, e são mais afeitas, portanto, aos aspectos lógico-semântico-pragmáticos.

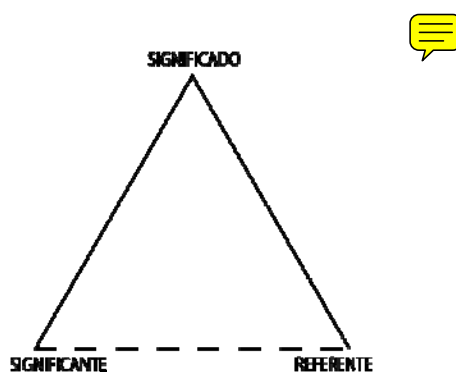


Figura 7 - Triângulo semântico

Pela proximidade de estrutura e de finalidade, é natural que surjam comparações entre tesouros e ontologias, busca-se, portanto, discuti-las com o fim de compreender melhor seus pontos de contato. É comum cair na tentação das comparações e julgar um tesouro melhor do que uma ontologia ou vice-versa, conforme o ponto de vista do observador e o lugar de onde fala, mas só se pode considerar um instrumento melhor do que o outro se eles são designados para resolver efetivamente o mesmo problema, o que

<sup>33</sup> Apenas para fins de ilustração, sem objetivos de análise, uma busca no *Google* realizada em 01 mar. 2010, pela expressão “tesouros e ontologias” (em língua portuguesa) resultou em 15.600 páginas.

não é o caso. Há, por exemplo, autores da ciência da computação (JASPER; USCHOLD<sup>34</sup>, 1999; FENSEL<sup>35</sup>, 2001 *apud* MOREIRA, 2003; FISCHER, 1998) que consideram os tesauros como tipos mais simples de ontologias, posição refutada pela maioria dos outros autores, tanto os oriundos da ciência da computação quanto aqueles que desenvolvem estudos no interior da ciência da informação (ALMEIDA; BAX, 2003; CAMPOS *et al.*, 2008; MARCONDES; CAMPOS, 2008).

Enquanto a produção de inferências lógicas é função explícita das ontologias, por meio de uma estrutura de relações, nos tesauros, estas inferências são menos explícitas ou menos formalizadas; são sugeridas.

Se na ciência da informação há ainda alguma confusão terminológica neste domínio, ela é ainda maior na ciência da computação (o que é compreensível, considerando-se o maior tempo de uso do termo), conforme demonstrado por Guarino e Giaretta (1995), Moreira (2003) e Garshol (2004).

Moreira (2003), em sua dissertação de mestrado, buscou esclarecer as diferenças entre tesauros e ontologias, questionando seu uso indistinto. Seus resultados apontam origens e propósitos distintos para os dois instrumentos. Enquanto o tesouro “nasceu como um instrumento prático para auxiliar na indexação e busca de documentos” (MOREIRA, 2003, p. 97), uma aplicação mais direcionada aos especialistas, as ontologias nasceram da “necessidade de descrever os objetos digitais e suas relações” (*loc. cit.*), uma aplicação mais direcionada aos procedimentos automatizados, às inferências computacionais por meio de agentes inteligentes. O ponto de contato entre os dois instrumentos refere-se ao fato de estarem relacionados com a descrição de alguma coisa.

Lancaster (2002) relaciona o sistema unitermo, criado por Mortimer Taube em 1951, com a aparição do tesouro, ainda que rudimentar. Sendo o conceito de tesouro um conceito

---

<sup>34</sup> JASPER R., USCHOLD, M. A framework for understanding and classifying ontology applications. In: INTERNATIONAL JOINT CONFERENCE ON ARTIFICIAL INTELLIGENCE, 1999, Stockholm (Sweden). **Proceedings...** Stockholm: [s.n.], 1999.

<sup>35</sup> FENSEL, D. OIL: an ontology infrastructure for the semantic web. **IEEE Intelligent Systems**, p. 38-45, mar./apr., 2001.

mais tradicional na ciência da informação (utilizado, pelo menos, desde a década de 50), é natural que apresente no seu interior uso mais claro do que na ciência da computação. O termo *ontologias*, conforme utilizado na ciência da computação, cuja popularização é mais recente, inclusive neste campo, ainda não foi definido com clareza. Aparentemente, observa Moreira (2003, p. 97), “tudo que modela um segmento da realidade pode ser denominado de ontologia, uma vez que é a palavra da moda. Neste caso, até mesmo os tesouros podem ser enquadrados como ontologias terminológicas”. Sua análise evidenciou, contudo, diferenças de propósitos entre os dois termos, conforme apontado anteriormente.

Naturalmente cada campo científico ou especialidade vê diferentemente um mesmo objeto, recorta-o conforme a epistemologia que lhe é própria. Por isso mesmo, por conta das diversas especializações é que se exige, num mundo tecnologizado e globalizado, cada vez mais o recurso da terminologia em busca da precisão nas comunicações. A comunicação técnico-científica ou as ações de cooperação entre sistemas de informação documentária, como qualquer outra forma de comunicação, dependem da negociação que entre si fazem seus atores a respeito do significado e esta negociação, sempre visando a precisão, se dá em domínios específicos do conhecimento.

Retoma-se a proposição clássica de Shannon (1948, tradução nossa) sobre o problema fundamental da comunicação, para localizar o problema das ontologias.

O problema fundamental da comunicação é o de reproduzir num ponto, de forma exata ou aproximada, uma mensagem selecionada noutro ponto. Frequentemente as mensagens possuem significado, isto é, referem-se ou estão relacionadas a algum sistema, a certas entidades físicas ou conceituais. Estes aspectos semânticos da comunicação são irrelevantes para o problema da engenharia.

No que Capurro (2003) denomina de *paradigma físico* nos processos (produção, coleta, organização, interpretação, armazenagem, recuperação, disseminação, transformação e uso) que envolvem a informação, tomar esta teoria como modelo implica a analogia entre “a veiculação física de um sinal e a transmissão de uma mensagem, cujos aspectos semânticos e pragmáticos intimamente relacionados ao uso diário do termo informação são explicitamente descartados por Shannon”, critica, textualmente, Capurro

(2003), que ainda cita como exemplo dessa influência problemática os experimentos realizados pelo *Cranfield Institute of Technology* em busca de medidas de revocação e precisão para avaliar sistemas automatizados de recuperação da informação.

Numa leitura apressada, pode-se apontar como um *erro* o fato de Shannon ter considerado como irrelevantes os aspectos semânticos da comunicação. O que tem sido apontado assim, contudo, desconsidera que sua conceituação visava, na verdade, outro paradigma, o paradigma físico da informação. Se há um *erro*, ele não está na teoria da informação proposta por Shannon, mas na apropriação indevida que dela se fez. Encontrou-se em Echeverría (2003, p. 84, tradução nossa) apoio substancial para este ponto de vista:

O entendimento prevalente de comunicação hoje é baseado na noção de transmissão de informação. Este é um conceito herdado da engenharia de comunicação e desenvolvido por C. Shannon, entre outros. Ocupa-se da comunicação entre máquinas - isto é, entre um transmissor e um receptor (como nos processos de transmissão por rádio). Este marco, apesar da sua utilidade na transmissão técnica mostra sua fragilidade quando utilizado para compreender a comunicação humana. A noção de transmissão de informação esconde, justamente, a natureza problemática do escutar humano.

A noção de transmissão de informação neste caso é diferente da que a ciência da informação e mesmo a ciência da computação moderna admitem. Há, naquele modelo, comunicação bem sucedida quando a informação é transmitida, independentemente de seu conteúdo. Ainda apoiado em Echeverría (2003), cita-se um exemplo: há comunicação bem sucedida quando a tela do televisor apresenta uma imagem clara e estável do que se passa no estúdio; não se questiona, neste caso, que *sentido* faz para o televisor a imagem recebida. No caso da comunicação humana, a noção de transmissão da informação funciona apenas como metáfora. Trata-se, entretanto, de uma metáfora inadequada, pois distorce o fenômeno que pretende revelar, uma vez que não se podem oferecer garantias a respeito da transmissão da informação.

Ainda que originalmente a teoria de Shannon refira-se aos aspectos técnicos da transmissão de informação, constitui-se numa importante contribuição para o conceito de informação que a ciência da informação desenvolveu posteriormente. O aspecto

matemático da informação, contudo, incluindo a sua quantificação, é mais fortemente abraçado pela ciência da computação, por este motivo foi confundida em alguns momentos com o processamento de dados. Há ecos da proposta de Shannon na dificuldade histórica para o trato da informação semântica neste campo. O tratamento semântico da informação requer a construção coerente de redes de relações entre os termos, razão pela qual se recorre novamente à terminologia.

## 6.3 INFORMAÇÃO DOCUMENTÁRIA E TERMINOLOGIA

O desenvolvimento terminológico é considerado como componente e consequente do desenvolvimento científico e tecnológico. Não há conhecimento científico sem terminologia e, por esta razão, a legitimação da ciência se dá por meio da legitimação de sua própria terminologia, como ensina Benveniste (1989, p. 252):

A constituição de uma terminologia própria marca, em toda ciência, o advento ou o desenvolvimento de uma conceitualização nova, assinalando, assim, um momento decisivo de sua história. Poder-se-ia mesmo dizer que a própria história particular de uma ciência se resume na de seus termos específicos. Uma ciência só começa a existir ou consegue se impor na medida em que faz existir e em que impõe seus conceitos, através de sua denominação. Ela não tem outro meio de estabelecer sua legitimidade senão por especificar seu objeto denominando-o, podendo este constitui uma ordem de fenômenos, domínio novo ou um modo novo de relação entre certos dados. O aparelhamento mente consiste, em primeiro lugar, de um inventário de termos que arrolam, configuram ou analisam a realidade. Denominar, isto é, criar um conceito é, ao mesmo tempo, a primeira e última operação de uma ciência.

Considerando-se que não há mais delimitação rígida de domínios do conhecimento, suas terminologias também não são mais tão facilmente demarcadas, mas sua identificação torna-se condição para os diálogos que a ciência e a tecnologia estabelecem entre si e com seus públicos. Sem que se conheça a abrangência terminológica de um determinado campo não se podem negociar sentidos ou mesmo significados em seu interior. Esta impossibilidade

de negociação reduz (se não inviabiliza) a comunicação eficaz. O termo, na linguagem especializada, assume função muito semelhante à atribuída à palavra, na linguagem natural. A palavra, na linguagem natural, também depende de negociação de sentido e essas negociações também se inserem em contextos regulatórios mais ou menos específicos. A diferença fundamental, então, está na pré-definição, no pré-acordo com relação ao grau de especificidade do domínio que o termo requer, com vistas a diminuir consideravelmente a ambiguidade da linguagem natural.

A intensidade da comunicação científica em nível nacional ou internacional tem favorecido cada vez mais a vulgarização de termos especializados, os quais deixam o domínio da ciência e da tecnologia e alcançam outras esferas da sociedade como um todo. Hoje o vocabulário técnico-científico-especializado, notadamente nas ciências humanas, não é tão facilmente distinto, pois os diversos campos da ciência não são também tão facilmente demarcados, como apregoeou o sonho positivista; algumas unidades transitam entre a linguagem natural e as linguagens de especialidades.

A facilidade de acesso aos veículos de informação ajuda igualmente a promover a fusão dos repertórios especializados com a linguagem natural. Este processo é denominado terminologização: a ressignificação das palavras da língua comum fazendo com que alcancem estatuto de termo (KRIEGER; FINATTO, 2004). O movimento contrário é o da vocabularização, a conversão de um termo em vocábulo. Há ainda o movimento de metaterminologização, quando se dá a passagem (conversão) de um termo de um campo para outro. Cita-se, como exemplo, o termo *vírus*, que se movimenta do campo das ciências da saúde para o da ciência da computação. Neste caso a conversão se deu sem a manutenção do núcleo sêmico comum, isto é, entre os termos *vírus (ciências da saúde)* e *vírus (ciência da computação)* não há intersecção semântica.

Diante da profusão de novas mídias e da grande variedade de ofertas (ou de potencialidades) de informação, a garantia de regularidade e de precisão na produção e na comunicação científica demanda novas propostas para a representação, busca e recuperação das informações produzidas. Quanto maior o número de documentos

disponíveis, maior preocupação manifesta-se no processo de recuperação com a separação do que é relevante daquilo que não é. Os critérios de avaliação de relevância têm preocupado a biblioteconomia já há muito tempo; ainda assim não formam consenso, a começar pelo próprio conceito de relevância e suas medidas de avaliação de sistemas de informação (CAPURRO, 2003). Se tradicionalmente os sistemas de informação documentária focaram e buscaram metodologias para antecipar (pela identificação do perfil) e satisfazer as necessidades de informação dos usuários, atualmente, em face das modernas tecnologias e da *enxurrada* de informações, do ponto de vista do pesquisador que busca a informação, a preocupação principal passa a ser conseguir organizar uma estratégia de busca que permita identificar e separar os documentos *não desejados*.

Cada comunidade científica possui seus próprios modelos terminológicos e seus próprios aparatos conceituais. Em se promovendo a interexportação desses modelos sem cuidados metodológicos necessários, pode-se incorrer no erro de não identificar as novas relações que os termos e conceitos estabelecerão entre si. Em outras palavras, corre-se o risco de obrigar o cientista pesquisador à necessidade constante de reconstruir conceitualmente seu campo de estudos, o que impediria a comunicação científica e, como consequência, qualquer avanço. A comunicação científica eficaz, como qualquer outra forma de comunicação, depende das negociações de sentido e dos acordos (explícitos ou tácitos, conforme a fase em que se encontra a inserção do novo conceito) a respeito de seus usos. Dessa forma, dispensa-se o cientista de “tentar construir seu campo de estudos começando pelos primeiros princípios e justificando o uso de cada conceito introduzido” (KUHN, 1978, p. 40), posto que pode adotar como certos alguns paradigmas orientadores.

Naturalmente que não se pode simplesmente engessar o conceito num princípio impraticável de univocidade de referência. Esta tem sido a maior crítica aos trabalhos pioneiros de Wüster. Na concepção wüsteriana, a terminologia expressa conceitos (estáveis, paradigmáticos, universais) e não significados (linguísticos, variáveis). Esta concepção, critica Krieger (2001a, p. 69), “está fundamentada num modelo positivista de ciência que ainda compreende a linguagem das ciências como lugar homogêneo e transparente, dada sua

função única de expressar as verdades científicas”. A questão, contudo, ainda não é pacífica e gera algumas discussões e controvérsias. A mesma autora, em outra obra, insere as linguagens documentárias no *modelo positivista*:

A linguagem de indexação objetiva, em última análise, mantém [sic] eficiência no processo de divulgação/busca da informação. Trata-se, portanto, de uma diretriz voltada ao alcance da univocidade comunicacional, reconhecida como condição essencial para proporcionar uma eficiente aproximação entre sistemas de informação e usuário (KRIEGER; FINATTO, 2004, p. 60).

A *univocidade comunicacional* desconsidera os diferentes significados e os sentidos que a linguagem no uso pode emprestar aos termos. Isso aparece também nas normas de elaboração de tesouros que, como observou Lara (2009), nem mencionam a polissemia. Se, conforme Wittgenstein, o critério orientador para a semântica é o uso diário, fica fácil deduzir que os sistemas de recuperação da informação forçam o pesquisador a utilizar a linguagem de forma não natural (BLAIR, 2003).

Neste sentido, é importante apontar o auxílio que as metodologias para construção de ontologias podem oferecer à linguagem documentária. Os tesouros, na condição de linguagem documentária, de código comutador, são instrumentos que visam a orientar as relações entre os conceitos e a atribuição de descritores pelo sistema de informação documentária e seu posterior uso para fins de recuperação. Na condição de código comutador, o tesouro lida, necessariamente, com a redução semântica. As experiências com o uso da linguagem natural podem até mesmo melhorar as condições de recuperação da informação (embora ainda não se conheçam resultados satisfatórios sobre isso), mas escapam das questões da linguística documentária, porque não envolvem a representação (ou a construção de informações documentárias). O mapeamento terminológico, ou seja, a identificação dos termos representativos do domínio, seus conceitos básicos e seus relacionamentos, enquanto fase da criação de ontologias, possibilita a identificação da informação no discurso, dos termos que devem ser modelados como classes, propriedades ou axiomas (BREITMAN, 2005), fixando as regras de inferência.

## 6.4 INFORMAÇÃO DOCUMENTÁRIA E REPRESENTAÇÃO DO CONHECIMENTO

Pode-se efetivamente representar o conhecimento para fins de recuperação da informação? Esta questão tem unido os esforços de pesquisadores da ciência da computação e da ciência da informação, mas tem origem mais remota; nasce na filosofia, mais precisamente estudada pela teoria do conhecimento. Do ponto de vista da ciência da informação, entretanto, é preciso completá-la com objetivos também pragmáticos, pois é preciso perguntar também pela qualidade e pela viabilidade desta possível representação.

No que se refere às ontologias, quando se opta pela representação de algo, realiza-se uma decisão de *design* (modelagem) do modo ou ponto de vista pelo qual as relações lógico-semânticas serão formalizadas. Para Gruber (1995), a questão sobre o que integra o universo da representação nas ontologias resolve-se com um axioma: em inteligência artificial o que existe é o que pode ser representado. Com exceção da postura radical de Gruber, não é tão diferente do que ocorre nas linguagens documentárias, invariavelmente desenhadas para atender objetivos institucionais.

O conceito de representação ainda está atrelado à ideia de conhecimento como representação do real, o que Santos (2003) considera como herança do positivismo; mas isso não pode ser aplicado aos tesouros ou às ontologias, pois a ideia de representação nestes casos é operacional e escapa da noção de mimese. Nas ontologias, como na documentação, a representação do conhecimento e a construção de informações documentárias (KOBASHI, 1994; LARA, 2009) tornam-se a própria essência deste campo e, pode-se dizer, um dos pilares da ciência da informação.

A informação documentária, ensina Kobashi (1994, p. 24), deve “de um lado, promover a identificação de itens informacionais que respondam de modo pertinente a uma dada pergunta e, de outro, deve permitir a tomada de decisão sobre a consulta ou não do

documento original”. Deve-se, então, tomar cuidados metodológicos na sua elaboração, para que se possam garantir “a equivalência de sentido entre o texto fonte e a sua representação” (*loc. cit.*). A equivalência de sentido, contudo, não se dá no mesmo nível em todos os itens gerados na análise documentária. Enquanto o resumo mantém relação de contiguidade e semelhança com o texto original (condição facilitada pelo uso do mesmo sistema semiótico), o índice depende da intermediação de outro sistema semiótico (uma linguagem documentária) para sua geração (LARA, 2009).

Qin e Paling (2001, tradução nossa) consideram que, no âmbito da comunidade de biblioteconomia e ciência da informação a representação do conhecimento opera como sinônimo de organização do conhecimento, ou seja, a representação do conhecimento é referida como “um processo de organizar o conhecimento em classificações, tesauros ou listas de cabeçalhos de assunto”.

Com o fim de posicionar as ontologias na representação do conhecimento, apresenta-se um quadro com a síntese das abordagens de Guarino (1995), Campos (2004) e Moreira, Alvarenga e Oliveira (2004), todos leitores de Ron Brachman, no qual se baseiam (Quadro 8). Brachman classificou a representação do conhecimento, inicialmente em quatro níveis que não incluíam o nível ontológico.

NÍVEL	ASPECTO PRINCIPAL	CARACTERÍSTICAS
Lógico	Formalização	Formalização; sem preocupação com a semântica dos termos.
Epistemológico	Estrutura	Estruturação do conhecimento.
Ontológico	Significado	Insere o conceito num contexto e restringe suas possibilidades de interpretação.
Conceitual	Conceitualização	Especificação da estrutura dos conceitos de um determinado domínio.
Linguístico	Linguagem	Associação entre conceitos e termos.

Quadro 8 - Níveis de representação do conhecimento

Tálamo e Lara (2006), apoiadas em Capurro, e Lara (2009) apoiada em Rorty, negam a concepção de representação como espelho, como “relação mimética entre representação e objeto/fenômeno representado” no interior da linguística documentária. A ideia de representação para a linguística documentária, sustentam Tálamo e Lara (2006, p. 210), “é

operacional e antecede a delimitação do conceito de informação documentária. A ideia pressuposta no conceito de representação tem o objetivo de mostrar que ela é algo que está em lugar de alguma coisa para alguém sob um determinado aspecto” (*op. cit.*, p. 211). Neste aspecto reforçam a proposta de Kobashi (1994) da substituição do termo *representação documentária* por *informação documentária* no intuito de descartar a ideia de mimese. O termo *informação documentária*, defendem Tálamo e Lara (2006, p. 211), tem a vantagem de “além de circunscrever melhor o conceito no interior de seu campo específico” constituir “um amadurecimento em relação à ideia de representação de conteúdos”.

Não se discorda da proposta, considera-se realmente o termo menos ambíguo, mas pretende-se ampliar a discussão porque a problemática em torno do conceito não se resolve e se considera a ideia de representação como seminal para a ciência da informação. Mesmo que a informação documentária não represente mimeticamente ou espelhe o objeto de informação, ela continua representando-o, notadamente quando se considera a função de referencialidade do termo, conforme Le Guern<sup>36</sup> (*apud* LARA, 1993, p. 79):

Para Le Guern, a principal diferença entre as palavras na língua – que constituem o léxico – e as palavras no discurso – as palavras em funcionamento – reside na questão referencial. Na língua as palavras não estão em relação imediata com as coisas, com o concreto; exprimem apenas um conjunto de propriedades, têm um significado, mas não uma referência; dizem respeito a propriedades, não a substâncias; a qualidades e não a objetos.

Busca-se uma nova concepção de representação que não seja tributária da crença ingênua numa relação de fidedignidade entre conteúdo de um documento e a expressão utilizada para sua representação, isso transformaria a linguagem documentária em nomenclatura. Lara (2001) já chamara a atenção para os perigos que tal concepção acarretaria.

A ideia de representação foi, realmente, necessária para fixar o *modus operandi* da documentação. Dado o grande incremento do número de documentos disponíveis, principalmente com o avanço da informática, tornou-se necessário, como estratégia de

---

<sup>36</sup> LE GUERN, M. Sur les relations entre terminologie et lexique. *Meta*, v. 34, n. 3, p. 340-343, 1989.

sobrevivência, envidar esforços nas tecnologias de organização da informação. Barreto (2002) denomina este período ‘era da gestão’. Pela impossibilidade verificada logo no início desta *era* de dar acesso ao conteúdo completo do documento “foi necessário estabelecer uma metodologia de reformatação do documento, baseado em sua substituição por indicadores do conteúdo do documento todo” (BARRETO, 2002, p. 71). Umas poucas palavras-chave e algum resumo pretendiam, neste caso, representar o conteúdo completo do documento. Esse *ocultamento da informação* baseava-se numa metalinguagem *usurpada* da linguagem natural. A *era da gestão*, sintetiza Barreto (2002, p. 72), “trouxe o esplendor das classificações, indexações, tesauros, medidas de eficiência na recuperação do documento como: *recall* e *precision*”.

## **7 A ÁRVORE E O RIZOMA: A PROPOSTA DE DELEUZE**

Não se tem com esta pesquisa qualquer pretensão de romper com a corrente que compreende as hierarquias, múltiplas ou não, como condicionantes do que hoje se entende como organização do conhecimento, conforme sua aplicação nas diversas linguagens documentárias. Pretende-se, sim, dar voz a outro discurso, provocar outras reflexões. A história das ideias ensina que duas grandes correntes de pensamento quase nunca são similares, pois não pertencem ao mesmo recorte epistemológico da realidade, por esta razão empreende-se este esforço de reflexão.

### **7.1 ANTES, UMA PEQUENA ENTRADA NA SOCIEDADE DA INFORMAÇÃO**

A informação, como parece estar se tornando cada vez mais consensual, é uma construção, muitas vezes um caminho para o conhecimento e envolve inúmeras variáveis ainda não mapeadas pela inteligência artificial ou pelos estudos sobre cognição. Considerados em suas limitações, contudo, os agentes inteligentes e o desenvolvimento de ontologias representam um avanço muito significativo em direção à sociedade do conhecimento que é almejada por todos, de modo geral. Distingue-se, brevemente, abaixo a

sociedade do conhecimento da sociedade da informação, com o fim de fundamentar a discussão e enfatizar o papel da tecnologia.

A chamada sociedade da informação com as suas características peculiares, mesmo que ainda longe de consenso, só é possível porque existem as tecnologias de informação e comunicação, com suas garantias de conectividade, instantaneidade e de eliminação da distância espacial. Trata-se de iniciativa de diversos países, que envolve não apenas a informação científica e tecnológica, mas também a que se dá em outros níveis, como a informação que se vincula ao exercício da cidadania, por exemplo. O objetivo do Programa Sociedade da Informação, conforme definido no *Livro Verde*, é

integrar, coordenar e fomentar ações para a utilização de tecnologias de informação e comunicação, de forma a contribuir para a inclusão social de todos os brasileiros na nova sociedade e, ao mesmo tempo, contribuir para que a economia do País tenha condições de competir no mercado global (TAKAHASHI, 2000).

Pode-se perceber, então, que a sociedade da informação não é um fim em si mesma, mas condição para a sociedade do conhecimento. Numa visão pragmática e utilitarista, não há sentido em fazer circular informação sem que se estabeleçam condições para a criação e a circulação do conhecimento. A promessa do conhecimento precisa considerar o indivíduo, deve assegurar-lhe conforto e condições tecnológicas para que possa apreender a informação.

Trata-se neste caso de evolução consequente; da mesma forma que o conhecimento é, neste quadro, posterior e superior à informação, a sociedade do conhecimento é posterior e superior à sociedade da informação, sintetiza Rojas (2005). As ênfases na administração e no necessário controle de estoques formaram a premissa técnica para a reflexão sobre a melhor forma de propiciar condições para a oferta de informações (BARRETO, 2002). Esta perspectiva da aquisição de conhecimento como algo posterior e complementar ao acesso à informação já estava presente anteriormente em outro texto do mesmo Barreto (1994), para quem

democratizar a informação não pode [...] envolver somente programas para facilitar e aumentar acesso à informação. É necessário que o indivíduo tenha

condições de elaborar este insumo recebido, transformando-o em conhecimento esclarecedor e libertador, em benefício próprio e da sociedade onde vive.

Barreto (1994; 1996) e Rojas (2005), entre outros autores, também concordam em advertir para o risco de que ao se falar sobre sociedade do conhecimento isso se faça como *slogan mercadológico*, com seu componente ideológico implícito.

Considerando-se, então, este fluxo, a ciência da informação tem importante papel a desempenhar. Na síntese de Smit (2007), a ciência da informação promove o encontro entre registros de informação e usuários. Esse encontro passa pela organização da informação, de modo a garantir sua permanência no tempo e no espaço. Não se refere aqui a qualquer tipo de informação, mas à informação registrada (em sua origem ou como informação documentária), mediada por uma linguagem.

A presença de uma metalinguagem é, aliás, uma das marcas distintivas da ciência da informação no trato com a informação. Em termos mais simples, a condição para que a informação se torne objeto da ciência da informação é a possibilidade de seu registro. O registro da informação e sua utilidade real ou potencial – ou seja, o interesse por sua recuperação num fluxo de conhecimento – é que vão garantir sua permanência.

O termo *recuperação da informação* foi cunhado por Calvin Mooers, em 1951, e diz respeito aos “aspectos intelectuais da descrição da informação e sua especificação para busca, e também de qualquer sistema, técnicas ou máquinas que são empregadas para realizar esta operação” (MOOERS<sup>37</sup> *apud* FERNEDA, 2003, p. 11). Como uma das acepções de recuperação remete à ideia de procurar/localizar algo tecnicamente perdido, é interessante acrescentar que esta busca se dá por algo que foi organizado, *ordenado* antecipadamente para que pudesse ser encontrado. Saliente-se, contudo, que as metodologias e tecnologias desenvolvidas e aprimoradas ao longo da história, utilizadas para a produção de informações documentárias com o fim de maximizar o fluxo de informação e a produção de

---

<sup>37</sup> MOOERS, Calvin. Zatocoding applied to mechanical organization of knowledge. *American Documentation*, v. 2, n. 1, p. 20-32, 1951.

conhecimentos, ainda seguem a mesma *ordem do livro* (ou do código), denunciada em tempos já relativamente remotos por Chartier (1994, p. 8):

O livro sempre visou instaurar uma ordem: fosse a ordem de sua decifração, a ordem no interior da qual ele deve ser compreendido ou, ainda, a ordem desejada pela autoridade que o encomendou ou permitiu a sua publicação. Todavia, essa ordem de múltiplas fisionomias não obteve a onipotência de anular a liberdade dos leitores. Mesmo limitada pelas competências e convenções, essa liberdade sabe como se desviar e reformular as significações que a reduziram.

Uma linguagem documentária, como o tesouro, permite representar, para fins documentários, a informação registrada. Para isso é necessário estruturar um sistema conceitual de relações que permita delimitar o universo de interpretação dos signos documentários. Neste sistema os conceitos não existem isoladamente, mas coexistem e condicionam sua compreensão ao tipo de relacionamento que estabelecem com outros conceitos.

## **7.2 A LÓGICA DA ÁRVORE COMO MODELO DE REPRESENTAÇÃO**

Os relacionamentos entre os conceitos, na proposta da terminologia wüsteriana, dividem-se em lógicos (gênero-espécie) e ontológicos (todo-parte). As relações lógicas são estabelecidas apenas no plano conceitual, não refletem a experiência concreta. Cita-se, como exemplo, a relação entre mamífero e leão. Já as relações ontológicas são relações indiretas entre conceitos porque resultam das propriedades que possuem os objetos do mundo empírico, como, por exemplo, a relação entre relógio analógico e ponteiro. Trata-se, como se vê, de uma tipologia de raiz binária, na qual todas as outras relações são derivadas das primeiras.

Arnzt e Picht (apud CABRÉ; MOREL; TEBÉ, 1996) redefinem a tipologia de Wüster adotando duas novas superclasses de relações: as relações hierárquicas e as relações não-hierárquicas. As relações hierárquicas englobam as relações lógicas e as relações ontológicas wüsterianas; as relações não-hierárquicas contemplam as relações conceituais regidas por critérios que não incluem as hierarquias, entre as quais se encontram, por exemplo, as relações causais.

A norma ISO 704 (INTERNATIONAL ORGANIZATION FOR STANDARDIZATION, 2000) segue, ao que parece, esta tipologia, agrupando as relações em: hierárquicas (genéricas, partitivas) e associativas. Essa dissociação, ainda dicotômica, rebate-se, certamente, nas condições e na qualidade da representação visual que se pode gerar dos relacionamentos conceituais que serão estabelecidos. Se as relações hierárquicas são facilmente demonstráveis por meio de ferramentas gráficas, isso não é tão tranquilo quando se trata das relações associativas, precisamente porque são mais abertas, menos previsíveis e potencialmente mais ricas.

Afora algumas experiências mais ricas em interatividade, como o *Visual Thesaurus*<sup>38</sup> (na verdade, um dicionário em interface hiperbólica), que permite criar mapas que representam os significados das palavras e identificam palavras relacionadas, ainda está muito presente o modelo representacional da árvore nas representações visuais dos produtos gerados pela lexicografia ou pela terminologia, apesar de todos os avanços em direção à complexidade verificados nestes campos.

Não é necessariamente a lógica da árvore o modelo ideal para a representação do conhecimento, considerando-se que esta representação não ocorre de modo isolado, mas

---

<sup>38</sup> Disponível em: <<http://www.visualthesaurus.com/>>. Acesso em: 01 mar. 2010.

no interior de redes de informação. Deleuze e Guattari (1995) contrapõem esta linha arborescente a uma lógica dos múltiplos singulares. Compreender se e como a ciência da informação pode apropriar-se de sua instigante filosofia é o desafio.

Gilles Deleuze (1925-1995) é um dos principais pensadores franceses da contemporaneidade. De modo semelhante ao que acontece quando são propostas idéias inovadoras, quando as estruturas assentadas precisam ser revistas, quando velhos paradigmas – assim como se faz com os tapetes – precisam ser batidos para que sejam limpos, o pensamento legitimamente filosófico de Deleuze instaura novos olhares para o modo como se compreende e, por extensão, como se representa o mundo. Discute-se, portanto, o conceito de rizoma<sup>39</sup> apresentado pelo autor e suas possíveis implicações na organização do conhecimento enquanto saber/poder/fazer da ciência da informação.

Deleuze caminha, ao longo de sua produção científica, da geração de obras sobre história da filosofia (Nietzsche, Kant, Bergson e Spinoza) e crítica literária (Proust, Zola, Kafka) para a filosofia crítica (Diferença e repetição, Lógica do sentido, O anti-Édipo: capitalismo e esquizofrenia e Mil platôs: capitalismo e esquizofrenia, estes dois últimos com Felix Guattari). O filósofo recebeu pouca atenção crítica na França. Em parte isso é creditado, advertem seus estudiosos (LECHTE, 2002), à sua rejeição sistemática da relação mestre-discípulo, ao fato de não reverenciar os pensadores que o antecederam (com exceção de Nietzsche), como fizeram seus contemporâneos. Para ele, ensina Lechte (2002), um filósofo que pensa, que cria um evento no pensamento, separa-se da história da filosofia e entra no

---

<sup>39</sup> Em botânica, chama-se rizoma a um tipo de caule que algumas plantas verdes possuem, cujo crescimento é horizontal e muitas vezes subterrâneo, podendo também ter porções aéreas. Certos rizomas, como em várias espécies de capim (gramíneas), servem como órgãos de reprodução vegetativa ou assexuada, desenvolvendo raízes e caules aéreos nos seus nós. Em outros casos, o rizoma pode servir como órgão de reserva de energia, na forma de amido, tornando-se tuberoso, mas com uma estrutura diferente de um tubérculo (VIDAL; VIDAL, 2003, p. 102).

deserto, conforme a compreensão nietzscheana deste termo<sup>40</sup>. A relação que Deleuze estabelece com as obras de Nietzsche, com os textos literários, o cinema, a música popular e a pintura “fizeram-no evoluir para um questionamento não destruidor, mas crítico, até mesmo desconstrutor, de todos os grandes saberes constituídos” (ROUDINESCO, 2007, p. 205).

A obra de Deleuze é, no mínimo, provocadora. Numa leitura idiossincrática, convida, de modo indireto, à reflexão sobre um dos temas mais relevantes para a ciência da informação: a organização do conhecimento por meio da crítica que faz ao estruturalismo. O conceito de rizoma, desenvolvido pelo autor juntamente com Felix Guattari, e sua já clássica oposição à árvore em especial, interessa de perto a este campo, já que relaciona fortemente a organização com a hierarquização e seus consequentes desdobramentos. A proposta do rizoma é explorada em *Mil Platôs*, recebendo maior ênfase (um platô inteiro, na verdade) no primeiro volume da obra, o qual será discutido neste capítulo.

*Mil Platôs* compartilha com o *Anti-Édipo*, também de Deleuze em parceria com Guattari, o subtítulo capitalismo e esquizofrenia. Entretanto, ainda que os autores afirmem tratar-se de continuação e final do que fora tratado no *Anti-Édipo*, suas recepções, dadas as condições contextuais foram distintas.

O *Anti-Édipo* ainda pertence a 68 (publicado em 1972), surge, portanto, numa “época agitada”, oposta, em certo sentido, à “calmaria já absoluta, a indiferença em que o outro surgiu”, em 1980 (DELEUZE; GUATTARI, 1995, p. 7). A data referida é da publicação original,

---

<sup>40</sup> “[...] Nietzsche encerrara a primeira parte de ‘Humano, Demasiado Humano’ com uma rapsódia ao peregrino filósofo e sua relação com a noite e com a manhã iminente. No aforismo número 638 ele escreve: *É verdade que tal pessoa há de sofrer noites difíceis, em que estará cansado e encontrará fechado o portal da cidade que deveria lhe oferecer descanso*. Isso é ruim, porque o deserto chega até o portal e a noite cai como um segundo deserto. [...]” (SAFRANSKI, 2005, p. 200).

Mil Platôs foi publicado no Brasil apenas quinze anos depois, em 1995, em cinco volumes concluídos em 1997.

No Anti-Édipo, os autores denunciam as “falhas de Édipo, do ‘papai-mamãe’ na psicanálise, na psiquiatria e até mesmo na antipsiquiatria, na crítica literária e na imagem geral que se faz do pensamento” (DELEUZE; GUATTARI, 1995, p. 7). Como se colocam contra a representação, rejeitam também o que Édipo representa na psicanálise.

Mil Platôs avança, constrói o conceito de multiplicidade, como rejeição e superação dos dualismos um/múltiplo, consciente/inconsciente, natureza/história, corpo/alma. Compõe-se de quinze platôs, conceito tomado de empréstimo a Gregory Bateson (antropólogo e sociólogo que trabalha conceitos de ecologia da mente e é um dos fundadores, ao lado de Wiener, do conceito de cibernética). O platô é definido por Bateson<sup>41</sup> (apud DELEUZE E GUATTARI, 1995, p. 47) como “uma região contínua de intensidades, vibrando sobre ela mesma, e que se desenvolve evitando toda orientação sobre um ponto culminante ou em direção a uma finalidade exterior”. Um platô está sempre no meio, nunca no início ou no fim. Um rizoma é feito de platôs. Rejeita, portanto, as dicotomias e a hierarquia como modelos de organização.

Um rizoma não começa nem conclui, ele se encontra sempre no meio, entre as coisas, inter-ser, intermezzo. A árvore é filiação, mas o rizoma é aliança, unicamente aliança. A árvore impõe o verbo “ser”, mas o rizoma tem como tecido a conjunção “e... e... e...” Há nesta conjunção força suficiente para sacudir e desenraizar o verbo ser (DELEUZE; GUATTARI, 1995, p. 35).

O volume um de Mil Platôs traz uma espécie de orientação, de guia para a leitura do texto. Nesta advertência apresenta e já desenvolve o conceito de rizoma, motivo pelo qual é preciso compreender um pouco melhor esta concepção de leitura. Os autores avisam que a

---

<sup>41</sup> BATESON. *Vers une écologie de l'esprit*: t. 1. Paris: Ed. du Seuil, 1977.

obra não deve ser lida seguindo-se uma abordagem tradicional. Deve, sim, ser entendida já no modelo rizomático. Deleuze ainda se refere à ideia do que se pode chamar, a sua revelia, uma nova estética da recepção de leitura em outra obra, numa passagem em que busca dessacralizar o livro:

[...] tratar um livro como se escuta um disco, como se olha um filme ou um programa de televisão, como se é tocado por uma canção: todo tratamento do livro que exigisse um respeito especial, uma atenção de outra espécie, vem de uma outra era e condena definitivamente o livro. Não há nenhuma questão de dificuldade nem de compreensão: os conceitos são exatamente como sons, cores ou imagens, são intensidades que convêm a você ou não, que passam ou não passam. Pop filosofia. Nada há a compreender, nada a interpretar (DELEUZE; PARNET, 1998, p. 10).

É preciso confessar, contudo, que nem sempre se experimenta tal fruição na leitura de Deleuze. Se for possível comparar sua obra com um disco este certamente não será um disco pop banal, como faz transparecer a citação acima.

Os modelos mais tradicionais de escritura e leitura são pautados na concepção linear início, meio e fim, ou, para preservar a terminologia dos estudos de gêneros textuais: introdução, desenvolvimento e conclusão. Nesta ordem canônica só se pode, teoricamente, avançar a partir da conclusão. A conclusão encerra, por assim dizer, um texto e abre novas possibilidades.

A concepção de livro apresentada pelos autores não tangencia e nem mesmo se relaciona com a concepção de documento (tomado aqui como equivalente de livro) como forma e como signo defendida por Pedauque (2004), desenvolvida anteriormente (Capítulo 6.1). Para Deleuze e Guattari o livro interessa apenas enquanto agenciamento em conexões com outros agenciamentos, “em relação com outros corpos sem órgãos”.

Não se perguntará nunca o que um livro quer dizer, significado ou significante, não se buscará nada compreender num livro, perguntar-se-á com o que ele funciona, em conexão com o que ele faz ou não passar intensidades, em que multiplicidades ele se introduz e metamorfoseia a sua, com que corpos sem órgãos ele faz convergir o seu. Um livro existe apenas pelo fora e no fora (DELEUZE; GUATTARI, 1995, p. 11).

Deleuze e Guattari propõem para a leitura de *Mil Platôs* a mesma estratégia de pensamento que pretendem apresentar e desenvolver ao longo da obra para a leitura do mundo: a abordagem rizomática. Nessa abordagem rompe-se com a linearidade aparente. Compreendendo que o livro não possui capítulos que se sucedem, mas platôs, ou zonas que possibilitam novas conexões, entendem os autores que o leitor pode, livremente, traçar seu percurso de leitura, saltar de um platô a outro, construindo seu rizoma conforme a orientação de sua própria concepção. Os platôs não possuem necessariamente início ou fim. Pode-se entrar ou sair deles por qualquer ponto; mas formam, ainda assim, uma rede, um rizoma, uma multiplicidade.

Por exemplo, uma vez que um livro é feito de capítulos, ele possui seus pontos culminantes, seus pontos de conclusão. Contrariamente, o que acontece a um livro feito de "platôs" que se comunicam uns com os outros através de microfendas, como num cérebro? Chamamos "platô" toda multiplicidade conectável com outras hastes subterrâneas superficiais de maneira a formar e estender um rizoma. Escrevemos este livro como um rizoma. Compusemo-lo com platôs. Demos a ele uma forma circular, mas isto foi feito para rir. Cada manhã levantávamos e cada um de nós se perguntava que platôs ele ia pegar, escrevendo cinco linhas aqui, dez linhas alhures. Tivemos experiências alucinatórias, vimos linhas, como fileiras de formiguinhas, abandonar um platô para ir a um outro (DELEUZE; GUATTARI, 1995, p. 32).

É impossível não fazer aqui paralelo com a concepção de hipertexto e com a geografia da internet, com o que há de concepção romântica na ideia abraçada por muitos estudiosos nos anos 1990 (o boom da internet) sobre a liberdade anárquica da rede. *Mil platôs* é de 1980, quando a internet ainda principiava, sem a visibilidade da web, que

ganharia desenvolvimento apenas na década seguinte Não há, portanto, na obra, qualquer referência explícita à ampla teia mundial. Sua configuração, entretanto, convida ao paralelo que se estabelece com base nos princípios do rizoma e nas características da internet, a partir da leitura de Hamman (1996), apresentado no Quadro 9. O agrupamento dos princípios, que serão mais detalhados no capítulo 7.4, segue a proposta de Deleuze e Guattari (1995).

1º e 2º princípios - Conexão e heterogeneidade	
RIZOMA	Qualquer ponto do rizoma pode ser conectado a qualquer outro ponto.
INTERNET	A descentralização como princípio fundamental no surgimento da internet. O envio de pacotes de informação de um computador para qualquer outro, em rotas pré-estabelecidas ou não. Qualquer computador, preparado para rede, pode conectar-se.
3º princípio – Multiplicidade	
RIZOMA	Contraposição às relações dicotômicas e hierárquicas; não existem pontos ou posições, apenas linhas em um rizoma.
INTERNET	À primeira vista, este princípio parece colocar em dúvida a utilização da internet como um exemplo de rizoma, pois o internauta vai de um ponto a outro, usando os endereços eletrônicos para encontrar e ler os <i>sites</i> . É preciso considerar, entretanto, que não se movimenta fisicamente entre os pontos, em vez disso permanece no mesmo local físico em todo o tempo de navegação. Não há pontos para ir que existam além do estado de <i>alucinação consensual</i> que é o ciberespaço, conforme a descrição de Gibson (1991).
4º princípio – Ruptura a-significante	
RIZOMA	Um rizoma não tem começo ou fim, pode ser rompido em qualquer lugar, pois retoma segundo uma ou outra de suas linhas e segundo outras linhas.
INTERNET	A internet foi projetada para resistir a um holocausto nuclear, seus protocolos permitem que os computadores que a compõem possam redirecionar os pacotes de informação em pontos de conflito.
5º e 6º princípios – Cartografia e decalcomania	
RIZOMA	O rizoma não é objeto de reprodução, porque não tem centro, hierarquia e corte significativo, então, suas ligações ilustram um mapa, e como tal possui várias entradas.
INTERNET	Existem muitos <i>links</i> entre computadores na internet, do mesmo modo, existem muitas entradas, já que se pode escolher qualquer <i>site</i> como entrada. Assim, o internauta cria mapas ligando páginas e movendo-se como um nômade que está navegando.

Quadro 9 - Comparação entre rizoma e internet

## 7.3 A MULTIPLICIDADE E A ÁRVORE

Um dos conceitos mais caros a Deleuze e Guattari em todo *Mil Platôs* é o da multiplicidade. Na concepção sumarizadora de Abreu Filho (DELEUZE; GUATTARI, 1998, p. 143):

A teoria da multiplicidade efetua uma interpretação do real que conjuga uma construção ontológica e uma leitura do mundo e da sociedade que surpreende com uma nova distribuição dos seres e das coisas: **não admite unidade natural**, uma vez que não se apóia em nenhuma necessidade e não visa a nenhum prazer; **não reconhece a falta**, uma vez que não se constitui em referência a uma unidade ausente (recusando, pois, a noção de desejo como falta); e **não aceita nenhuma transcendência** - seja na origem, como idéia ou modelo, seja no destino, como sentido historicamente desenvolvido. A perspectiva da imanência e o conceito de multiplicidade fazem do pensamento uma atividade ética - sem modelos e finalidades transcendentais - avessa a qualquer conforto moral ou orientação histórica (grifos nossos).

A multiplicidade pode ser contraposta à concepção de unicidade disseminada pelo uso da estrutura arborescente como metáfora para a representação do conhecimento, utilizada, pelo menos, desde a Idade Média, podendo-se citar a célebre Árvore de Porfírio (cuja base é aristotélica, conforme tratado no capítulo 2.2) e todos os seus desdobramentos como exemplos.

A longa permanência desse modelo de representação ocorre em função de dois aspectos compreensíveis (BURKE, 2003): a) a abordagem dicotômica simplificada revelada na oposição entre os elementos pela presença/ausência de determinada propriedade e b) a naturalização do convencional, pela apresentação da cultura como se fosse natureza. Apenas no século XVI aparece a noção de sistema, como representação mais complexa.

Para Deleuze e Guattari (1995, p. 29) toda lógica da árvore é uma lógica do decalque e da reprodução. No volume 4 de *Mil Platôs* Deleuze e Guattari (1997) criticam mais duramente a metáfora da árvore e seu devir condicionante. Compreendem o devir como

“um movimento pelo qual a linha libera-se do ponto, e torna os pontos indiscerníveis: rizoma, o oposto da arborescência, livrar-se (sic) da arborescência”.

A lógica da arborescência, e a dualidade que a constitui, aplicada às linguagens documentárias, numa visão clássica, dá ideia de que há desdobramentos a partir dos dipolos<sup>42</sup> de ligação. Citando textualmente Deleuze e Guattari (1995, p. 25):

A árvore ou a raiz inspiram uma triste imagem do pensamento que não pára de imitar o múltiplo a partir de uma unidade superior, de centro ou de segmento. Com efeito, se se considera o conjunto galhos-raízes, o tronco desempenha o papel de segmento oposto para um dos subconjuntos percorridos de baixo para cima: um tal segmento será um “dipolo de ligação”, diferentemente dos “dipolos-unidades” que formam os raios que emana de um único centro.

Isso está refletido na análise que Lara (1993) realizou de algumas linguagens documentárias (tesauros e classificações decimais). A partir da análise de alguns casos, concluiu que o uso de mais de uma característica de divisão para a construção de hierarquias é predominante na CDD e na CDU, o que implica a coordenação de termos que não possuem a mesma natureza. Este problema, aliás, já havia sido detectado por Coyaud<sup>43</sup> (*apud* LARA, 1993). A autora observou a mesma estruturação hierárquica dos domínios na análise dos tesauros, com predomínio das relações gênero-espécie e todo-parte.

Há aqui um paradoxo. Credita-se a Aristóteles a supremacia da hierarquia ainda presente nos modelos de organização do conhecimento em detrimento de modelos contemporâneos alternativos. O uso de mais de uma característica de divisão e a presença de relações lógicas e associativas demonstram, entretanto, mais influência da “base empírica do registro” (LARA, 1993) do que da lógica aristotélica.

---

<sup>42</sup> A interação química dipolo-dipolo ou dipolo permanente refere-se, basicamente, à força de atração que ocorre entre duas moléculas polares que se ligam por seus respectivos pólos, ou seja o pólo positivo de uma molécula liga-se ao pólo negativo de outra.

<sup>43</sup> COYAUD, M. *Introduction a l'étude des langages documentaires*. Paris: C. Klincksieck, 1966.

## 7.4 REVENDO AS ESTRUTURAS

O conceito de multiplicidade é, sob certos aspectos, contraposto ao conceito de estrutura. A estrutura impede o crescimento quando engessa as possibilidades num feixe pré-programado de possibilidades. Dessa forma “toda vez que uma multiplicidade se encontra presa numa estrutura, seu crescimento é compensado por uma redução das leis de combinação” (DELEUZE; GUATTARI, 1995, p. 13).

Apresentam, com o fim de elucidação, as principais características aproximativas dos rizomas concentradas em alguns princípios: 1º e 2º - Princípios de conexão e de heterogeneidade; 3º - Princípio de multiplicidade; 4º - Princípio de ruptura a-significante e 5º e 6º - Princípio de cartografia e de decalcomania. Cada um deles é detalhado nos itens subsequentes.

### **1º e 2º - Princípios de conexão e de heterogeneidade**

Como garantia de que todas as possibilidades estejam abertas, “qualquer ponto de um rizoma pode ser conectado a qualquer outro e deve sê-lo”. Este princípio é contrário, então, a qualquer forma de hierarquia ou de estrutura nos moldes como são mais conhecidas e utilizadas nos processos de representação do conhecimento. Retomando as críticas ao estruturalismo, Deleuze e Guattari (1995, p. 14-15) usam a gramaticalidade de Chomsky para evidenciar seu poder disciplinador ou homogeneizador:

o símbolo categorial S que domina todas as frases, é antes de tudo um marcador de poder antes de ser um marcador sintático [...]. Você constituirá frases gramaticalmente corretas, você dividirá cada enunciado em sintagma nominal e sintagma verbal (primeira dicotomia)

e assim por diante. Contrariamente à estrutura, que depende do relacionamento biunívoco entre os pontos para o estabelecimento de uma identidade relacional, os rizomas traçam linhas incertas, pontos de fuga e de desterritorialização.

Este desenho rizomático está presente hoje nas representações que se constroem da internet e sua hipermidialidade. O termo hipermídia é resultado de uma sutileza conceitual. A designação genérica hipertexto e sua natural vinculação com o elemento textual, indicada pelo próprio termo, não representam com precisão o tipo de processo predominante nos sistemas de informação modernos. Com a ampla circulação, em proporções que se equivalem, de arquivos sonoros, imagens e textos, a anteriormente utilizada designação *multimídia* ganhou maior abrangência e criou-se a partir desta fusão (hipertexto+multimídia) o termo hipermídia, com a finalidade de representar objetos que utilizam a tecnologia multimidiática (texto, som e imagem) com ligações, ou nós, entre si.

A riqueza da internet, enquanto rede de comunicação planetária que é, está na abundância; na possibilidade que esta tecnologia cria para que pessoas de variadas nacionalidades, de variados idiomas ou culturas, lançando mão de variadas mídias, possam interagir.

Neste ponto coloca-se um dos principais problemas atuais da ciência da informação: como garantir a organização do conhecimento e o acesso à informação registrada em suportes tão diferentes, mutantes, fluidos? Há condições suficientes para interação (entre pessoas), mas ainda insuficientes para a interatividade (pessoas e máquinas ou máquinas entre si) quando se trata do fluxo da informação documentária. Na perspectiva do ciberespaço, a subjetividade (quer seja do indivíduo mesmo ou da instituição) nas representações hierarquizadas, assumiria a perspectiva do uno e condenaria os sistemas de informação ao isolamento ou à formação de pequenos agrupamentos que só poderiam dialogar entre si.

### **3º - Princípio da multiplicidade**

É importante enfatizar que o conceito de multiplicidade não se opõe, quer contrária ou contraditoriamente, ao conceito de unidade, pois dessa forma estabelecer-se-ia simplesmente uma continuidade das dicotomias. Deleuze e Guattari (1995, p. 15) preferem tratar o múltiplo como substantivo; tratado como multiplicidade “ele não tem mais

nenhuma relação com o uno como sujeito ou como objeto, como realidade natural ou espiritual, como imagem e mundo”. A divisão estanque sujeito e objeto (outra dicotomia), tão cara ao positivismo, estaria superada.

Aos elos que são estabelecidos como garantia de multiplicidades, ao “crescimento das dimensões, numa multiplicidade que muda necessariamente de natureza à medida que ela aumenta suas conexões”, Deleuze e Guattari (1995, p. 16) denominam agenciamento.

Os pensadores sabem, no entanto, que não se apaga simplesmente o sujeito com sua inserção numa teia de agenciamentos. A noção de unidade continua aparecendo “quando se produz numa multiplicidade uma tomada de poder pelo significante ou um processo correspondente de subjetivação” (DELEUZE; GUATTARI, 1995, p. 16). No ambiente atual dos sistemas de informação, marcado por virtualidades a questão do sujeito-autor tem sido amplamente discutida, como já fora apontado por Foucault e explorado por Coelho Netto (1994). A partir da indagação de Foucault, ‘*Qu'est-ce qu'un auteur?*’, Coelho Netto apresenta questões interessantes e (ainda) atuais sobre o que o filósofo, pensava, em 1959, a respeito desta entidade. Uma delas é negação do autor por Foucault<sup>44</sup> (*apud* COELHO NETTO, 1994), para quem existe apenas uma *função-autor*, responsável pela geração de um texto, “instauradores de discursos”, “formadores de discursividade”, o indivíduo funciona como catalisador desse processo.

Nos atuais modelos de produção e uso de informação, possibilitados pela complexa conjunção de tecnologias de informação e de comunicação que se convencionou chamar simples e resumidamente de internet, torna-se cada vez mais comum a produção coletiva de conhecimentos.

#### **4º - Princípio de ruptura a-significante**

Uma estrutura, na definição clássica de Piaget (1979, p. 6)

---

<sup>44</sup> FOUCAULT, M. (1994), *Qu'est-ce qu'un auteur?* In: \_\_\_\_\_. *Dits et écrits*. Paris: Gallimard, 1994.

é um sistema de transformações que comporta leis enquanto sistema (por oposição às propriedades dos elementos) e que se conserva ou se enriquece pelo próprio jogo de suas transformações, sem que estas conduzam para fora de suas fronteiras ou façam apelo a elementos exteriores. Em resumo, uma estrutura compreende os caracteres de totalidade, de transformações e de auto-regulação.

Na visão pós-estruturalista de Deleuze e Guattari (1995, p. 17) é preciso investir contra a rigidez da estrutura. No modelo estruturalista auto-regulador qualquer ruptura torna-se necessariamente significativa e fere todo o sistema, podendo, inclusive, atravessá-lo. Um rizoma, por sua vez, esclarecem os autores “pode ser rompido em um lugar qualquer, pois retoma segundo uma ou outra de suas linhas e segundo outras linhas”. Em outra passagem (p. 13) sustentam: “Toda vez que uma multiplicidade se encontra presa numa estrutura, seu crescimento é compensado por uma redução das leis de combinação”.

## **5º e 6º - Princípio de cartografia e de decalcomania**

“Um rizoma não pode ser justificado por nenhum modelo estrutural ou gerativo. Ele é estranho a qualquer idéia de eixo genético ou de estrutura profunda” (DELEUZE; GUATTARI, 1995, p. 20). Esta rejeição deve-se à compreensão de que o eixo ou a estrutura “são antes de tudo princípios de decalque, reproduzíveis ao infinito”. A própria ideia de eixo, como linha imaginária ou real que determina o centro de um corpo é estranha à concepção do rizoma. A ideia de eixo aqui remete novamente à lógica da árvore, criticada vigorosamente pelos pensadores em estudo: “toda lógica da árvore é uma lógica do decalque e da reprodução” (DELEUZE; GUATTARI, 1995, p. 21). A árvore, prosseguem “articula e hierarquiza os decalques”, reproduz a si mesma quando pensa reproduzir o outro.

O relativo *insucesso* de grandes projetos nos campos da biblioteconomia e da documentação, como a universalização da classificação ou a criação de um centro mundial de documentação, pode ser tributado, em parte, a sua condição inerente de reprodução, de decalque a partir de um eixo principal, representado pela ideologia manifesta em adjetivos como *universal* e *mundial*, que tendem a ignorar a multiplicidade.

## 7.5 AS PROVOCAÇÕES DE DELEUZE

Deleuze foi saudado por Foucault (*apud* DELEUZE, 1992, p. 111) como instaurador de um novo pensamento, como um marco na filosofia:

Produziu-se uma fulguração, que levará o nome de Deleuze. Um novo pensamento é possível, de novo o pensamento é possível. Ele está aí, nos textos de Deleuze, saltitante, dançante, diante de nós, entre nós... Um dia talvez o século será deleuzeano.

Fazendo coro com Foucault, guardadas as devidas proporções, acredita-se que as ideias colhidas do pensamento de Deleuze indicadas neste capítulo possam ser aprofundadas e se tornar contributivas para uma releitura dos modelos de representação e, conseqüentemente, recuperação da informação no âmbito da ciência da informação. Tratam-se de forças quase contraditórias: de um lado a rejeição da ideia de representação em favor da multiplicidade e de seus agenciamentos, de outro a representação como fundamento epistemológico. Até que ponto, pode-se perguntar, a lógica do decalque interfere positiva ou negativamente nos processos usuais de representação da informação? Que contribuições as ontologias oferecem neste contexto?

A virtualização (desmaterialização) dos documentos e sua desterritorialização trouxeram novos desafios para o subcampo da organização da informação e seu conjunto de técnicas mais apropriado ao tratamento de documentos estabelecidos numa relação tempo-espço. O pensamento rizomático deleuzeano permite instaurar um novo olhar para as redes de informação, para uma nova pragmática das ações de informação sem a necessidade de adoção de um modelo regulado por um eixo específico, o que é, aliás, condição necessária para a organização e recuperação da informação no ciberespaço.

Algumas questões impõem-se neste novo cenário: qual a melhor opção para disponibilização do objeto de informação eletrônico se o conceito de estoque tornou-se instável, se não há mais árvores que possam abrigá-los? Ainda são necessárias a descrição e

a representação do documento (documento e seus metadados não estão fisicamente separados) ou eles mesmos se bastam? Se as técnicas de organização da informação tradicionais referem-se fortemente à informação tridimensional, como adaptar-se à ausência de fronteiras? É possível repensar a construção de informações documentárias para encontrar algo semelhante ao mundo *Google*™ no tratamento da informação sintática e ao mesmo tempo rico em possibilidades semânticas como o rizoma deleuzeano? Provocações...

## 8 CONSIDERAÇÕES FINAIS

Estabeleceu-se como objetivo geral para esta reflexão investigar a interface teórico-prática entre a linguística documentária, a terminologia e as ontologias (filosóficas, como aporte teórico, e computacionais, como conjunto de metodologias). A discussão partiu de um ponto de observação preciso, do interior do qual se analisou a problemática: os subsídios teórico-práticos que as ontologias oferecem para a construção de informações documentárias em sistemas de informação documentária.

Tomou-se como pressuposto, com base na investigação da literatura sobre ontologias produzidas pela ciência da informação e pela ciência da computação, que sem o estabelecimento de ontologias comuns não se poderão formar bases de conhecimento que garantam a qualidade das trocas informacionais necessárias entre os diferentes sistemas de informação interconectados por redes de informação. A interdependência inerente aos sistemas de informação, considerados seus objetivos, impede a insistência em modelos isolados ou nucleares de organização do conhecimento. Observou-se que, conquanto compartilhem alguns objetivos comuns, ainda há pouco diálogo entre a ciência da informação (e, no seu interior, a linguística documentária) e a ciência da computação (ontologias), principalmente desta última para com a primeira.

Para a consecução do objetivo geral foi preciso analisar as condições de produção, desenvolvimento, implementação, uso e integração de ontologias com base no referencial teórico da ciência da informação, tendo em vista os benefícios que as ontologias e os tesouros (produtos documentários) poderiam permutar. Dessa forma, em alguns momentos foi necessário evidenciar as relações que guardam entre si a ciência da informação e a ciência da computação. Este diálogo teve como base a convicção de que a aplicação dos

recursos da informática não resolve, por si, todos os problemas decorrentes da organização e do fluxo da informação, como quer acreditar o senso comum, e que, por outro lado, não é possível, considerando-se a base tecnológica atual, resolver estes problemas sem o auxílio do tratamento automático da informação.

Os métodos automáticos de organização da informação consideram basicamente os seus aspectos sintáticos e estatísticos, possivelmente por conta da facilidade com que os *softwares* aplicativos podem operar com este tipo de insumo. Defende-se, para dar conta dos aspectos semânticos ausentes, que os aportes da terminologia respondem, com maior acuidade, aos problemas de acesso menos quantitativo e mais qualitativo à informação.

A abordagem automática quantitativa é utilizada com base no argumento de uma pretensa objetividade, o mesmo princípio de objetividade, aliás, que norteou algumas teorias iniciais referentes à análise documentária, que ensinavam a indexar o *verdadeiro conteúdo* do documento e não aquele que o autor gostaria de ter escrito ou aquele que é resultado da *interpretação* do indexador. No âmbito dos sistemas de informação documentária, a abordagem automática quantitativa da informação, embora se preste ao levantamento de dados de importância para um trabalho de construção de filtros semânticos, não responde adequadamente às necessidades relativas ao tratamento documentário-informacional dos discursos, quer em relação aos conteúdos ou ao atendimento das necessidades territorializadas de informação.

Se a terminologia clássica se mostra potente para apoiar as tarefas documentárias (e de construção de ontologias) nos campos cuja terminologia é estável, é a teoria comunicativa da terminologia que se mostra mais adequada para tratar de campos onde o sistema terminológico está em formação ou transformação. Como se afirmou anteriormente, esta teoria busca compreender os termos, suas significações e suas relações em situações de uso a partir de metodologias que observam as características dos discursos nos quais há flutuação designacional e definicional.

Interessou também neste trabalho, verificar a contribuição das ontologias para o desenvolvimento de tesauros e vice-versa. Acredita-se que haja um conjunto de conhecimentos desenvolvido na ciência da informação referentes à organização e recuperação da informação e à organização e fluxo do conhecimento que é resultante de um longo processo de desenvolvimento, visando alcançar o seu aperfeiçoamento. Neste sentido são questionadas algumas pesquisas sobre ontologias por seu caráter absolutamente utilitarista, voltados que são à formalização sem considerar as particularidades dos campos de conhecimento e de atividade. Sobre esse aspecto, observou-se que, com raras exceções, como Sowa (2000) e Smith (2004), por exemplo, as discussões sobre ontologias não são realizadas a partir de fundamentação teórico-filosófica. Defende-se nesta reflexão, ao contrário – e na linha dos últimos autores citados – que a construção de ontologias concretas e as ontologias filosóficas não são eventos completamente independentes que guardam entre si apenas a similaridade de denominação. A discussão sobre categorias e categorização na ciência da computação, nem sempre possui, pelo que pôde observar nos limites desta reflexão, a ênfase que recebe na ciência da informação no âmbito dos estudos sobre organização e representação do conhecimento.

No que diz respeito às interfaces entre a ciência da computação e a ciência da informação, foi possível confirmar que, do lado da ciência da informação – em especial no que tange à construção de linguagens documentárias – pode haver uma profícua apropriação de metodologias para a construção de instrumentos documentários mais precisos, flexíveis e dinâmicos a partir do conhecimento acumulado, na ciência da computação, com a construção de ontologias de domínio e suas regras de inferências e relacionamentos.

A pesquisa permitiu verificar que tesauros e ontologias não possuem a mesma natureza, pois suas concepções e suas finalidades são distintas, mas podem ser observados sob um mesmo prisma: o da representação do conhecimento em contextos de linguagens especializadas (no recorte temático desta reflexão) com o objetivo de organizar de forma eficaz os fluxos da comunicação científica.

O tesauro, cujo surgimento coincide com o da ciência da informação, é resultante da necessidade de resolver os problemas de fluxo de informação (acesso e uso) decorrentes da rápida ampliação da produção do conhecimento científico e tecnológico neste período. A consequente dificuldade que este contexto trouxe para a organização e recuperação da informação científica também colaborou para sua aceitação e uso. O tesauro possibilitou, por meio do estabelecimento de rede lógico-semântica de conceitos, maior independência do pesquisador e maior eficácia nas buscas que necessita realizar em bases de dados. Este novo arranjo possibilitou organizar os termos e conceitos de um determinado campo e facilitar, dessa forma, o acesso à informação.

Um dos fatores que limitam os tesauros, entretanto, refere-se à ausência de regras de inferência explicitamente formuladas. O modo como as relações entre os conceitos são estabelecidas, cuja base precípua é a garantia literária, bem como a estrutura de sua apresentação, possibilitam que o tesauro assuma também função pedagógica. Esta função, contudo, é sempre tributária das necessidades de informação e das condições de interpretação que o usuário manifesta. Com exceção das notas de aplicação, portanto, não há indicações explícitas sobre o ponto de vista pelo qual o conceito é formulado. Como resultante disso, o tesauro precisa admitir alguma estabilidade do conceito para que possa operar. Espera-se que o sistema conceitual que o tesauro estabeleceu repita-se nos documentos futuros.

A estruturação dos tesauros em linguagem apropriada ao seu uso na *web* em formatos legíveis por máquina e passíveis de tratamento automático ainda é, como se pôde detectar na literatura estudada, incipiente. Sendo menos formalizados que as ontologias, os tesauros podem ser adaptados mais facilmente a contextos diferenciados dentro de um mesmo campo conceitual. Estas adaptações, contudo, carregam consigo a necessidade de interpretar (ou reinterpretar) sua modelagem conceitual, o que gera prejuízos a sua concepção semântica. A modelagem conceitual possibilitada pelas ontologias, pelo seu alto grau de formalização, tende a ser mais consistente nos aspectos que envolvem a

interoperabilidade, porque incorpora semântica de um modo mais formalizado, o que não significa, evidentemente, que consegue resolver todos os problemas de significação.

O fundamento filosófico da aplicação de ontologias foi discutido com base no estudo das categorias ontológicas presentes na filosofia clássica (Aristóteles e Kant) e nas propostas contemporâneas (Deleuze). O *leitmotiv* desta discussão foram as categorias, como fundamento das ontologias filosóficas e da representação documentária.

Na condição de instrumentos de modelização do conhecimento, as ontologias computacionais precisam envolver, além dos aspectos cognitivos e linguísticos, os aspectos teórico-filosóficos. A escolha das formas de organização (as relações conceituais) das ontologias implica, então, necessariamente, uma visão que se rebate, de certo modo, numa posição filosófica. Dessa forma, não se deve adotar como ponto de partida os *mundos substitutos* construídos dentro de modelos de *software* existentes. Pelo contrário, é preciso, concorda-se com Smith (2004), abordar a realidade em si, aproveitando a riqueza e a diversidade das descrições científicas possíveis em face das formas de organização reconhecidas pelas comunidades discursivas.

A lógica arborescente das hierarquias ainda é, certamente, orientadora de diversos discursos, dentre os quais o das ciências da informação e da computação. As transformações que os recursos da telemática e da hipermídia provocam, contudo, geram novas possibilidades de organização. Deleuze e Guattari (1995), por sua proposta do rizoma, foram tratados nesta pesquisa como provocadores de uma reflexão sobre a validade do modelo hierárquico e sobre as possibilidades de sua ampliação. É interessante notar que rizoma e raiz (a árvore) não são tratados como opostos. Esta aparente dualidade é recusada pelos autores, pois assumir isto implicaria em negar a multiplicidade. Um rizoma pode originar uma cadeia de raízes (hierarquias). Nas palavras dos autores: “existem estruturas de árvore ou de raízes nos rizomas, mas, inversamente, um galho de árvore ou uma divisão de raiz podem recomeçar a brotar em rizoma” (DELEUZE; GUATTARI, 1995, p. 23). Não se trata, então, de estabelecer classificações: *isto é um rizoma e isto não é*, pois modelo rizomático

possui também seu próprio despotismo e “não para de erigir e de se entranhar, [...] de se alongar, de romper-se e retornar” (op., cit., p. 30).

Deste ponto de vista, os sistemas de informação documentária e também as ontologias são menos arborescentes, então, do que se pensa comumente, pois, embora não possam ainda prescindir das hierarquias, provocam agenciamentos não previstos, fugas de interpretações, modos particulares ou coletivos diferenciados daquele que o sistema propõe como forma de organizar o conhecimento. Num sistema de informação documentária tradicionalmente concebido, as relações conceituais são dependentes de alguma estabilidade para seu perfeito funcionamento, por este motivo estes novos agenciamentos raramente são incorporados, uma vez que isto implicaria uma dinamicidade que poderia comprometer as relações conceituais estabelecidas.

Estabelecer relações conceituais, fixar, de alguma forma, algo que está sempre em movimento, equivale a retirar delas, pode-se afirmar, na trilha de Deleuze e Guattari (1995), o que possuem de mais rico: o devir. A árvore, exemplificam, não é verde, ela verdeja.

As inter-relações entre ontologia filosófica, ontologia computacional e representação documentária baseada em categorias de assunto também foram objeto desta reflexão.

É certo que as fronteiras entre os campos científicos, bem como a própria concepção de ciência, veem sendo paulatinamente dissolvidas, mas elas ainda existem e insistem. As questões sobre disciplinaridade (a aproximação da realidade por meio de um recorte possível) e interdisciplinaridade (os espaços contraditórios decorrentes) ainda fomentam calorosos debates. Desse modo, ainda que a ciência da informação e a ciência da computação (que já fora confundida com a *informática*) não cheguem a um acordo nem mesmo sobre o que é *informação* e ainda discutam territórios, é forçoso aceitar: a ciência da informação, por meio da biblioteconomia, principalmente, vem desenvolvendo metodologias para a organização da informação (os catálogos, por exemplo) e para a organização do conhecimento (as classificações, por exemplo), desde que isso se tornou uma preocupação humana. Por outro lado, a ciência da computação avançou consideravelmente

no trato da informação sintática (resolveu questões de fluxo e ajustou algumas propostas da catalogação para a descrição dos metadados) e desenvolveu metodologias e perspectivas interessantes para a representação e a organização do conhecimento, como as que oferecem as ontologias e sua interface com a *web* semântica. Em síntese: a construção de ontologias não pode prescindir do tratamento terminológico-conceitual, como compreendido pela terminologia e pela ciência da informação, acumulado nos referenciais teóricos e nas metodologias para construção de linguagens documentárias. A construção de linguagens documentárias mais flexíveis não pode ignorar o modelo de representação das ontologias, com mais predisposição para a formalização e para a interoperabilidade.

Naturalmente são problemas que os estudos futuros sobre representação do conhecimento e recuperação da informação deverão enfrentar. As redes de informação estão dadas. Tem-se a impressão, mas não se trata de nenhuma aposta, nem mesmo de um exercício de futurologia, de que em seu aspecto tecnológico não se poderá esperar um crescimento exponencial proporcional ao seu crescimento dos últimos dez anos.

A representação do conhecimento é o aspecto que une, por semelhança de propostas, a ciência da computação, por meio das ontologias, e a ciência da informação, por meio dos tesauros. Se há alguma identidade na origem da proposta, contudo, ela não existe no destinatário, que não é o mesmo. Os tesauros são construídos para funcionar como instrumento de representação do conhecimento e recuperação da informação em atividades de busca realizadas por seres humanos (especialistas). As ontologias são construídas para funcionar como instrumentos de representação do conhecimento e recuperação da informação em atividades de busca delegadas a agentes inteligentes.

O computador (o que computa) opera, como não poderia deixar de ser, de forma mais eficaz em ambientes quantitativos de recuperação de informação, justamente porque há nestes mais previsibilidade. Comparadas as funções de representação e recuperação da informação, tem-se que a segunda apresenta mais instabilidade e um número maior de variáveis. A relativa estabilidade atribuída à representação neste caso refere-se ao fato de que esta toma como ponto de partida interesses institucionais determinados e

estabelece, por este meio, o ponto de vista pelo qual o objeto será focalizado. Na recuperação da informação, o usuário parte do que Belkin (1980) denominou de *estado anômalo do conhecimento* e navega pelas estruturas de informação em busca do modo preciso como foi organizada a informação de que necessita. A atividade de seleção que um usuário não determinado qualquer realiza é, inerentemente, uma atividade intelectual e sempre inserida em contextos dinâmicos e, por este motivo, mais difícil de ser formalizada por meio de um algoritmo. O processo de representação do conhecimento em sistemas de informação documentária, sendo institucionalizados, são mais passíveis de mapeamento ontológico e de formalização.

Considerando-se que terminologia busca compreender os termos no uso, conforme a aplicação que se lhes dá nos contextos dos campos de especialidades e que as ontologias dependem de uma formalização de uma conceitualização compartilhada por especialistas em domínios relativamente determinados, é possível propor o tratamento terminológico na fase de organização das ontologias e, de outro lado, que as ontologias possam formalizar as terminologias (os usos) dos especialistas.

Na realização de busca por informação numa base de dados local, ou com alcance limitado, a presença ou a ausência de um determinado documento específico é um fator coercitivo. A capacidade das redes de elevar o percentual de revocação a números impensáveis, ao mesmo tempo em que lhe confere extremo dinamismo, torna o acesso ao documento uma eventualidade aberta. Em outras palavras, o usuário pode não ter o *documento*, mas dispõe de formas alternativas (rizoma) para adquirir a mesma informação.

Em síntese, os sistemas de informação documentária não podem prescindir das tecnologias desenvolvidas para o aperfeiçoamento das ontologias, pois perderiam, dessa forma o momento histórico, como perderam, num sentido vago, com os metadados. É preciso, então, pensá-los com as ontologias. Ajustar as metodologias desenvolvidas pela ciência da computação para os objetivos da ciência da informação e incorporar os fundamentos da ciência da informação aos procedimentos da ciência da computação no que tange ao universo dos fluxos da informação e do conhecimento.

## REFERÊNCIAS

- ALMEIDA, Maurício B.; BAX, Marcello P. Uma visão geral sobre ontologias: pesquisa sobre definições, tipos, aplicações, métodos de avaliação e de construção. **Ciência da Informação**, v. 32, n. 3, p. 7-20, set./dez., 2003. Disponível em: <<http://www.scielo.br/pdf/ci/v32n3/19019.pdf>>. Acesso em: 28 fev. 2010.
- ARAÚJO, Vânia Maria R. H. Sistemas de informação: nova abordagem teórico-conceitual. **Ciência da Informação**, v. 24, n. 1, p. 54-76, jan./abr., 1995. Disponível em: <<http://revista.ibict.br/ciinf/index.php/ciinf/article/view/532/484>>. Acesso em: 01 mar. 2010.
- ARMS, William Y. *et al.* A spectrum of interoperability: the site for science prototype for the NSDL. **D-Lib Magazine**, v. 8, n. 1, jan., 2002. Disponível em: <<http://www.dlib.org/dlib/january02/arms/01arms.html>>. Acesso em: 01 mar. 2010.
- AUBERT, Francis Henrik. Língua como estrutura e como fato histórico-social: consequências para a terminologia. In: ALVES, Ieda Maria (org.) **A constituição da normalização terminológica no Brasil**. 2.ed. São Paulo: FFLCH/CITRAT, 2001. p. 11-16
- BARRETO, Aldo A. A condição da informação. **São Paulo em Perspectiva**, v. 16, n. 3, p. 67-74, 2002.
- \_\_\_\_\_. A questão da informação. **São Paulo em Perspectiva**, v. 8, n. 4, 1994.
- \_\_\_\_\_. Pensando a pós-graduação em informação no Brasil. In: ENCONTRO NACIONAL DOS CURSOS DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO E BIBLIOTECONOMIA, 12, 1992, São Paulo. **Anais...** São Paulo: ANCIB, 1993.
- BARROS, Lídia A. **Curso básico de Terminologia**. São Paulo: Edusp, 2004.
- BELKIN, Nicholas J. Anomalous state of knowledge as a basis for information retrieval. **The Canadian Journal of Information Science**, v. 5, p. 133-140, 1980.
- BENVENISTE, Émile. **Problemas de lingüística geral II**. Campinas: Pontes, 1989.

BERNERS-LEE, T., HENDLER, J & LASSILA, O. The semantic web. **Scientific American**, may, 2001. Disponível em <<http://www.sciam.com/article.cfm?articleID=00048144-10D2-1C70-84A9809EC588EF21>> Acesso em: 15 out. 2007.

BICUDO, Maria Aparecida V.; TAMIRI, Leonor Maria; SAMPAIO, Helena. Educação superior: graduação e pós-graduação. In: Fundação de Amparo à Pesquisa do Estado de São Paulo. **Indicadores de ciência, tecnologia e inovação em São Paulo**. São Paulo: FAPESP, 2002. Disponível em: <<http://www.fapesp.br/indct/cap03/cap03.htm>> Acesso em: 04 out. 2006

BLAIR, David. Information retrieval and the philosophy of language. **Annual Review of Information Science and Technology (ARIST)**, v. 37, p. 3-50, 2003.

\_\_\_\_\_. **Wittgenstein, language and information: back to the rough ground!** Dordrecht: Springer, 2006.

BLOOR, David. **Wittgenstein: a social theory of knowledge**. London: MacMillan, 1983.

BRAGA, Maria Lúcia S. As três categorias peircianas e os três registros lacanianos. **Psicologia USP**, v. 10, n. 2, p. 81-91, 1999. Disponível em: <[http://www.scielo.br/scielo.php?pid=S0103-65641999000200006&script=sci\\_arttext](http://www.scielo.br/scielo.php?pid=S0103-65641999000200006&script=sci_arttext)>. Acesso em: 28 fev. 2010.

BREITMAN, Karin K.. **Web semântica: a internet do futuro**. Rio de Janeiro: LTC, 2005.

BURKE, Peter. **Uma história social do conhecimento: de Gutenberg a Diderot**. Rio de Janeiro: Jorge Zahar, 2003.

CABRÉ, Maria Teresa. La terminología hoy: concepciones, tendencias y aplicaciones. **Ciência da Informação**, v. 24, n. 3, 1995. Disponível em: <<http://revista.ibict.br/index.php/ciinf/article/view/487/442>>. Acesso em: 01 mar. 2010.

\_\_\_\_\_. **La terminología: representación y comunicación: elementos para una teoría de base comunicativa y otros artículos**. Barcelona: Universitat Pompeu Fabra, 1999.

\_\_\_\_\_. **La terminología: teoría, metodología, aplicaciones**. Barcelona: Antàrdida/Empúries, 1993.

\_\_\_\_\_; MOREL, Jordi; TEBÉ, Carles. Las relaciones conceptuales de tipo causal: un caso práctico. In: SIMPOSIO IBEROAMERICANO DE TERMINOLOGÍA, 5., 1996, Cidade do México. **Anais eletrônicos...** Cidade do México: RITerm, 1996. Disponível em: <<http://www.riterm.net/actes/5simposio/cabre6.htm>> Acesso em: 14 jul. 2008.

CAMPOS, Maria Luiza A. Modelização de domínios de conhecimento: uma investigação de princípios fundamentais. **Ciência da Informação**, Brasília, v. 33, n. 1, p. 22-32, jan./abr., 2004. Disponível em: <<http://revista.ibict.br/index.php/ciinf/article/viewFile/56/53>>. Acesso em: 01 mar. 2010.

\_\_\_\_\_. et al. O estado das pesquisas do Grupo - Ontologia e Taxonomias: aspectos teóricos e metodológicos. In: SEMINÁRIO DE PESQUISA EM ONTOLOGIA NO BRASIL, 2., 2009, Rio de Janeiro. **Anais...** Rio de Janeiro, 2009.

\_\_\_\_\_. et al. O uso de tesauro como base terminológica para a elaboração de ontologias de domínio: uma experiência com o domínio do Folclore e Cultura Popular. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO (ENANCIB), 9., 2008, São Paulo. **Anais...** São Paulo: Universidade de São Paulo, 2008.

CAPLAN, Priscilla. **Metadata fundamentals for all librarians**. Chicago: American Library Association, 2003.

CAPURRO, Rafael. Epistemologia e ciência da informação. Tradução: Ana Maria Rezende Cabral, Eduardo Wense Dias, Isis Paim, Ligia Maria Moreira Dumont, Marta Pinheiro Aun e Mônica Erichsen Nassif Borges. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 5., 2003, Belo Horizonte. **Anais...** Belo Horizonte: Escola de Ciência da Informação, UFMG, 2003.

CHARTIER, Roger. **A ordem dos livros: leitores, autores e bibliotecas na Europa entre os séculos XIV e XVII**. Brasília: UnB, 1994.

CHAUÍ, Marilena S. **Convite à filosofia**. 7.ed. São Paulo: Ática, 2000.

CINTRA, Anna Maria M. Elementos de linguística para estudos de indexação. **Ciência da Informação**, v. 12, n. 1, p. 5-22, 1983. Disponível em: <<http://revista.ibict.br/index.php/ciinf/article/view/1526/1144>>. Acesso em: 01 mar. 2010.

\_\_\_\_\_. et al. Linguagem documentária e terminologia. In: ALVES, Ieda Maria (org.) **A constituição da normalização terminológica no Brasil**. 2.ed. São Paulo: FFLCH/CITRAT, 2001. p. 17-22

COELHO NETTO, José Teixeira. O autor, ainda. **Imagens**, n. 3, p. 69-73, dez. 1994.

COSTA, Leonardo C.; MARCONDES, Carlos Henrique. Um ambiente para edição, extração e representação do conhecimento contido em artigos científicos publicados na web. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO (ENANCIB), 9., 2008, São Paulo. **Anais...** São Paulo: Universidade de São Paulo, 2008.

DAHLBERG, Ingetraut. Teoria do conceito. **Ciência da Informação**, Rio de Janeiro, v. 7, n. 2, p. 101-107, 1978. Disponível em: <<http://revista.ibict.br/index.php/ciinf/article/viewFile/1680/1286>>. Acesso em: 01 mar. 2010.

DELEUZE, Gilles. **Conversações: 1972-1990**. Rio de Janeiro: 34, 1992.

\_\_\_\_\_. **Diferença e repetição**. Rio de Janeiro: Graal, 1988.

\_\_\_\_\_; GUATTARI, Felix. **Mil platôs: capitalismo e esquizofrenia: vol. 1.** Rio de Janeiro: 34, 1995.

\_\_\_\_\_; \_\_\_\_\_. **Mil platôs: capitalismo e esquizofrenia: vol. 4.** Rio de Janeiro: 34, 1997.

\_\_\_\_\_; \_\_\_\_\_. **Mil platôs: capitalismo e esquizofrenia.** Rio de Janeiro: 34, 1995-1997. 715 p. Resenha de: ABREU FILHO, Ovídio. **Mana**, v. 4, n. 2, p. 143-146, 1998.

\_\_\_\_\_; PARNET, Claire. **Diálogos.** São Paulo : Escuta, 1998.

ECHEVERRÍA, Rafael. **Ontologia del lenguaje.** 7.ed. Santiago: J.C. SÁEZ, 2003.

ECO, Umberto. **Conceito de texto.** São Paulo: Edusp, 1984.

FERNEDA, Edberto. **Recuperação da informação: análise sobre a contribuição da ciência da computação para a ciência da informação.** Tese (Doutorado em Ciências da Comunicação) – Universidade de São Paulo, Escola de Comunicações e Artes, 2003.

FERRATER MORA, José. **Dicionário de Filosofia.** 4.ed. Lisboa: Dom Quixote, 1978.

FINATTO, Maria José B. Terminologia e ciência cognitiva. In: \_\_\_\_\_. MACIEL, Anna Maria B. (Orgs.). **Temas de Terminologia.** Porto Alegre: Editora da UFRGS, 2001. p. 141-149.

FISCHER, D. H. From thesauri towards ontologies? In: STRUCTURES AND RELATIONS IN KNOWLEDGE ORGANIZATION: International ISKO Conference, 5., 1998, Lille. **Proceedings...** Würzburg: Ergon, 1998. Disponível em: [http://www.ipsi.fraunhofer.de/topas/publications/Fischer\\_1998.pdf](http://www.ipsi.fraunhofer.de/topas/publications/Fischer_1998.pdf) Acesso em: 03 fev. 2010.

FRANCHINI, A. S.; SEGANFREDO, Carmen. **As 100 melhores histórias da mitologia: deuses, heróis, monstros e guerras da tradição greco-romana.** 9.ed. Porto Alegre: LP&M, 2007.

FREITAS, K. F et al. Tesouro como base terminológica para a elaboração de ontologia de domínio: uma aplicação no domínio do Folclore e Cultura Popular. In: SEMINÁRIO DE PESQUISA EM ONTOLOGIA NO BRASIL, 1., 2008, Niterói. **Anais...** Niterói, 2008.

GARCÍA GUTIÉRREZ, A. L. **Estructura lingüística de la documentación: teoría y método.** Murcia: Universidad de Murcia, 1990.

GARSHOL, Lars Marius. Metadata? Thesauri? Taxonomies? Topic Maps! Making sense of it all. **Journal of Information Science**, v. 30, n. 4, p. 378-391, 2004.

GIBSON, William. **Neuromancer.** São Paulo: Aleph, 1991.

GÓMEZ-PÉREZ, Asunción. Ontological engineering: a state of the art. **Expert Update**, v. 2, n. 3, p. 33-43, 1999.

GRUBER, Thomas R. Toward principles for the design of ontologies used for knowledge sharing. **International Journal Human-Computer Studies**, v. 43, n. 5/6, p. 907-928, 1995. Disponível em: <<http://www.itee.uq.edu.au/~infs3101/Readings/OntoEng.pdf>>. Acesso em: 28 fev. 2010.

GUARINO, Nicola. The ontological level. In: CASI, R; SMITH, B; WHITE, G. (Eds.). **Philosophy and the cognitive science**. Viena: Holder-Pichler-Tempsky, 1995.

\_\_\_\_\_; GIARETTA, Pierdaniele. Ontologies and knowledge bases: towards a terminological clarification. In: MARS, N. (ed.) **Towards very large knowledge bases: knowledge building and knowledge sharing**. Amsterdam: IOS, 1995. p. 25-32. Disponível em: <<http://www.loa-nr.it/Papers/KBKS95.pdf>>. Acesso em: 10 dez. 2006.

HAMMAN, Robin B. Rhizome@internet: using the internet as an example of Deleuze and Guattari's "rhizome". 1996. Disponível em: <<http://www.swinburne.infoexchange.net.au/media/halm316/gallery/david/pg11b.htm>>. Acesso em: 01 mar. 2010.

ILARI, Rodolfo. O estruturalismo linguístico: alguns caminhos. MUSSALIM, Fernanda; BENTES, Anna Christina (Orgs.). **Introdução à linguística: fundamentos epistemológicos**, vol. 3. São Paulo: Cortez, 2004. p. 53-91.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. ISO 704: terminology work – principles and methods. Geneva, 2000.

IYER, Hemalata. Cognition and categories. In: \_\_\_\_\_. **Classificatory structures: concepts, relation and representation**. Frankfurt: Indeks Verlag, 1995. p. 40-59.

JACOB, Elin K. Ontologies and the semantic web. **Bulletin of the American Society for Information Science and Technology**, apr./may, p. 19-22, 2003.

KALFOGLOU, Y.; SCHORLEMMER, M. IF-Map: an ontology mapping method based on Information Flow theory. **Journal on Data Semantics**, v. 1, n. 1, p. 98-127, 2003.

KANT, Immanuel. **Crítica da razão pura**. São Paulo: Nova Cultural, 1996. (Os Pensadores).

KOBASHI, Nair Y. **A elaboração de informações documentárias: em busca de uma metodologia**. Tese (Doutorado em Ciências da Comunicação) – Departamento de Biblioteconomia e Documentação, Escola de Comunicações e Artes, Universidade de São Paulo, São Paulo, 1994.

\_\_\_\_\_; SMIT, Johanna W.; TÁLAMO, Maria de Fátima G. M. A função da terminologia na construção do objeto da Ciência da Informação. **DataGramaZero**, v. 2, n. 2, abr., 2001. Disponível em: <[http://www.datagramazero.org.br/abr01/Art\\_03.htm](http://www.datagramazero.org.br/abr01/Art_03.htm)> Acesso em: 22 fev. 2005.

KOSTINA, Irina. La manifestación del contenido del término en un mismo ámbito profesional desde perspectivas distintas: en búsqueda del patrón lingüístico. In: JORNADA INTERNACIONAL SOBRE LA INVESTIGACIÓN EN TERMINOLOGÍA Y CONOCIMIENTO ESPECIALIZADO, 2001, Barcelona. **Anais...** Barcelona: Universitat Pompeu Fabra, 2003. p. 26-33.

KRIEGER, Maria da Graça. O termo: questionamentos e configurações. In: \_\_\_\_\_; MACIEL, Anna Maria B. (Orgs.). **Temas de Terminologia**. Porto Alegre: Editora da UFRGS, 2001a. p. 62-81.

\_\_\_\_\_. Terminologia em contextos de integração: funcionalidade e fundamentos. In: \_\_\_\_\_; MACIEL, Anna Maria B. (Orgs.). **Temas de Terminologia**. Porto Alegre: Editora da UFRGS, 2001b. p. 214-225.

\_\_\_\_\_; FINATTO, Maria José B. **Introdução à terminologia**: teoria e prática. São Paulo: Contexto, 2004.

KUHN, Thomas S. **A estrutura das revoluções científicas**. 2.ed. São Paulo: Perspectiva, 1978.

LANCASTER, Frederick W. **El control del vocabulário em la recuperación de información**. 2.ed. València: Universitat de València, 2002.

LARA, Marilda L. G. Linguagem documentária e terminologia. **Transinformação**, v. 16, n. 3, p. 231-240, set./dez. 2004. Disponível em: <<http://revistas.puc-campinas.edu.br/transinfo/include/getdoc.php?id=217&article=72&mode=pdf>>. Acesso em: 01 mar. 2010.

\_\_\_\_\_. **Linguística documentária**: seleção de conceitos. 2009. 178 f. Tese (Livre-docência) – Escola de Comunicações e Artes, Universidade de São Paulo, São Paulo, 2009.

\_\_\_\_\_. O unicórnio (o rinoceronte, o ornitorrinco...), a análise documentária e a linguagem documentária. **DataGramaZero**, v. 2, n. 6, dez. 2001. Disponível em <[http://www.datagramazero.org.br/dez01/Art\\_03.htm](http://www.datagramazero.org.br/dez01/Art_03.htm)>. Acesso em: 22 fev. 2005.

\_\_\_\_\_. **Representação documentária**: em jogo a significação. 1993. 133 f. Dissertação (Mestrado em Ciências da Comunicação) – Universidade de São Paulo, Escola de Comunicações e Artes, Departamento de Biblioteconomia e Documentação, 1993.

LECHTE, John. **50 pensadores contemporâneos essenciais**: do Estruturalismo à Pós-Modernidade. Rio de Janeiro: Difel, 2002.

LOPES, Edward. **Fundamentos da linguística contemporânea**. 19 ed. São Paulo: Cultrix, 2005.

LOPES, Maria Immacolata V. **Pesquisa em comunicação**. 6.ed. São Paulo: Loyola, 2001.

MAEDCHE, Alexander. **Ontology learning for the semantic web**. Boston: Kluwer Academic Publisher, 2002.

MAI, Jens-Erik. The future of general classification. **Cataloging & Classification Quarterly**, v. 37, n. 1/2, p. 3-12, 2003.

MARCONDES, Carlos Henrique et al. Representing and coding the knowledge embedded in texts of health science web published articles. In: CONFERENCE ON ELECTRONIC PUBLISHING, 2007, Vienna (Austria). **Proceedings...** Vienna: [s.n.], 2007.

\_\_\_\_\_. From scientific communication to public knowledge: the scientific article web published as a knowledge base. In: CONFERENCE ON ELECTRONIC PUBLISHING, 9., 2005, Leuven (Belgium) **Proceedings...** Leuven: Katholieke Universiteit Leuven, 2005.

\_\_\_\_\_; CAMPOS, Maria Luiza A. Ontologia e web semântica: o espaço da pesquisa em ciência da informação. **PontodeAcesso**, v. 2, n. 1, p. 107-136, jun./jul., 2008.

\_\_\_\_\_; SAYÃO, Luís Fernando. Integração e interoperabilidade no acesso a recursos informacionais eletrônicos em C&T: a proposta da Biblioteca Digital Brasileira. **Ciência da Informação**, v. 30, n. 3, p. 24-33, set./dez., 2001. Disponível em: <http://revista.ibict.br/index.php/ciinf/article/viewFile/190/167>>. Acesso em: 01 mar. 2010.

MOREIRA, Alexandra. **Tesouros e ontologias**: estudo de definições presentes na literatura das áreas das ciências da computação e da informação, utilizando-se o método analítico-sintético. 2003. 150 f. Dissertação (Mestrado em Ciência da Informação) – Universidade Federal de Minas Gerais, Escola de Ciência da Informação, Belo Horizonte, 2003.

\_\_\_\_\_; ALVARENGA, Lídia; OLIVEIRA, Alcione P.. O nível do conhecimento e os instrumentos de representação: tesouros e ontologias. **DataGramaZero**: revista de Ciência da Informação, v. 5, n. 6, dez. 2004. Disponível em [http://www.dgz.org.br/dez04/Art\\_01.htm](http://www.dgz.org.br/dez04/Art_01.htm)> Acesso em: 20 nov. 2006.

MOREIRA, Walter. **Bibliotecas tradicionais x bibliotecas virtuais**: modelos de recuperação da informação. 1998. 113 f. Dissertação (Mestrado em Biblioteconomia e Ciência da Informação) – Departamento de Pós-graduação em Biblioteconomia, Pontifícia Universidade Católica de Campinas, Campinas, 1998. Disponível em: [http://eprints.rclis.org/7670/1/BibliotecaTradicionalXBibliotecaVirtual\\_ModelosDeRecuperacaoDaInformacao.pdf](http://eprints.rclis.org/7670/1/BibliotecaTradicionalXBibliotecaVirtual_ModelosDeRecuperacaoDaInformacao.pdf)>. Acesso em: 28 fev. 2010.

\_\_\_\_\_. Lexicologia, terminologia, ontologia e representação documentária: estudos de interface por meio de análise de periódicos de Ciência da Informação. **Biblios** (Lima), v. 8, p. 4, 2007. Disponível em [http://www.bibliosperu.com/articulos/27/27\\_06.pdf](http://www.bibliosperu.com/articulos/27/27_06.pdf)> Acesso em: 02 fev. 2008.

\_\_\_\_\_. Os colégios virtuais e a nova configuração da comunicação científica. **Ciência da Informação**, Brasília, v. 34, n. 1, p.57-63, jan./abr. 2005. Disponível em: <<http://revista.ibict.br/index.php/ciinf/article/view/635/564>>. Acesso em: 28 fev. 2010.

NAHRA, Cinara; WEBER, Ivan H. **Através da lógica**. 4.ed. Petrópolis: Vozes, 2001.

NASCIMENTO, Marta S. O. et al. A ontologia na ciência da informação. **Revista Digital de Biblioteconomia e Ciência da Informação**, v. 5, n. 1, p. 13-39, jul./dez., 2007. Disponível em: <<http://www.sbu.unicamp.br/seer/ojs/viewarticle.php?id=105>> Acesso em: 28 fev. 2010.

NATIONAL INFORMATION STANDARDS ORGANIZATION (NISO). **Understanding metadata**. Bethesda: NISO Press, 2004. Disponível em: <<http://www.niso.org/publications/press/UnderstandingMetadata.pdf>> Acesso em: 22 jan. 2010.

OTLET, Paul. **El tratado de documentación**: el libro sobre el libro: teoría y práctica. Tradução Maria Dolores A. García. Murcia: Editum, 1996.

PEDAUQUE, Rogelio T. **Documento**: forma, signo y medio, reformulaciones de lo digital. 2004. Disponível em: <[http://archivesic.ccsd.cnrs.fr/docs/00/06/24/58/PDF/sic\\_00001160.pdf](http://archivesic.ccsd.cnrs.fr/docs/00/06/24/58/PDF/sic_00001160.pdf)> Acesso em: 29 jan. 2010.

PEIRCE, Charles S. On a new list of categories. **Proceedings of the American Academy of Arts and Sciences**, n. 7, p. 287-298, 1868. Disponível em <<http://www.peirce.org/writings/p32.html>> Acesso em: 13 dez. 2006.

PEROJO, Keilyn R.; LEON, Rodrigo R. Web semántica: un nuevo enfoque para la organización y recuperación de información en el web. **Acimed**, Ciudad de La Habana, v. 13, n. 6, 2005. Disponível em: <<http://scielo.sld.cu/pdf/aci/v13n6/aci030605.pdf>> Acesso em: 15 ago. 2008.

PIAGET, Jean. **O Estruturalismo**. 3.ed. São Paulo: Difel, 1979.

PINHEIRO, Lena Vânia R.; LOUREIRO, José Mauro Matheus. Traçados e limites da ciência da informação. **Ciência da Informação**, v. 24, n. 1, p. 42-53, 1995. Disponível em: <[http://www.dabiblioufma.hpg.com.br/documentacao/art04Pinheiro\\_Loureiro.pdf](http://www.dabiblioufma.hpg.com.br/documentacao/art04Pinheiro_Loureiro.pdf)>. Acesso em: 28 fev. 2010.

POMBO, Olga. Da classificação dos seres à classificação dos saberes. **Leituras**: revista da Biblioteca Nacional de Lisboa, n. 2, p. 19-33, primavera, 1998. Disponível em: <<http://www.educ.fc.ul.pt/hyper/resources/opombo-classificacao.pdf>> Acesso em: 09 nov. 2006.

PORFÍRIO DE TIRO. **Isagoge**: introdução às categorias de Aristóteles. Trad. e com. Bento Silva Santos. São Paulo: Attar, 2002.

QIN, Jian; PALING, Stephen. Converting a controlled vocabulary into an ontology: the case of GEM. *Information Research*, v. 6, n. 2, 2001. Disponível em: <<http://informationr.net/ir/6-2/paper94.html>>. Acesso em: 01 mar. 2010.

REITZ, Joan M. **Dictionary for Library and Information Science**. Westport: Libraries Unlimited, 2004.

RODRIGUES, Adriano D. **Comunicação e cultura**: a experiência cultural na era da informática. Lisboa: Presença, 1994.

ROJAS, Miguel Ángel R. Relación entre los conceptos: información, conocimiento y valor: semejanzas y diferencias. **Ciência da Informação**, Brasília, v. 34, n. 2, p. 52-61, maio/ago., 2005. Disponível em: <<http://revista.ibict.br/index.php/ciinf/article/view/628/560>>. Acesso em: 01 mar. 2010.

ROUDINESCO, Elisabeth. Giles Deleuze: variações antiedipianas. In: \_\_\_\_\_. **Filósofos na tormenta**: Canguilhem, Sartre, Foucault, Althusser, Deleuze e Derrida. Rio de Janeiro: Zahar, 2007. p. 204-217.

RUSSELL, Bertrand. **História do pensamento ocidental**: a aventura dos pré-socráticos a Wittgenstein. Rio de Janeiro: Ediouro, 2001.

SAFRANSKI, Rüdiger. **Nietzsche**: biografia de uma tragédia. São Paulo: Geração Editorial, 2005.

SALATIEL, José Renato. Peirce e Kant sobre categorias: parte I: dedução metafísica e reviravolta semiótica. **Cognitio-Estudos**: Revista Eletrônica de Filosofia, v. 3, n. 1, p. 79-88, jan./jun. 2006. Disponível em <[http://www.pucsp.br/pos/filosofia/Pragmatismo/cognitio\\_estudos/cog\\_estudos\\_v3n1/cog\\_est\\_v3\\_n1\\_luisi\\_maria\\_t07\\_65\\_70.pdf](http://www.pucsp.br/pos/filosofia/Pragmatismo/cognitio_estudos/cog_estudos_v3n1/cog_est_v3_n1_luisi_maria_t07_65_70.pdf)> Acesso em: 30 nov. 2006.

SALES, Rodrigo de; CAFÉ, Lígia. Os tesouros e as ontologias sob um olhar terminológico. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO (ENANCIB), 9., 2008, São Paulo. **Anais...** São Paulo: Universidade de São Paulo, 2008

SANTOS, Bento S. Apresentação. In: PORFÍRIO DE TIRO. **Isagoge**: introdução às categorias de Aristóteles. Tradução e comentários: Bento Silva Santos. São Paulo: Attar, 2002. p. 9-16.

SANTOS, Boaventura S. **Introdução a uma ciência pós-moderna**. 4.ed. Rio de Janeiro: Graal, 2003.

\_\_\_\_\_. **Um discurso sobre as ciências**. 13.ed. Porto: Afrontamento, 2002.

SARACEVIC, Tefko. Ciência da informação: origem, evolução e relações. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 1, n. 1, p. 41-62, jan./jun. 1996. Disponível em:

<<http://www.eci.ufmg.br/pcionline/index.php/pci/article/viewFile/235/22>>. Acesso em: 01 mar. 2010.

SARAMAGO, José. **Todos os nomes**. São Paulo: Companhia das Letras, 1998.

SAYÃO, Luis Fernando; MARCONDES, Carlos Henrique. O desafio da interoperabilidade e as novas perspectivas para as bibliotecas digitais. **Transinformação**, Campinas, v. 20, n. 2, p. 133-148, maio/ago., 2008. Disponível em: <<http://revistas.puc-campinas.edu.br/transinfo/viewarticle.php?id=215>>. Acesso em: 01 mar. 2010.

SHANNON, Claude E. A mathematical theory of communication. **The Bell System Technical Journal**, v. 27, jul./oct., p. 379–423, 623–656, 1948. Disponível em: <<http://plan9.bell-labs.com/cm/ms/what/shannonday/shannon1948.pdf>> Acesso em: 18 jun. 2008.

SILVA, Daniela L.; SOUZA, Renato R.; ALMEIDA, Maurício B. Uma comparação de metodologias para construção de ontologias e vocabulários controlados. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO (ENANCIB), 9., 2008, São Paulo. **Anais...** São Paulo: Universidade de São Paulo, 2008.

SMIT, Johanna. Ciência da Informação: definições, missões e desafios. 20 jun. 2007. [Notas de aula].

SMITH, Barry. Ontology. In: FLORIDI, Luciano (Ed.). **The blackwell guide to the philosophy of computing and information**. Malden: Blackwell, 2004. p. 155-166.

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

STEIN, Ernildo. Vida de obra. In: HEIDEGGER, Martin. **Conferências e escritos filosóficos**. São Paulo: Nova Cultural, 2000. p. 5-10.

TAKAHASHI, Tadao (Org.). **Sociedade da informação no Brasil: livro verde**. Brasília: Ministério da Ciência e Tecnologia, 2000. Disponível em <<http://www.mct.gov.br/index.php/content/view/18940.html>>. Acesso em: 22 nov. 2008.

TÁLAMO, Maria de Fátima G. M. et al. Instrumentos de controle terminológico: limites e funções. In: SIMPÓSIO LATIIONAMERICANO DE TERMINOLOGIA, 2., 1992, Brasília. **Anais...** Brasília: União Latina/ IBICT, 1992. Disponível em: <<http://www.riterm.net/actes/2simposio/talamo.htm>> Acesso em: 12 fev. 2010.

\_\_\_\_\_; LARA, Marilda L. G. O campo da Linguística Documentária. **Transinformação**, Campinas, v. 18, n. 3, p. 203-211, set./dez., 2006. Disponível em: <<http://revistas.puc-campinas.edu.br/transinfo/viewarticle.php?id=182>>. Acesso em: 01 mar. 2010.

VAN DER LAAN, Regina Helena; FERREIRA, Glória Isabel S. Tesouros e terminologia. In: CONGRESSO BRASILEIRO DE BIBLIOTECONOMIA E DOCUMENTAÇÃO, 19, 2000, Porto Alegre. **Anais Eletrônicos...** Porto Alegre, 2000.

VIDAL, Maria Rosária R.; VIDAL, Waldomiro. **Botânica-organografia**: quadros sinóticos ilustrados de fanerógamos. Viçosa: UFV, 2003.

VOGT, Carlos. Divulgação e cultura científica. **ComCiência**, n. 100, 10 jul. 2008. Editorial. Disponível em: <<http://www.comciencia.br/comciencia/?section=8&edicao=37&id=436>> Acesso em: 10 jul. 2008.

VYGOTSKY, Lev S. **Pensamento e linguagem**. 3.ed. São Paulo: Martins Fontes, 1991.

# APÊNDICES

## APÊNDICE A – ARTIGOS SOBRE ONTOLOGIA EM PERIÓDICOS INTERNACIONAIS DE CIÊNCIA DA INFORMAÇÃO

PERIÓDICO	2004		2005		2006		2007		2008		Total	
											N	%
AI Magazine			v. 26, n. 1	5							7	2,68
	v. 25, n. 3	1	v. 26, n. 3	1								
AI Communications					v. 19, n. 1	1					4	1,53
			v. 18, n. 3	1								
	v. 17, n. 4	2										
Anales de Documentacion	v. 7	1									1	0,38
Applied Intelligence					v. 24, n. 3	1					1	0,38
Artificial Intelligence			v. 165, n. 2	1							1	0,38
Artificial Intelligence and Law	v. 12, n. 4	7									7	2,68
Artificial Intelligence in Medicine	v. 32, n. 1	1			v. 36, n. 1	1					5	1,92
					v. 36, n. 2	1						
			v. 34, n. 3	1	v. 37, n. 3	1						
Aslib Proceedings: New Information Perspectives									v. 60, n. 1	1	1	0,38
Biblioteche Oggi			v. 23, n. 5	1							1	0,38
BiD: Textos Universitaris de Biblioteconomia i Documentacio									n. 20	1	1	0,38
Bollettino AIB	v. 44, n. 2	1									1	0,38
Bulletin of Library and Information Science					n. 56	1					1	0,38

Cataloging & Classification Quarterly	v. 37, n. 1/2	1						2	0,77
			v. 39, n. 3/4	1					
CIT - Journal of Computing and Information Technology			v. 13, n. 4	1				1	0,38
Documentaliste - Sciences de l'Information							v. 45, n. 3	1	1 0,38
Electronic Library						v. 25, n. 2	1		1 0,38
Expert Systems	v. 21, n. 4	1							1 0,38
Faslname-Ye Ketab/Library and Information Studies						v. 18, n. 2	1		1 0,38
First Monday						v. 12, n. 12	1	v. 13, n. 8	1 2 0,77
IEEE Intelligent Systems	v. 19, n. 1	2			v. 21, n. 1	3		v. 23, n. 1	1
			v. 20, n. 2	1					
	v. 19, n. 3	2			v. 21, n. 3	1			
	v. 19, n. 4	6							
			v. 20, n. 5	1	v. 21, n. 5	1			
	v. 19, n. 6	1	v. 20, n. 6	1	v. 21, n. 6	1			
Informacijos Mokslai/Information Sciences			n. 34	1					1 0,38
Informatie Professional			v. 9, n. 2	1					
	v. 8, n. 10	1							
					v. 10, n. 12	1			3 1,15
Information and Software Technology					v. 48, n. 7	1			
			v. 47, n. 12	1					2 0,77
Information Management and Computer Securit							v. 16, n. 2	1	1 0,38
Information Processing and Management					v. 43, n. 4	3	v. 44, n. 2	1	
							v. 44, n. 5	1	5 1,92
Information Research					v. 12, n. 2	1			1 0,38
Information Studies						v. 13, n. 1	1		
					v. 12, n. 3	1			
					v. 12, n. 4	1			
Information Systems	v. 29, n. 1	1							
					v. 31, n. 4/5	1			
			v. 30, n. 5	1					
							v. 32, n. 6	1	
			v. 30, n. 7	1					
							v. 32, n. 8	1	
Information Visualization					v. 5, n. 4	1			1 0,38

Information: Wissenschaft & Praxis									v. 59, n. 2		2	11	4,21		
	v. 55, n. 3								1						
			v. 57, n. 5		1	v. 58, n. 5		1							
			v. 56, n. 5/6		1										
					v. 57, n. 6/7			5							
Intelligent Information Systems							v. 3, n. 1		1			1	0,38		
Interacting with Computers	v. 16, n. 3	1	v. 17, n. 3	1							2	0,77			
International Journal of Data Warehousing and Mining									v. 4, n. 1		1	1	0,38		
International Journal of Distance Education Technologies			v. 3, n. 1		1			v. 5, n. 1		1			3	1,15	
			v. 3, n. 2		1										
International Journal of Electronic Government Research									v. 4, n. 1		1	1	0,38		
International Journal of Human-Computer Studies	v. 60, n. 1	1			v. 64, n. 1		1					8	3,07		
					v. 62, n. 5		4								
					v. 64, n. 10		1								
					v. 64, n. 11		1								
International Journal of Information Management			v. 25, n. 1		1							1	0,38		
International Journal of Web Information Systems			v. 1, n. 2		1					v. 4, n. 2		2	3	1,15	
International Journal of Web Services Research			v. 2, n. 2		1	v. 3, n. 2		1					5	1,92	
			v. 2, n. 3		2										
					v. 3, n. 4			1							
Internet Research: Electronic Networking Applications and Policy					v. 16, n. 5		1					1	0,38		
Item					n. 42		3					3	1,15		
Journal of Agricultural & Food Information					v. 6, n. 4		1					1	0,38		
Journal of Database Management					v. 17, n. 1		1					6	2,30		
	v. 15, n. 2						4			v. 19, n. 4				1	
Journal of Enterprise Information Management					v. 19, n. 2		1					1	0,38		
Journal of Information Science									v. 34, n. 1		1	4	1,53		
									v. 34, n. 2		1				
					v. 32, n. 5				1						
									v. 34, n. 6					1	

Journal of Information Science and Technology Association							v. 57, n. 5		1				2	0,77			
										v. 58, n. 7		1					
Journal of Information Technology					v. 21, n. 4		1						1	0,38			
Journal of Knowledge Management							v. 11, n. 1		1				5	1,92			
			v. 9, n. 5		3					v. 12, n. 6		1					
Journal of the American Society for Information Science and Technology			v. 56, n. 1		1					v. 59, n. 1		1	12	4,60			
	v. 55, n. 3		1														
			v. 56, n. 5		1			v. 59, n. 5		1							
	v. 55, n. 7		1			v. 58, n. 6		1									
									v. 59, n. 9		1						
			v. 56, n. 11		1					v. 59, n. 11		1					
										v. 59, n. 12		1					
										v. 59, n. 13		1					
Journal of the China Society for Scientific and Technical Information	v. 23, n. 1		1	v. 24, n. 1		1	v. 25, n. 1		1	v. 26, n. 1		2	v. 27, n. 1		1	33	12,64
	v. 23, n. 2		1	v. 24, n. 2		1	v. 25, n. 2		3				v. 27, n. 2		2		
	v. 23, n. 3		1				v. 25, n. 3		2	v. 26, n. 3		1	v. 27, n. 3		1		
	v. 23, n. 4		1				v. 25, n. 4		1	v. 26, n. 4		2	v. 27, n. 5		1		
										v. 26, n. 5		1					
	v. 23, n. 6		2	v. 24, n. 6		2	v. 25, n. 6		3				v. 27, n. 6		2		
Knowledge-based systems							v. 20, n. 1		2				16	6,13			
					v. 19, n. 2		1			v. 21, n. 2		1					
										v. 21, n. 3		1					
			v. 19, n. 4		1												
			v. 19, n. 5		4												
	v. 18, n. 6		1														
	v. 18, n. 7		1	v. 19, n. 7		1											
					v. 19, n. 8		1			v. 21, n. 8		2					
Knowledge Organization			v. 32, n. 1		1								5	1,92			
	v. 31, n. 3		1			v. 33, n. 3		1									
	v. 31, n. 4		1						v. 35, n. 4		1						
Library Management			v. 26, n. 4		2								2	0,77			
Library Resources & Technical Services			v. 49, n. 4		2								2	0,77			
Library Trends	v. 52, n. 3		1											1	0,38		
Machine Translation			v. 19, n. 2		1								1	0,38			
Medical Informatics & The Internet in Medicine			v. 30, n. 2		1	v. 31, n. 2		1					2	0,77			

Nauchno-tekhnicheskaya informatsiya					n. 1	1				15	5,75
			n. 2	1	n. 2	2					
							n. 3	1			
			n. 4	1							
	n. 5	1									
			n. 6	1							
					n. 8	1					
			n. 9	2							
			n. 10	1							
			n. 11	2							
			n. 12	1							

New Review of Hypermedia and Multimedia	v. 10, n. 1	1							1	0,38		
Online Information Review					v. 30, n. 3	1				1	0,38	
Profesional de la Informacion						v. 16, n. 6	5	v. 17, n. 1	3	8	3,07	
Revista Interamericana de Bibliotecologia					v. 29, n. 2	1				1	0,38	
Serials Librarian	v. 46, n. 3/4	1							1	0,38		
Signum	n. 3	1							5	1,92		
											n. 5	2
	n. 8	2										
The Learning Organization			v. 12, n. 5	3					3	1,15		
Toshokan-Kai (The Library World)					v. 58, n. 2	1				1	0,38	
Tudomanyos es Muszaki Tajekoztatas	v. 51, n. 5	1								2	0,77	
						v. 54, n. 10	1					
VINE: The Journal of Information and Knowledge Management Systems								v. 38, n. 4	1	1	0,38	
Webology	v. 1, n. 2	1				v. 4, n. 2	1			3	1,15	
				v. 3, n. 3	1							
Zagadnienia Informacji Naukowej	v. 2, n. 84	1							1	0,38		

55	66	66	32	42	26
2004	2005	2006	2007	2008	1

## APÊNDICE B – GRUPOS DE PESQUISA SOBRE ONTOLOGIAS NO BRASIL

Obs: O quadro apresentado a seguir é resultado de busca realizada no Diretório dos Grupos de Pesquisa no Brasil, do CNPq. Não houve intenção de realizar uma busca exaustiva, mas apenas localizar alguns grupos de pesquisa cuja produção pode interessar aos estudos sobre ontologias na ciência da informação. É possível que alguns grupos tenham sido omitidos, pois verificou-se inconsistência no sistema, fator impeditivo da precisão na busca.

Nome: <b>Redes de conhecimento e informação</b>
Instituição: Universidade Estadual de Londrina – Centro de Educação Comunicação e Artes
Líder: Maria Inês Tomaél
Linhas de pesquisa: Compartilhamento da informação e do conhecimento. Uso de ontologias para a representação do conhecimento.
Nome: <b>Ontologia e taxonomia: aspectos teóricos e metodológicos</b>
Instituição: Universidade Federal Fluminense – Departamento de Ciência da Informação
Líder: Maria Luiza de Almeida Campos
Linhas de pesquisa: <i>Datawarehouse</i> . Fluxos e mediações sócio-técnicas da informação.
Nome: <b>Estudos e pesquisa em mediação e comunicação da informação</b>
Instituição: Universidade Federal da Bahia – Departamento de Documentação e Informação
Líderes: Henriette Ferreira Gomes e Maria Isabel de Jesus Sousa
Linhas de pesquisa: Informação e contextos sócio-econômicos.
Nome: <b>Educação, conhecimento e tecnologia</b>
Instituição: Universidade Federal do Rio Grande – Instituto de Ciências Humanas e da Informação
Líder: Gisele Vasconcelos Dziekaniak
Linhas de pesquisa: Informação, educação e sociedade. Organização do conhecimento.
Nome: <b>Gestão da informação e do conhecimento</b>
Instituição: Universidade Federal de Minas Gerais - Escola de Ciência da Informação
Líder: Ricardo Rodrigues Barbosa e Marta Macedo Kerr Pinheiro
Linhas de pesquisa: Ontologias organizacionais.
Nome: <b>Fenomenologia do conhecimento e antropologia filosófica</b>
Instituição: Pontifícia Universidade Católica do Rio Grande do Sul – Departamento de Pós-Graduação em Filosofia

Líder: Ernildo Jacob Stein e Luiz Carlos Santuário
Linhas de pesquisa: Ontologia, lógica e linguagem.
Nome: <b>Ética e Metafísica</b>
Instituição: Universidade Federal do Rio Grande do Sul – Instituto de Filosofia e Ciências Humanas
Líder: Balthazar Barbosa Filho
Linhas de pesquisa: Estudos de lógica e ontologia.
Nome: <b>Engenharia de software</b>
Instituição: Universidade de São Paulo – Instituto de Ciências Matemáticas e de Computação
Líder: Paulo Cesar Masiero
Linhas de pesquisa: Ontologias.
Nome: <b>Dataware</b>
Instituição: Universidade Federal do Rio de Janeiro – Departamento de Ciência da Computação
Líder: Maria Luiza Machado Campos e Paulo de Figueiredo Pires
Linhas de pesquisa: Gestão do conhecimento. Ontologias.
Nome: <b>A construção de objetos nas ciências humanas</b>
Instituição: Universidade de Caxias do Sul – Departamento de Filosofia
Líder: Sofia Inês Albornoz Stein
Linhas de pesquisa: Linguagem e ontologia.
Nome: <b>Modelagem Conceitual e Arquiteturas de Sistemas de Informação</b>
Instituição: Universidade Federal do Espírito Santo
Líderes: João Paulo Andrade Almeida e Giancarlo Guizzardi
Linhas de pesquisa: Ontologias e modelos conceituais de referência.