

Da recuperação da informação à recuperação do conhecimento: reflexões e propostas

João de Pontes Junior

Bibliotecário do Museu da Cidade de São Paulo da Secretaria Municipal de Cultura/Prefeitura da Cidade de São Paulo. Bacharel em Biblioteconomia pela Pontifícia Universidade Católica de Campinas. Mestre em Ciência da Informação pelo Programa de Pós-Graduação do Centro de Ciências Humanas e Sociais Aplicadas/Pontifícia Universidade Católica de Campinas, como Bolsista CAPES

Rodrigo Aquino de Carvalho

Professor Efetivo do Curso de Biblioteconomia do Instituto de Ciências Humanas e da Informação da Universidade Federal do Rio Grande. Bacharel em Ciência da Informação com habilitação em Biblioteconomia pela Pontifícia Universidade Católica de Campinas. Mestre em Ciência da Informação pelo Programa de Pós-Graduação do Centro de Ciências Humanas e Sociais Aplicadas/Pontifícia Universidade Católica de Campinas, como Bolsista CAPES

Alexander William Azevedo

Professor Assistente do Departamento de Ciência da Informação (DCI) da Universidade Federal de Pernambuco (UFPE). Bacharel em Ciência da Informação com habilitação em Biblioteconomia pela Pontifícia Universidade Católica de Campinas. Mestre em Ciência da Informação pelo Programa de Pós-Graduação do Centro de Ciências Humanas e Sociais Aplicadas/Pontifícia Universidade Católica de Campinas, como Bolsista CAPES

Apresenta reflexões e propostas da recuperação da informação ao que hoje é chamado de recuperação do conhecimento. Este último se trata de um termo relativamente novo, que, ainda, tem muito a ser discutido e estudado. Sendo assim, o que artigo apresenta é um debate sobre a recuperação da informação, em que explanamos sobre os vários autores que tratam o assunto, bem como sobre algumas ferramentas para a recuperação da informação, como classificação, tesouros,

taxonomias e ontologias, que, também, poderiam ser utilizadas para a recuperação do conhecimento, de forma tão eficiente quanto para recuperação da informação. O que se pretende não é definir o escopo da recuperação do conhecimento, mas lançar uma discussão sobre o tema, a fim de angariar novas contribuições e definições sobre o assunto.

Palavras-chave: *Recuperação da informação; Recuperação do conhecimento; Tesouro; Classificação; Taxonomia; Ontologia.*

Information retrieval to knowledge retrieval: reflections and proposals

Presents ideas and proposals of information retrieval to what is now called the recovery of knowledge. The latter is a relatively new term that still has much to be discussed and studied. Thus, the article presents is that a debate on information retrieval, in which expounded on the various authors treat the subject as well as some tools for information retrieval and classification, thesauri, taxonomies and ontologies, which could also be used for the recovery of knowledge as efficiently as for information retrieval. The intention is not to define the scope of the RC, but to launch a discussion on the subject, in order to raise new contributions and definitions on the subject.

Keywords: *Information etrieval; Knowledge Retrieval; Thesaurus; Classification; Taxonomy; Ontology.*

Recebido em 09.11.2010 Aceito em 06.07.2013

1 Introdução

Quando pensamos na relação entre a informação e o conhecimento, percebemos que isso “se faz na linha contínua que une os pólos da convergência e da divergência” (TALAMO, 2004). Podemos afirmar, portanto, que, em alguns momentos, tais termos são tratados como sinônimos, dependendo da área do conhecimento que esteja lidando com eles. O certo é que, para muitos, estamos na Sociedade da Informação e/ou do Conhecimento, ou seja, esses ativos possuem grande influência nos rumos da sociedade, pois atualmente a denominam, sendo, em vários

setores (educação, saúde, economia, segurança, etc.), as principais balizas de diferenciação e organização.

Sendo assim, esse trabalho objetiva refletir sobre a recuperação da informação (RI) ao que atualmente é conhecido como a recuperação do conhecimento (RC), sob o ponto de vista da Ciência da Informação (CI). A RI já se encontra bem desenvolvida na CI, inclusive essa última, no seu início, já se denominou RI, já a RC se trata de uma proposta nova.

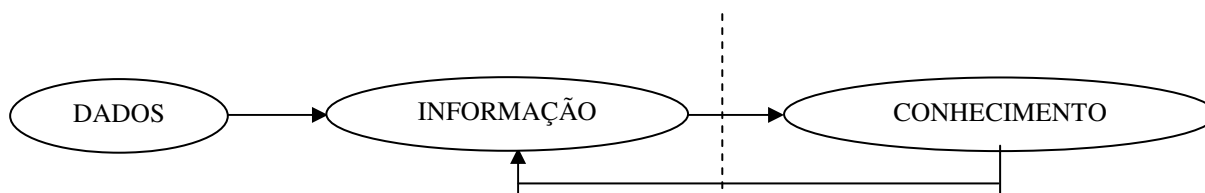
Para a CI, a informação apresenta-se como algo que vem depois do dado. Dados, portanto, são símbolos ou representações sem contexto, segundo Zins (2007), qualificáveis e qualificáveis. A partir de uma contextualização, tratamento e interpretação podem ser transformados em informação.

Barreto (1996, p. 2) define a informação como “estruturas significantes com a competência de gerar conhecimento no indivíduo, em seu grupo ou a sociedade”. Nessa proposta, temos a informação como um elemento anterior e gerador de conhecimento e, esse último, podemos notar que é algo particular de um indivíduo.

Ou seja, podemos entender o conhecimento como a alteração cognitiva de um sujeito em contato com a informação, sendo que essa informação pode ser interpretada e é somada com as vivências e experiências desse indivíduo. Isso é evidenciado por Batista, Costa e Álvares (2007), na interpretação da equação de Brookes (1980), quando dizem que o conhecimento de um indivíduo (vivências e experiências) em contato com a informação gera uma estrutura modificada de conhecimento.

Portanto, parece claro que há dependência da informação em relação ao conhecimento e, vice-versa. Pois, essa última se alimenta da primeira, que, por sua vez, não tem razão de existir, senão para se transformar em conhecimento. A Figura 1 exemplifica essa relação.

Figura 1 – Relação entre dado, informação e conhecimento



Fonte: CARVALHO; PONTES; AZEVEDO.

1.1 Estoques/Fontes (coisificação¹ da informação) Informação absorvida pelo indivíduo

Sendo assim, a partir dos conceitos apresentados e da figura acima, parece evidente que criar meios para a RI não parece algo tão

¹ Termo utilizado em: MIRANDA, A. A ciência da informação e a teoria do conhecimento objetivo: um relacionamento necessário. In: MIRANDA, A. Ciência da informação: teoria e metodologia de uma área em expansão. Brasília: Thesaurus, 2003. p. 173-187.

complicado, uma vez que lidamos com um objeto “no sentido tangível, mensurável, deteriorável física e intrinsecamente, com volume, peso, preço e outras propriedades administráveis” (MIRANDA, 2003, p. 177). Mas, como fazer a RC, haja vista, ser o conhecimento um bem individual? Existem propriedades administráveis para o conhecimento, para que haja sua efetiva recuperação? Esse texto não responde essas questões de forma definitiva e nem satisfatória, mas apresenta uma proposta e alguns pontos que merecem reflexões por parte da CI.

2 Recuperação da informação

A história da comunicação e da transmissão da informação pode ser contada por diferentes marcos/fatos históricos, desde tradição da comunicação oral, escrita alfabética, a imprensa de Gutenberg, o advento das revistas científicas, as comunidades científicas, fatos que alteraram de forma significativa o modo produção social do conhecimento, através da circulação da informação e da sistematização da produção do conhecimento científico (MEADOWS, 1999).

Oliveira (2005, p. 10) afirma que a Revolução Industrial, ocorrida no final do século XIX na Europa e nos Estados Unidos América (EUA), despertou notadamente o crescimento do volume de informações registradas, sendo que “[...] várias tentativas foram feitas para realizar um levantamento bibliográfico universal. A iniciativa mais importante foi assumida pelos advogados belgas Paul Otlet e Henri La Fontaine”.

Segundo Mattelart (2002), Otlet e La Fontaine desenvolveram técnicas e ferramentas de organização e tratamento de uma grande massa de conhecimento registrado, evidenciando seu trabalho de Documentação, pelo fato de prepararem uma classificação universal, com intuito de gerar técnicas que acompanhasse a evolução da explosão documentária ou informação com base no sistema de classificação de Dewey.

Seguindo o raciocínio cronológico, Vannevar Bush, que foi uma autoridade no campo da pesquisa durante a Segunda Guerra Mundial, coordenando mais de seis mil cientistas, em seu artigo publicado em julho de 1945, pela *The Atlantic Review*, "*As We May Think*" (*Como Nós Pensamos*), apontou a necessidade da evolução tecnológica como uma maneira de enfrentar o problema do fluxo de produção científica e da complexidade da recuperação da informação, fato observado anteriormente pelo próprio Paul Otlet.

Portanto, com as questões recorrentes à Segunda Guerra Mundial, que englobam desde as inovações tecnológicas e científicas até a explosão da informação, fez com que a década 1950 ficasse marcada por diversos encontros e conferências técnico-científicos, na qual uma massa crítica de cientistas, engenheiros e empresários começaram a trabalhar com entusiasmo nos problemas expostos por Bush (1945), ou seja, Recuperação da Informação (*Information Retrieval*).

A RI, segundo Saracevic (1992, p. 2), é expressa como “[...] como aquilo que abrange os aspectos intelectuais da descrição da informação e suas especificações para busca e também quaisquer sistemas, técnicas ou máquinas que são utilizados para realizar a operação”, além de ter relacionamento com campos do conhecimento, como psicologia, arquitetura da informação, linguística, semiótica, ciência da informação, ciência da computação, biblioteconomia e estatística.

Face esta questão, a RI estabeleceu-se como um campo do conhecimento científico que objetivou equacionar três problemas enunciados pela explosão da informação: *representação* da informação, *especificação* da busca da informação e *criação* de mecanismo para recuperação (MOOERS, 1951).

Lancaster (1993, p. 202) resgata a ênfase histórica do desenvolvimento dos sistemas de RI em dois tópicos fundamentais: uma delas funciona baseada em termos de indexação tirada de um vocabulário controlado e outra linha da disponibilidade de textos completos, de forma legível, a fim de que o computador realize as buscas por palavras e combinações de palavras nos textos.

Hjørland²(2002) *apud* Robredo (2003), em seu trabalho proferido no ano de 2002, em duas Conferências Internacionais: Information Science and Knowledge Organization (ISKO7) e na Conception of library and Information Science (CoLIS4), aponta que a teoria da recuperação da informação abrange seus estudos e aplicações (ferramentas) em questões como a organização do conhecimento, classificações especializadas, tesouros e as ontologias.

Nesta perspectiva, cabe ressaltar que a RI, como um campo que abrange diversos domínios, desde a Ciência da Informação até Ciência da Computação, possui ferramentas de organização e recuperação da informação e conhecimento, como: classificação, tesouros, taxonomia e ontologias, nos quais é possível estabelecer diferentes estudos teóricos e múltiplas visões do desdobramento de suas aplicações tidas como fundamentais na RI.

² HJØRLAND, B. The methodology of constructing classification scheme: a discussion on the state-of-the-art. In: LÓPEZ-HUERTAS; M. J. (Ed). *In: INTERNATIONAL ISKO CONFERENCE, 7.*, Würzburg, 2002. *Proceedings...* Germany: Ergon Verlag, 2002. p. 450-456. (Advances in Knowledge Organization, v.8).

HJØRLAND, B. Principia Informatica: foundational theory of information and principles of information services. *In: RAYA, F. et al.* (Eds.). *Emerging frameworks an methods. International Conference on Conceptions of Library and information Science, 4.*, Seattle, 2002. *Proceedings...* Greenwood Village, Colorado: Libraries Unlimited, 2002. p. 109-121.

2.1 Tesouros

O tesouro organiza as informações através da denominada linguagem documentária, trata-se da relação de um termo entre si mesmo. O tesouro é utilizado para a indexação de documentos normalmente de assuntos e áreas específicas. Em um primeiro momento, aparecem os catálogos impressos e manuais, após a automação das bases de dados, os catálogos *online*.

Para definir melhor a função dos tesouros, Currás³ (1995) *apud* Moreira, Alvarenga e Oliveira (2004) nos diz que tesouro é “[...] uma linguagem especializada, normalizada, pós-coordenada, usada com fins documentários, nos quais os elementos linguísticos que compõem – termos simples ou compostos – encontram-se relacionados, entre si, sintática e semanticamente”.

Normalmente os tesouros possuem uma base estruturada e hierárquica, conforme exemplo no quadro abaixo. Sua construção é feita de forma hierarquizada e os termos baseados em conhecimento específicos. Souto (2003) acrescenta, ainda, a importância que tem a aplicação dos tesouros como linguagem controlada nas possíveis inconsistências na recuperação da informação em bases de dados. O autor nos apresenta, também, a estrutura de um tesouro.

Quadro 1 – Estrutura de um tesouro

Termo	Biblioteconomia
Nota de aplicação	Termo incluído em 18/5/01
Equivalência	UP Documentação
Termo genérico	TG Ciência da Informação
Termo específico	TE Bibliotecário
Termo associado	TA Arquivística

Fonte: SOUTO (2003).

Moreira, Alvarenga e Oliveira (2004) apresentam uma definição feita pela Organização Nacional de Padrões da Informação (NISO), em um documento dito como as Instruções (normas) para a construção, formato e gerenciamento do tesouro monolíngua, pelo qual o tesouro é definido como:

[...] vocabulário controlado e arranjado em uma ordem conhecida e estruturada que sua equivalência, homografia, hierarquia e relação associativa entre termos são demonstradas claramente e identificada por um relacionamento padronizado de indicadores que são empregados reciprocamente.

Zhonghong, Chaudhry e Sattar (2006) afirmam, ainda, que os “tesouros foram criados no ambiente *online* e usados para indexar a matéria dos assuntos dos documentos. Eles também podem ser usados para ajudar nas buscas dos usuários”.

³ CURRÁS, E. *Tesaurus, linguagens terminológicas*. Brasília: IBICT, 1995.

A definição de que os tesouros foram criados em ambientes *online*, contrapõem com outros autores, que descrevem em suas pesquisas a criação dos tesouros bem antes da utilização em ambientes *online*. Peter Mar Roget cria o termo em 1852, o que, na verdade, era chamado de dicionário de tesouros pelo próprio autor. Roget estabelece um denominador comum para um tipo de vocabulário que relata seus próprios termos, usando certo tipo de relacionamento semântico. Em 1951, Mortimer Taube, com a ajuda de Lancaster, cria o sistema unitermo, evoluindo-se como o primeiro tesouro em 1959, concebido pelo Centro de Engenharia da Informação de DuPont (MOREIRA; ALVARENGA; OLIVEIRA, 2004).

Neste mesmo artigo, os autores fazem, ainda, uma diferenciação entre taxonomias, classificação e tesouros, deixando clara a posição de cada um dos modelos de classificação, na organização da informação e do conhecimento. Segundo os autores, os tesouros foram concebidos para a organização de informação disponibilizada na rede, como informação que possua um assunto específico para que possam ter seu vocabulário controlado (ZHONGHONG; CHAUDHRY; SATTAR, 2006).

Os tesouros são tão importantes como meios de recuperação de informação, principalmente na *Web*. Vogel [2006?] coloca a possibilidade de definirmos as taxonomias a serem utilizadas em portal corporativo, utilizando tesouros para definir um vocabulário controlado, podendo ser desde uma lista de sinônimos como até mesmo uma linguagem documentária ou informacional, permitindo hierarquias propostas pela taxonomia e o tesouro.

2.2 Classificação

Os sistemas de classificação foram um dos primeiros modelos de organização da informação e do conhecimento que sem tem registro. A história dos sistemas de classificação bibliográfica aparece logo após a segunda guerra no século XIX, quando os estudiosos e cientistas se viram em meio a uma "babel" de documentos, resultados de pesquisas, relatórios, entre outros, sem organização alguma e, neste momento, começam a surgir excelentes exemplos de sistemas que organizassem estes documentos, como o de Cutter, a Classificação Universal de Dewey (CDD) e o Código de Classificação Universal (CDU) (BARBOSA, 1969).

Para Piedade (1983), o homem está sempre fazendo algum tipo de classificação em sua vida, seja em coisas ou ideias, para que, inclusive, possa entender e organizar melhor o espaço em que vive. Diz, ainda, que classificar é o ato de dividir em classes ou grupos, segundo diferenças e semelhanças, seguindo critérios e disposição em grupos, utilizando para isto uma metodologia específica.

Por ser um sistema de classificação hierárquico e estruturado, o sistema de classificação se assemelha a outros tipos de organização do

conhecimento, como, por exemplo, os tesouros e taxonomias. Mai⁴ (2004, p. 93) *apud* Zhonghong, Chaudhry e Sattar (2006, p. 6) destaca que os sistemas de classificação bibliográfica estão "intimamente ligados à comunidade científica e a seu discurso". Os autores dizem, ainda, que "[...] os sistemas de classificação e tesouros estão ampla e essencialmente ligados ao ambiente embasado em documentos e mais obrigatório dentro da comunidade acadêmica, embora eles agora tenham sido também expandidos para o ambiente *Web*".

Segundo Zhonghong, Chaudhry e Sattar (2006), a utilização dos sistemas de classificação é praticamente restrita aos meios acadêmicos e bibliotecas, já que, para ser utilizado na *Web*, se tornaria inviável. Os autores nos indicam, também, que, para o ambiente *Web*, o mais correto seria a atualização de tesouros e taxonomias, os quais serão apresentados individualmente nesta seção, com suas peculiaridades.

O Quadro 2 mostra algumas classificações desenvolvidas para utilização em bibliotecas e centros de documentações, visando à recuperação de documentos.

Quadro 2 – Sistemas de classificação

SISTEMAS DE CLASSIFICAÇÃO	ESTRUTURA
<p>CLASSIFICAÇÃO DECIMAL DE DEWEY-CDD</p> <p>Criador: Melvil Dewey</p> <p>Ano de criação: 1876</p>	<p>Compreende em 9 grandes classes de assuntos, a notação é construída de números decimais e foi usado um mínimo de três algarismos, tais como 100 e 200. O único sinal gráfico utilizado é um ponto após o terceiro algarismo, sendo que o único valor deste sinal é facilitar a leitura.</p>
<p>CLASSIFICAÇÃO DECIMAL UNIVERSAL-CDU</p> <p>Criadores: Paul Otlet e Henri LaFontaine</p> <p>Ano de criação: 1892</p>	<p>Esta classificação é baseada na CDD, havendo um detalhamento das subdivisões dos assuntos e com notação mista com sinais gráficos, letras ou palavras o que representa os conceitos na classificação e expressa sua ordenação. Assim como a CDD a CDU também possui 9 grandes classes do conhecimento.</p>
<p>CLASSIFICAÇÃO DE CUTTER</p> <p>Criador: Charles Ammi Cutter</p> <p>Ano de criação: 1876</p>	<p>A classificação expansiva é mista, constituída de letras maiúsculas, algarismos arábicos e ponto.</p>
<p>CLASSIFICAÇÃO BIBLIOTECA DO CONGRESSO</p> <p>Criador: a própria Biblioteca</p> <p>Ano de criação: 1800</p>	<p>Classificação mista, constituída de letras maiúsculas, algarismos arábicos, de 1 a 9999, em ordem aritmética, e um sinal gráfico o ponto. O conhecimento está dividido em 20 classes, indicados por letras maiúsculas. Classificação que permite expansão.</p>
<p>CLASSIFICAÇÃO DE BROWN</p> <p>Criador: James Duff Brown e John Henry Quinn</p> <p>Ano da criação: 1894</p>	<p>Como outras este modelo de classificação possui notação mista, constituindo-se de letras de A a X, algarismos na ordem aritmética, letras minúsculas e sinais gráficos. As classes são divididas em 11 grandes classes, algumas vezes representando as ciências como subdisciplinas outras vezes grandes divisões destas.</p>
<p>CLASSIFICAÇÃO DE BLISS</p> <p>Criador: Henry Evelyn Bliss</p> <p>Ano da criação: 1908</p>	<p>O autor dividiu o conhecimento em quatro grandes classes, que fazem saber: Filosofia Ciência, História e Tecnologia e Artes. Seu sistema e notação eram mistos, constituídos por letras maiúsculas e minúsculas, algarismos arábicos e sinais gráficos. Para Bliss, sua notação não era inteiramente facetada, mas permitia a ampliação de síntese.</p>
<p>CLASSIFICAÇÃO DE RANGANATHAN</p> <p>Criador: Shiyali Ramamrita Ranganathan</p> <p>Ano da criação: 1924</p>	<p>A notação da classificação de dois pontos é mista, utilizando algarismos arábicos, letras minúsculas e maiúsculas, letras gregas e sinais gráficos, somando cerca de 70 caracteres. O conhecimento humano é dividido em 41 classes principais, representadas em textos de Ranganathan pela sigla (MC).</p>

Fonte: Adaptação de PIEDADE (1983).

⁴ MAI, J. Classification of the web: challenges and inquieres. *Knowledge organization*, v. 31, p. 92-96. 2004.

2.3 Taxonomias

Os cientistas se veem às voltas, cada vez mais, com o grande número de informações que se encontram disponíveis na *Web* e, ao mesmo tempo, com a preocupação de como manipular esta informação, em como controlar e arquitetar sua estrutura, facilitando a recuperação de informações por parte do usuário.

Oriunda das ciências da vida, mais especificamente da Biologia, no princípio, utilizada para a classificação de seres vivos, a taxonomia veio para o meio de digital com o intuito de cumprir sua parte na organização de informações, disponíveis em portais corporativos das organizações e em portais de busca que possuem diretórios, como, por exemplo, o Yahoo, também por possibilitar filtros de pesquisa, classificação automática, entre outros. Isso ocorre por serem as taxonomias organizações mais específicas que possuem como elementos-chave uma estrutura hierárquica baseada em etiquetas

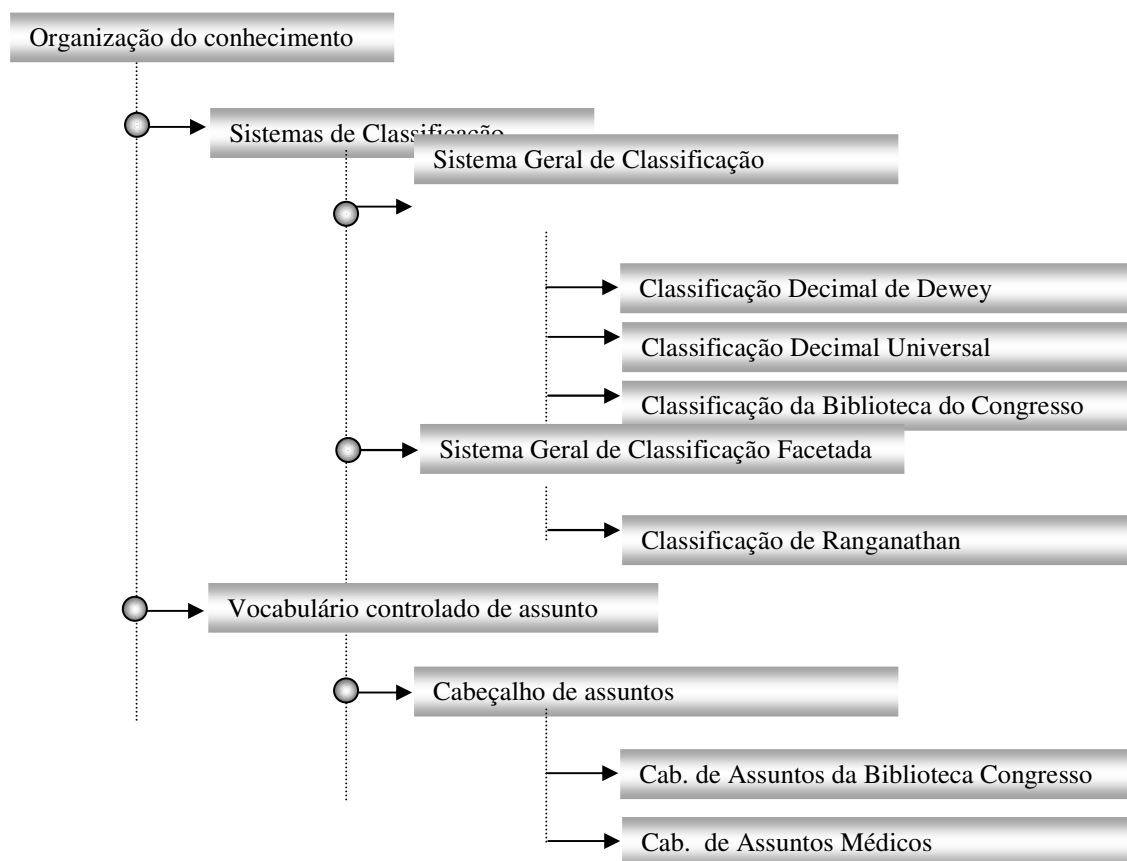
Atualmente, as taxonomias são ferramentas essenciais para a arquitetura da informação, mesmo havendo, ainda, um falta de clareza sobre seu papel e intenção, pois podem referir-se a coisas diferentes de diferentes contextos (ZHONGHONG; CHAUDHRY; SATTAR, 2006).

Apesar das taxonomias serem um sistema de classificação de conceitos para ambiente *Web*, não podemos descartar as diferenças que esse ambiente possui em relação aos outros sistemas, como tesouros, a própria classificação e as ontologias. Talvez, o que mais se aproxime e, em alguns momentos, sendo utilizado para definição de termos controlados, seja o desejo do arquiteto de redes é o tesouro.

Concoran⁵ (2002) *apud* Zhonghong, Chaudhry e Sattar, (2006) argumenta sobre o poder de recuperação e busca de informações em ambientes *Web*, sendo esta uma das vantagens que possui as taxonomias.

⁵ CORCORAN, M. *Industry insights: taxonomies: hope or hype?* 2002. Disponível em: <<http://www.ebscohost.com/thisTopic.php?marketID=1&topicID=1>>. Acesso em: 6 out. 2005

Figura 2 – Exemplo de taxonomia unidimensional



Fonte: Adaptação de ZHONGHONG; CHAUDHRY; SATTAR (2006).

2.4 Ontologias

Com a eterna preocupação de que organização da informação disponível no ambiente *Web* seja recuperada com maior precisão e que o usuário não se perca nas buscas, se faz, cada vez mais necessário, a criação de ferramentas que busquem e tratem a informação, de maneira eficaz. Pensando nisto, as ontologias surgem em alguns sistemas que trazem a informação como foco principal, estando estes disponíveis, em sua grande maioria, na Internet, sendo representadas, hierarquicamente, através de palavras-chave e conceitos. Esta hierarquia de conceitos pode ser vista, inclusive, como uma taxonomia (FREITAS, 2006?).

Moreira, Alvarenga e Oliveira (2004) dizem que as ontologias podem ser utilizadas em diversas aplicações, como se fossem um sistema multiagente, fazendo com que estes agentes troquem informações entre si, ou seja, fazendo o que podemos também chamar de interoperabilidade de sistemas. Os autores mencionam ainda alguns problemas que podem se beneficiar das ontologias os quais se fazem saber: a representação do conhecimento, o reuso do conhecimento, o compartilhamento do conhecimento, a aquisição e integração do conhecimento; o processo de linguagem natural; a tradução automática; a interoperabilidade entre sistemas, agentes companhias ou pessoas.

Segundo Moreira, Alvarenga e Oliveira (2004) as ontologias possuem quatro principais tipos de definições: como um *sistema conceitual informal*, no qual está abaixo da base de conhecimento particular, o que, para esse caso, a ontologia pertence ao nível conceitual e não ao nível simbólico; como uma *representação de um sistema conceitual* via uma teoria lógica, que define a ontologia como um tipo de base especial do conhecimento (nível simbólico) e conhecimento independente de um estado de domínio particular; como um *vocabulário usado por uma teoria lógica*, neste caso o nível de formalização desse vocabulário pode variar de uma ontologia a outra; e como uma *especificação de conceitualização*, fazendo com que uma conceitualização possa ser entendida como um conjunto de comitês ontológicos (nível simbólico), a última interpretação é que faz com que o número de seguidores pertencentes à comunidade de Inteligência Artificial seja significativo.

Assim como nas taxonomias, os tesouros estão muito próximos das ontologias e podem até operar no mesmo nível, porém, apresentam diferentes propostas: a primeira é direcionar para o domínio do registro de conceitos, objetivando as deduções automatizadas, enquanto que a segunda é direcionada para a comunicação entre o usuário e a linguagem documentária.

No intuito de apresentar alguma solução para que a organização da informação e do conhecimento e, também, de forma que a CI se aproprie destes conceitos efetivamente, de maneira que o usuário inicial ou final saiba o que ele realmente está usando para recuperar ou arquitetar sua informação, após a apresentação e definição de algumas ferramentas utilizadas principalmente na *Web*, na arquitetura, organização e recuperação da informação e de maneira a ficar claro os focos e aplicações de cada uma delas, bem como suas diferenças, segue uma tabela, a qual, de forma resumida, apresenta as especificidades de cada um dos instrumentos que facilitam a RI apresentados.

Quadro 3 – Diferenças entre classificação, tesouros, taxonomias e ontologias

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

Fonte: Adaptação de ZHONGHONG; CHAUDHRY; SATTAR (2006).

3 Recuperação do conhecimento

Abordar as principais definições relacionadas à organização do conhecimento, tais como: estrutura do conhecimento, domínio do conhecimento, conceitos; com o objetivo de estabelecer uma base teórica que permita a discussão e o entendimento sobre recuperação do conhecimento.

A estrutura do conhecimento, segundo Anderson⁶ (1992; 1997) *apud* Tsai; Chou (2005), é a forma na qual um indivíduo constitui seus conhecimentos, seguindo princípios organizacionais derivados de experiências prévias existentes.

Para este trabalho, é considerado que a estrutura de conhecimento pode ser correlacionada com redes ou esquemas neurais, que, provavelmente, não obedecem a estruturas hierárquicas. Logo após, o termo conceito, que pode ser visto como o esquema que representa esta forma de estruturação do conhecimento (Figura 3).

Para Tsai e Chou (2005), o termo conceito “refere-se à idéias, pensamentos, partes de conhecimento e proposições”. Concordando com essa afirmação, Alvarenga (2003), define conceitos como “unidades de conhecimento” e complementa, enfatizando sua importância, no momento em que ressalta que “são elementos chave para os atos de representação, comunicação e preservação de conhecimentos”.

Shera⁷ (*apud* ALVARENGA, 2003) afirma que cada nova experiência ou sensação seria classificada pelo cérebro de acordo com um acervo de experiências passadas, diminuindo a incerteza, determinando o processo de aprendizagem do indivíduo.

Correlacionando a afirmativa da autora com os conceitos anteriores, as novas experiências seriam bases para novos conceitos, que estariam, por sua vez, organizados (classificados) dentro de uma estrutura de conhecimento, sendo esta originária de experiências prévias do indivíduo.

Segundo Tsai e Chou (2005), a estrutura de conhecimento de cada indivíduo está organizada ao redor de um conceito central, que está fortemente ligado a outros conceitos dentro de um domínio de conhecimento.

O termo domínio, definido no dicionário Aurélio século XXI (FERREIRA, 1999), é descrito como o “Âmbito de uma arte ou ciência”, ou seja, é a área ou o escopo que uma determinada arte ou ciência tem por foco de atividade.

A conceituação de domínio de conhecimento está, segundo Campos (2004), intimamente ligada ao termo conceito, “como é possível perceber, o entendimento do conceito está refém de um contexto que é classificado

⁶ ANDERSON, O. R. Some interrelationships between constructivist models of learning and current neurobiological theory, with implications for science education. *Journal of research science teaching*, v. 29, p.1037-58, 1992.

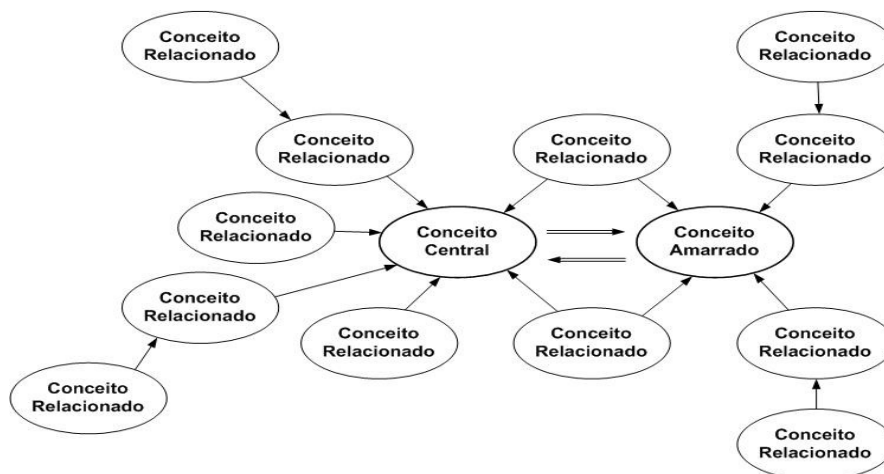
ANDERSON, O. R. 1997. A neurocognitive perspective on current learning theory and science instructional strategies. *Science education*, v. 81, p. 67-89, 1997.

⁷ SHERA, J. H. Pattern, structure and conceptualization in classification. In: International Study Conference on Classification for Information Retrieval, 1957, Dorking. *Proceedings...* London: ASLIB, 1957.

por conta dos [outros] conceitos que nele estão inseridos”, dessa forma, o entendimento de um conceito está restrito a seu contexto, ou seja, a seu domínio de conhecimento.

Além da suposição da existência de um conceito central presente na estrutura de conhecimento de um indivíduo, Tsai e Chou (2005) consideram a existência de um conceito amarrado, ou seja, um conceito que também está ligado a vários outros, mais de forma menos acentuada que o central. A figura 3 ilustra o exposto acima.

Figura 3 – Um modelo de estrutura de conhecimento



Fonte: TSAI; CHOU (2005).

A recuperação do conhecimento é realizada por meio do uso de conceitos centrais e amarrada pode ajudar a recuperar todos os conceitos relacionados a um domínio de conhecimento (TSAI; CHOU, 2005).

Isto posto, a recuperação do conhecimento pode ser entendida como o processo realizado por um indivíduo, dentro de sua esfera privada, com o objetivo de lembrança de um dado conceito, previamente assimilado e associado a sua estrutura de conhecimentos, por meio de métodos que promovam a recuperação eficiente e efetiva destes conceitos, dentro de um domínio de conhecimento.

O estudo da forma com que o conhecimento é organizado pode promover um entendimento melhor das estruturas envolvidas no processo de recuperação de um conhecimento prévio, isto pode permitir que se estabelecessem possíveis correlações entre os processos de recuperação do conhecimento com os de recuperação da informação, objetivando a melhoria do desempenho desses últimos.

4 Considerações finais

O sistema de classificação proposto tanto por Dewey quanto por Otlet, apresentam uma forma essencialmente hierárquica de organização da informação, que está em divergência com o modo na qual se dá a organização do conhecimento, segundo a opinião dos autores analisados neste trabalho.

A estrutura do conhecimento aqui analisada sinaliza para métodos de RI que possam ser mais flexíveis, como é o caso da taxonomia, e, ainda, que possam conter relações entre os conceitos, ou seja, que considerem as ligações (nós ou links) entre os conceitos. Em uma reflexão, um possível caminho seria considerar que esses links poderiam ser estabelecidos pelo próprio usuário, ficando registrados na rede; dessa forma, seria possível a representação de uma estrutura de conhecimentos confeccionada por múltiplos usuários, que, por sua vez, possibilitaria a detecção de conceitos centrais e amarrados, o que vai de encontro com a proposta de Vannevar Bush, em que pese o fato da tecnologia se apresentar como a extensão do pensamento humano.

A identificação de tais conceitos, dentro de um domínio de conhecimento, poderia contribuir tanto para os processos de RC quanto de RI, apontando, assim, para uma nova forma de organização da informação, a partir da análise das estruturas do conhecimento. Sendo assim, pensando principalmente na era digital, há um destaque para as taxonomias e ontologias, que facilitam a manipulação da informação para usuários e sistemas, respectivamente. Podemos afirmar que isso pode facilitar a RC.

Parece claro que há dependência da informação em relação ao conhecimento e vice-versa, como já afirmado anteriormente, ao longo do texto. Não é possível fazer uma análise que isole tais conceitos, pois o conhecimento se alimenta da informação, que, por sua vez, não tem razão de existir, senão para se transformar em conhecimento.

Portanto, o trabalho levantou questões pertinentes a RC com relação a RI, tentando desvendar possíveis elos e conceitos que pudessem defini-las. Ocorre que pouca literatura há sobre o assunto, devido, talvez, à subjetividade do conhecimento no entendimento da CI e a difícil identificação de propriedades administráveis desse ativo.

A partir das reflexões e propostas expostas neste trabalho, espera-se que novas pesquisas sejam realizadas sobre a temática, visando ampliar as investigações sobre conceitos e aplicações da RI em relação à RC, além do que, espera-se que futuras pesquisas focalizem outros contextos, com novas questões nas estruturas do conhecimento.

Referências

- ALVARENGA, L. Representação do conhecimento na perspectiva da ciência da informação em tempo e espaço digitais. *Encontros Bibli: Revista Eletr. Biblioteconomia Ciência Informação*, Florianópolis, n. 15, 1º sem., 2003. Disponível em: <<http://www.periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2003v8n15p18/5233>>. Acesso em: 3 dez. 2007.
- BARBOSA, A. P. *Teoria e prática dos sistemas de classificação bibliográfica*. Rio de Janeiro: IBICT, 1969.

BARRETO, A. de A. A eficiência técnica e econômica e a viabilidade de produtos e serviços de informação. *Ciência da Informação*, Brasília, v. 25, n. 3, 1-18, 1996.

BUSH, V. As we may think. *Atlantic Monthly*, v. 176, n. 1, p. 101-108, 1945. Disponível em: <http://www.theatlantic.com/unbound/flashbks/computer/bushf.htm>. Acesso em: 2 dez. 2007

BATISTA, F. F.; COSTA, S. M. de S.; ALVARES, L. Gestão do conhecimento: a realização da proposta de Brookes para a Ciência da Informação? In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB, 8., 2007, Salvador, BA. *Anais...* Salvador, 2007.

BROOKES, Bertram. The foundations of information science. Part I. Philosophical aspects. *Journal of Information Science*, v. 2, p. 125-133, 1980.

CAMPOS, M. L. de A. Modelização de domínios de conhecimento: uma investigação de princípios fundamentais. *Ciência da Informação*, Brasília, v. 33, n. 1, p. 22-32, 2004.

FERREIRA, A. B. de H. *Novo aurélio século XXI: o dicionário da língua portuguesa*. 3. ed., 4. imp., rev. e ampl. Rio de Janeiro: Nova Fronteira, 1999.

FREITAS, F. L. G. de. *Ontologias e a web semântica*. 2006?. Disponível em: <http://www.inf.unisinos.br/~renata/cursos/topicosv/ontologias-ws.pdf>. Acesso em: 3 dez. 2007

LANCASTER, F. W. *Indexação e resumos: teoria e prática*. Brasília: Briquet de Lemos/Livros, 1993.

MATTELART, A. *História da sociedade da informação*. São Paulo: Loyola, 2002.

MEADOWS, A. J. *A comunicação científica*. Brasília: Briquet de Lemos, 1999.

MIRANDA, A. A ciência da informação e a teoria do conhecimento objetivo: um relacionamento necessário. In: MIRANDA, A. *Ciência da informação: teoria e metodologia de uma área em expansão*. Brasília: Thesaurus, 2003. p.173-187.

MOOERS, C. Zatocoding applied to mechanical organization of knowledge. *American Documentation*, v. 2, p. 20-32, 1951.

MOREIRA, A.; ALVARENGA, L.; OLIVEIRA, A. de P. Thesaurus and ontology: a study of the definitions found in the Computer and Information Science Literature, by means of an analytical-synthetic method. *Knowledge Organization*, v. 31, n. 4, p. 231-244, 2004.

OLIVEIRA, M. de. Origem e evolução da ciência da informação. In: OLIVEIRA, M. de. *Ciência da Informação e Biblioteconomia: novos*

conteúdos e espaços de atuações. Belo Horizonte: Editora UFMG, 2005. p. 9-28.

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

SARACEVIC, T. Information science: origin, evolution and relations. In: VAKKARI, P.; CRONIN, B. (Eds). *Conceptions of library and Information Science*. COLIS Conference Tampere, 1991, Finland. *Proceedings...* Los Angeles: Taylor Graham, 1992.

SOUTO, L. F. Recuperação de informações em bases de dados: usos de tesouro. *Transinformação*, Campinas, v. 15, n. 1, p. 73-81, jan./abr. 2003.

TÁLAMO, M. de F. G. M. A pesquisa: recepção da informação e produção do conhecimento. *DataGramZero*, v. 5, n. 2, abr. 2004. Disponível em: <http://www.datagramazero.org.br/abr04/Art_01.htm>. Acesso em: 10 dez. 2007.

TSAI C. C.; CHOU Y. J. The role of "core" and "anchored" concepts in knowledge recall: a study of knowledge organization of learning thermal physics. *Knowledge organization*, v. 32, n. 4, p. 143-158, 2005.

VOGEL. M. J. M. *Taxonomias: alguns conceitos e algumas confusões*. [2006?]. Disponível em: <<http://portais.integra.com.br/sites/terraforum/Biblioteca/libdoc00000124v002>

[Taxonomia_%20conceitos_confusoes.pdf](#)>. Acesso em: 30 nov. 2007.

ZHONGHONG, W.; CHAUDHRY, A. S.; KHOO, C. Potential and prospects of taxonomies for content organization. *Knowledge Organization*, v. 33, n. 3, p. 160-169, 2006.

ZINS, C. Conceptual approaches for defining data, information, and knowledge. *Journal of the American Society for Information Science and Technology*, v. 58, n. 4, p. 479-493, 2007. Disponível em: <<http://www.success.co.il/is/dik.html>>. Acesso em: 3 set. 2007.